A Citizens-Editors Model of News Media

Sobbrio, Francesco

IMT Institute For Advanced Studies, Lucca

October 2009

A Citizens-Editors Model of News Media*

Francesco Sobbrio†

October 19, 2009

Abstract

We model a market for news where profit maximizing media outlets choose their editors from a population of rational citizens. We show that when information acquisition is costly, liberal (conservative) citizens find optimal to acquire information from a media outlet having a liberal (conservative) editor. Consequently, we show that depending on the distribution of citizens' ideological preferences, a media outlet may choose to hire a non-moderate editor even in a monopolistic market. Moreover, the higher the degree of competition in the market for news, the more likely that media outlets will hire non-moderate editors. Finally, less moderate editors are more likely to be hired in a news market where the opportunity cost of acquiring information for citizens is low.

JEL Classification: D72, D81, D83

Key Words: Media Bias, Information Acquisition, Valence, Competition

*I am very grateful to Juan Carrillo and Andrea Mattozzi for many useful suggestions. I would also like to thank Francesco De Sinopoli, Mark Dincecco, Paolo Pin, Guofu Tan, Edilio Valentini, Simon Wilkie, Kenneth Wilburn and seminar audiences at the 7th Workshop in Media Economics, 2009 European Meeting of the Econometrics Society, the 2009 Meeting of the Association for Public Economic Theory, the 8th Journées Louis-André Gérard-Varet, the 2009 Meeting of the European Public Choice Society, University of Bologna, University of Padova and IMT Lucca. All remaining errors are mine.

†Department of Economics, IMT Lucca Institute for Advanced Studies, Piazza San Ponziano 6, 55100, Lucca, Italy. Email: f.sobbrio@imtlucca.it

1 Introduction

In regulating the market for news in the US, the Federal Communication Commission pursues three strategic policy goals: competition, diversity and localism.[1] Despite the self-evident importance for democratic decision-making of fostering an efficient market of information, such policy goals still lack a sound theoretical foundation and an analysis of their consequences on consumers' welfare and on the optimal media ownership's rules.

This paper is the first to show the presence of a direct link between competition and diversity in a market for news where consumers are rational (i.e., they do not derive any exogenous utility from receiving biased information), they share the same prior beliefs and media outlets are just profit-maximizers. More specifically, our analysis suggests that a higher degree of competition leads to more viewpoint diversity in the form of having different media outlets hiring editors with different ideological preferences.

We model the market for news as driven by the demand for information of citizens. More specifically, citizens have to choose between two alternative candidates (or policies). Citizens differ in their idiosyncratic preferences but they all equally value the *valence* (i.e., quality) of alternative candidates (or public benefit of alternative policies). Citizens may acquire some information about the quality of different candidates by watching news reports. News reports are produced by editors hired by media outlets from the populations of citizens. That is, once hired by a media outlet, a citizen-editor can gather (costly) information about the candidates' quality and, then, report it to the viewers.

We show that editors with different idiosyncratic preferences have different optimal information acquisition strategies. A moderate editor (i.e., one who is *ex-ante* indifferent between the two candidates) uses a "balanced" information acquisition strategy. The amount of evidence in support of the leftist candidate that she requires in order to stop collecting information and produce a report in favor of such candidate is the same as the one she requires to produce a report in favor of the rightist candidate. Instead, a non-moderate editor acquires information in a "slanted" way. That is, a small amount of evidence in support of the leftist candidate is sufficient to induce a leftist editor to stop investing in information acquisition and produce a report in favor of that candidate. On the other hand, such editor would produce a report in favor of the rightists candidate only after having collected a large amount of evidence in support of that candidate. Moreover, the more moderate an editor is the more information, on average, she collects.

In order to access news reports, citizens have to pay an opportunity cost. Hence, in choosing whether to watch a media outlet report or not and, if so, which of them to watch, a citizen will take into account two different components. She will consider how

[1] Source: http://www.fcc.gov/mediagoals/

much information the editor of a media outlet may have collected before producing a news report. At the same time, she will also take into account how *pivotal* the information gathered by an editor could be for her final choice. Suppose, for example, that a liberal citizen has to decide whether to watch a media outlet having a moderate editor or one having a liberal editor. This type of citizen knows that the moderate editor is the one who, on average, will produce a more informative report (i.e., the average amount of evidence contained in a report is higher). However, such citizen also knows that the liberal editor is the one who may be more likely to produce a report that will be relevant for her final decision. Indeed, a liberal citizen will vote for a conservative candidate only when the evidence in favor of such candidate is very strong. A liberal editor is the one who will collect more evidence in support of the conservative candidate before producing a favorable report. Instead, a report coming from a moderate editor in favor of the conservative candidate may not contain enough evidence to convince a liberal citizen to change her *ex-ante* ranking of preferences over candidates. Hence, a liberal citizen may find optimal to acquire information from a media outlet having an editor with similar idiosyncratic preferences even though such citizen does not have any exogenous preferences for like-minded sources of information.

Media outlets anticipate this behavior by citizens and hence they choose which editor to hire taking into account the expected demand for news reports produced by editors with different idiosyncratic preferences. That is, by choosing a more leftist, moderate or rightist editor, media outlets implicitly choose their "product" location in the political space. We show that when the distribution of citizens is such that the number of extremists citizens is higher than the one of moderate citizens, a media outlet may choose to hire a non-moderate editor even in a monopolistic market. Hence, even though citizens do not derive any exogenous utility from acquiring biased information and the media outlet is just maximizing profits, the endogenous acquisition of costly information may induce a media outlet to choose an editor whose optimal information acquisition strategy is "slanted" in favor of the alternative *ex-ante* preferred by a subset of citizens. This is true even in the case where all citizens share the same *ex-post* ranking of preferences over candidates.

We also show that even in the case where citizens are uniformly distributed in the policy space, it exists a threshold in the number of media outlets present in the market for news above which media outlets hire non-moderate editors. More specifically, the lower the opportunity cost of watching news by citizens, the more citizens care about candidates' quality and the lower is the cost of acquiring information by editors, the more likely that media outlets hire non-moderate editors, for a given number of media outlets present in the market for news.

Overall, our results suggest that we should expect more moderate editors to prevail in news markets where the opportunity cost that citizens have to incur to access information is high. Indeed, when such opportunity cost is high, the expected benefit of watching news reports for extremist citizens may be lower than the cost. Hence, media outlets will be more likely to hire moderate editors since the bulk of the demand for news comes from moderate citizens. Instead, when the opportunity cost is low, even extremist citizens may find convenient to watch news reports when such news reports come from an editor with similar idiosyncratic preferences. Hence, a media outlet may find optimal to "locate" its news product to capture this demand for news by non-moderate citizens (i.e., hire a non-moderate editor). A clear application of such result is represented by the market for news in the broadcast media sector with respect to the press. The opportunity cost of watching a report from a broadcast media is arguably lower than the one of reading a newspaper. Our analysis thus suggests that we should expect to find more moderate editors in the press than in the broadcast media sector. At the same time, we should expect more extremist citizens watching broadcast media and a higher overall demand for broadcast media with respect to the one faced by the press.

1.1 Related Literature

A recent empirical literature has shown the presence of systematic bias in the market for news using a variety of instruments to measure such bias (e.g., Groseclose and Milyo 2005, Ho and Quinn 2008, Gentzkow and Shapiro 2009). In parallel, a fast growing theoretical literature has tried to rationalize the presence of such systematic bias in the media. This literature has, so far, identified two different forces creating a bias in media reports. The first one is a "supply-driven" bias: media bias may be derive from the idiosyncratic preferences of journalists (Baron 2006), owners (Djankov et al. 2003, Anderson and McLaren 2007), governments (Besley and Prat 2006) or advertisers (Ellman and Germano 2008). The second one is a "demand-driven" bias. Part of this literature assumes that consumers like to receive information confirming their bias and thus media just reflect and confirm the bias of their audience (Mullainathan and Shleifer 2005, Bernhardt et al. 2008). On the other hand, Gentzkow and Shapiro (2006) show that even when consumers do not like biased information, if media outlets have reputation concerns and there is uncertainty on the quality of media outlets, in presence of heterogeneous prior beliefs different media outlets operating in the same market may find optimal to bias their reports according to the prior beliefs of different segments of consumers.[2] Finally, Chan and Suen (2008) show that media bias emerges when media outlets observe the

[2]See also Burke (2008) for a model of media bias with no reputation concerns and with heterogeneous prior beliefs.

state of the world but they are exogenously constrained to report coarse information.

Our model provides a demand-driven rationale for media bias without relying on any exogenous preferences for biased news confirming individuals' beliefs (as in Mullainathan and Shleifer 2005), without heterogeneous prior beliefs (as in Gentzkow and Shapiro 2006 and Burke 2008) and without exogenous coarsening of information (as in Chan and Suen 2008). In our model, the only force behind the individual willingness to acquire information from a like-minded source is the cost of acquiring information. Our results are consistent with the empirical results of Gentzkow and Shapiro (2009). Using zip-code level data on newspaper circulation in the US, they show that the demand for right-wing newspaper is higher in markets with a higher proportion of Republicans. Moreover, they find that ownership has little or no role in media slant.[3] Our paper suggests that such findings may not be the result of behavioral preferences for biased news but they may rather be the result of the demand for costly information by rational individuals and the consequent optimal "ideological location" of news by profit maximizing media outlets.[4]

Formally, our model of optimal acquisition of information by citizen-editors is related with the one of Brocas and Carrillo (2008) on systematic errors in decision making. In their setting individuals have to decide how much information they want to collect before taking an action whose utility depends on the state of the world. Given any exogenous amount of information, all individuals would choose the same action. However, in presence of endogenous information acquisition different individuals would have different probabilities of choosing a given action. More specifically, they show that individuals favor actions with large payoff-variance. Our setting differs in that we assume that all actions have the same variance in payoffs for any citizen-editor and such variance is equal across citizen-editors. Moreover, in our model citizen-editors differ in their *ex-ante* ranking of actions even when they share the same *ex-post* ordinal preferences over actions.[5]

Our paper is also related to Suen (2004) on the self-perpetuation of biased beliefs. Suen focuses on a situation where information acquisition is not costly but the presence of heterogeneous subjective beliefs and coarse information lead to a "short-run" polarization of beliefs. Instead, our setting involves a situation where information is not coarse, people share the same subjective beliefs but the presence of a cost in information gathering and heterogeneous idiosyncratic preferences may lead to a "long-run" polarization of beliefs.

The paper is organized as follows. Section 2 describes the model and the structure of the

[3] More specifically, they find that "the slant of co-owned papers is only weakly (and statistically insignificantly) correlated to a newspaper's political alignment" (Gentzkow and Shapiro, 2009, page 5).

[4] Calvert (1985) was the first to point out the positive value of a biased source of information for a rational decision-maker.

[5] Notice also that in their model the cost of acquiring information is embedded in the discount factor. Their results do not apply in presence of a per unit cost of sampling since individuals differ only in the variance of their payoffs but not in their *ex-ante* ranking between actions.

game. Section 3 derives the optimal information acquisition strategy by citizen-editors. Section 4 discusses the demand for news. Section 5 contains the results on the optimal choice of editors by media outlets. Section 6 concludes. All the proofs are provided in the appendix.

2 The Model

2.1 Citizens

There is a continuum of *citizens* of measure one who have to make a decision regarding a single issue or policy P. Without loss of generality, we assume the policy space to be $\Psi = [0, 1]$. There are only two possible alternative candidates/policies L and R (i.e., $P = \{L; R\}$) where $L = 0$ and $R = 1$. There are two possible states of the world $s \in \{l, r\}$, where the prior probability of the state of the world being $s = r$ is assumed to be common knowledge and it is denoted by q.

Citizens care about the ideological distance between their idiosyncratic preferences and the candidates' policy platforms. That is, citizens want to minimize the euclidean distance between their policy preferences and the ones of the chosen candidate. At the same time citizens also care about the *valence* (i.e., quality) of the candidates. The valence component is captured by an additive constant in the citizen's utility function. That is, regardless of her idiosyncratic policy preferences, each citizen gets an extra positive payoff when she chooses the high valence candidate and a negative one when the low valence candidate is chosen.[6] Hence, citizen i's utility function is:

$$u_i(P, x_i) = \delta I_s I_p - |P - x_i| \tag{1}$$

where x_i represents the idiosyncratic policy preference of citizen i. Moreover, $\delta \in (0, \frac{1}{2}]$ and:

$$I_s = \begin{cases} 1 \text{ if } s = l \\ -1 \text{ if } s = r \end{cases} \quad \text{and} \quad I_p = \begin{cases} 1 \text{ if } P = L \\ -1 \text{ if } P = R \end{cases} \tag{2}$$

As a consequence, candidate L gives a higher utility to citizens when the state of the world is l than when the state is r (*viceversa* for candidate R).[7] In other words, while L and R represent the alternative political platforms of two candidates, 2δ can be seen

[6] As usual in the literature on the demand for news media (e.g., Mullainathan and Shleifer 2005, Gentzkow and Shapiro 2006, Chan and Suen 2008) we assume that citizens receive utility from choosing a given candidate/alternative *per se*.

[7] For a similar specification of the voters' utility function see, for example, Aragones and Palfrey (2002).

as the difference in the *valence* of the two candidates in each state of the world.[8] The idiosyncratic preferences of citizens are distributed with a common knowledge c.d.f. $F(x)$ with density function $f(x)$ where $supp\,[f(x)] = [0,1]$. Hence, the state contingent utilities of citizen i are as follows:

$$u_i(L) = \begin{cases} \delta - x_i & \text{if } s = l \\ -\delta - x_i & \text{if } s = r \end{cases} \quad \text{and} \quad u_i(R) = \begin{cases} -\delta + x_i - 1 & \text{if } s = l \\ \delta + x_i - 1 & \text{if } s = r \end{cases} \tag{3}$$

Notice also that for any citizen i the two candidates have the same variance in payoffs and such variance is equal across citizens since:

$$u_i(L|s = l) - u_i(L|s = r) = u_i(R|s = r) - u_i(R|s = l) = 2\delta \qquad \forall i$$

Let $\Sigma = \{\sigma_l, \sigma_r\}$ be the signal space. The signal likelihood function is as follows:

$$\Pr(\sigma_l|s = l) = \Pr(\sigma_r|s = r) = \theta \tag{4}$$

where $\theta \in \left(\frac{1}{2}, 1\right)$ represents the precision of the signal. Suppose now that citizens receive n_l signals σ_l and n_r signals σ_r on the state of the world. Then the citizens' posterior beliefs are:

$$\Pr(s = r|n_l, n_r) = \frac{\Pr(n_l, n_r|s = r)\Pr(s = r)}{\Pr(n_l, n_r|s = r)\Pr(s = r) + \Pr(n_l, n_r|s = l)\Pr(s = l)}$$

Thus

$$\Pr(s = r|n_l, n_r) = \frac{q\theta^{n_r - n_l}}{q\theta^{n_r - n_l} + (1-q)(1-\theta)^{n_r - n_l}}$$

Therefore, denoting $n = n_r - n_l$ we can write the citizens' posterior beliefs as:

$$\mu(n) = \frac{1}{1 + \frac{1-q}{q}\left(\frac{1-\theta}{\theta}\right)^n} \tag{5}$$

Hence, citizen i prefers candidate R to candidate L whenever:

$$\mu(n) > \frac{1}{4\delta}(2\delta - 2x_i + 1) = \mu(\hat{n}_i) = \hat{\mu}_i \tag{6}$$

That is $\hat{n}_i$ is the difference in the number of signals in favor of state r which makes citizen i being indifferent between candidates R and L. Notice that for $\delta = \frac{1}{2}$ we always have

[8] As an alternative interpretation of the model, L and R can be seen as two alternative policies (e.g. implementing Kyoto's protocol or not). Hence, if the state of the world is l then the public benefits/cost ratio of policy L is higher than the one of R (*viceversa* if $s = r$). That is, if the state of the world is l policy L is the most efficient one.

that $\hat{\mu}_i > 0$. Hence for $\delta = \frac{1}{2}$ all citizens would prefer candidate L when $s = l$ and candidate R when $s = r$. That is, when $\delta = \frac{1}{2}$, *ex-post* all citizens have the same ranking of preferences over candidates. Instead, for $0 < \delta < \frac{1}{2}$ there will be some "stubborn" citizens who will always vote for the same candidate regardless of the state of the world.[9] Moreover:

$$\frac{\partial u_i(R)}{\partial \mu(n)} = -\frac{\partial u_i(L)}{\partial \mu(n)} = 2\delta, \;\; \forall i$$

that is, the utility functions of citizens i and j are always parallel. We can thus represent the utilities of citizens as follows:

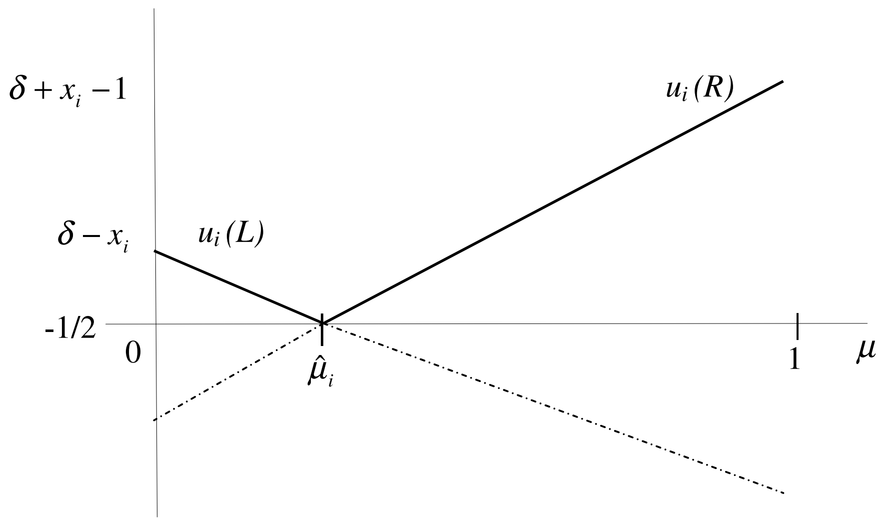

Figure 1. Utility of citizen i for $x_i > 1/2$

For any exogenously given $\mu(n) \in (0, 1)$, different citizens may have different ranking of preferences regarding candidates L and R. More specifically:

$$\hat{\mu}_{\frac{1}{2}} = \frac{1}{2} \quad \text{and} \quad \frac{\partial \hat{\mu}_i}{\partial x_i} < 0 \tag{7}$$

Thus, citizens with more "rightist" preferences require less evidence in favor of R in order to choose that candidate with respect to moderate citizens. Notice also that:

$$u_i(L|\hat{\mu}_i) = u_i(R|\hat{\mu}_i) = -\frac{1}{2} \qquad \forall i$$

Hence the expected utilities of citizens i and j for $x_j = 1/2 < x_i$ are as follows:

[9] Notice that assuming $\delta \in (0, \frac{1}{2}]$ is without loss of generality. The same results would hold in a model where $\delta \in \mathbb{R}^+$ and *supp* $[f(x)] = \mathbb{R}$.

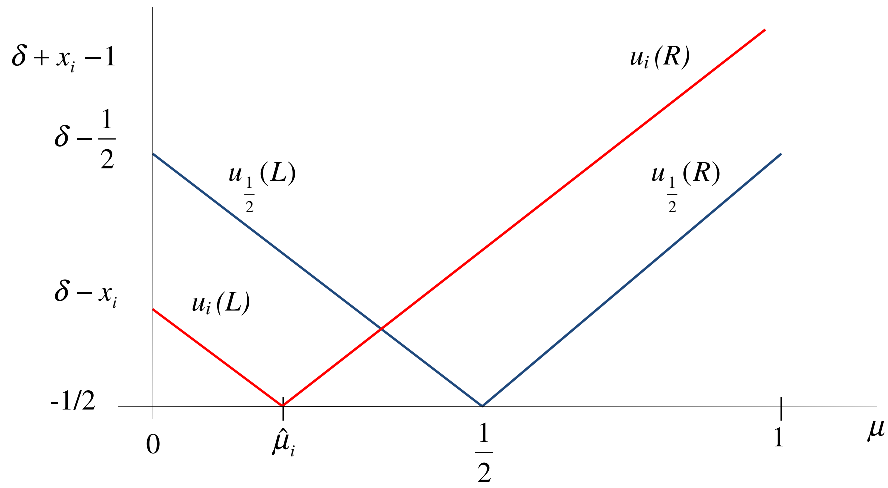

Figure 2. Expected utilities of citizens i and j for $x_j = \frac{1}{2} < x_i$

Notice also that when a citizen cares more about the true state of the world (i.e., when the *valence* component is larger), her indifference threshold is closer to the one of a moderate citizen. That is:

$$\frac{\partial \hat{\mu}_i}{\partial \delta} = \frac{(2x_i - 1)}{4\delta^2} \begin{cases} < 0 & \text{if } \ x_i < \frac{1}{2} \\ > 0 & \text{if } \ x_i > \frac{1}{2} \end{cases} \tag{8}$$

In other words, the more citizens care about the quality of different candidates, the less evidence in favor of the least ideologically closer candidate they require in order to vote for her.

2.2 The Game

There is a media industry composed by $K \geq 1$ media outlets. We assume that each media outlet wants to maximize its viewership in order to maximize its advertising revenues. In order to produce news reports each media outlet has to hire an editor from the population of citizens. Once hired, the citizen-editor is endowed with a (costly) technology that allows her to collect evidence on the state of the world. More specifically, we assume that an editor has to incur a cost c any time she decides to get a signal on the state of the world (e.g., effort she has to exert to acquire information, opportunity cost of sending reporters to investigate an issue, etc.).[10] The media outlet will then produce a news report based on the editor's optimal sampling strategy. Citizens will then decide whether to access a media outlet's report by paying an opportunity cost C or not. If they decide to watch a media outlet's report they update their beliefs using Bayes' rule. Hence, the demand of news reports that a media outlet faces is a function of the type of editor that it has hired. That is, given an editor with idiosyncratic preferences x_e, the profit function of

[10] By "editor" we refer to what is usually called "Editor-in-Chief" for a newspaper and "Managing Editor" in the broadcast media sector. More in general, the model applies to the choice of a profit maximizing media outlet regarding the type of journalists to be hired.

media outlet k is $\Pi_k(x_e) = D_k(x_e)$, where $D_k(x_e)$ is the demand for the news report produced by the media outlet. To avoid the presence of exogenous asymmetries, we focus on symmetric distribution of citizens' idiosyncratic preferences that are monotone in the sub-intervals $x \in [0, \frac{1}{2}]$ and $x \in [\frac{1}{2}, 1]$.[11] To summarize, the timing of the game is as follows:

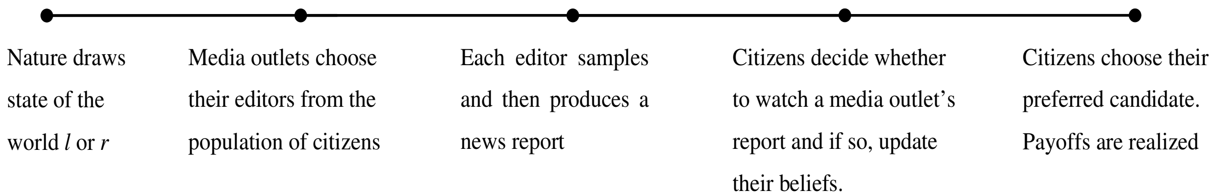

Figure 3. Timing of the Game

We now turn to the analysis of the optimal strategy of a citizen-editor (i.e., her optimal sampling strategy). Then we derive the demand of news reports by citizens (i.e., $D_k(x_e)$) as a function of an editor's optimal sampling strategy. Finally, we analyze the profit-maximizing strategy of media outlets (i.e., which type of editor maximizes the profits of the media outlet) and discuss the results.

3 Optimal Information Acquisition by Citizens-Editors

Suppose that a media outlet has hired a citizen with idiosyncratic preferences x_i to work as its editor. We denote by $\tau_{i,m}(n)$ the decision of a citizen-editor i given that she has already drawn $m = \{0, 1, ...\infty\}$ signals and given a current difference of signals in favor of r equal to n. Given any m and n, the choice set of citizen-editor i is $\Gamma_m(n) = \{L, R, d\}$. Thus she can choose candidate L or R or she can pay c and draw another signal on the state of the world (i.e., choose $\tau_{i,m}(n) = d$, where d stands for "draw").

Hence, an editor faces a trade-off between the cost of acquiring a signal and the utility she gets from the informative content of each signal. Her problem is thus to find an optimal stopping rule. More specifically, the value function that editor i maximizes after m draws, given a current difference of signals in favor of state r equal to n, is the following:

$$V_i(n) = \begin{cases} \max \begin{Bmatrix} \delta(1 - 2\mu(n)) - x_i; \\ v(n)V_i(n+1) + (1 - v(n))\, V_i(n-1) - c \end{Bmatrix} & \text{if } \mu(n) < \hat{\mu}_i \\ \max \begin{Bmatrix} \delta(2\mu(n) - 1) - (1 - x_i); \\ v(n)V_i(n+1) + (1 - v(n))\, V_i(n-1) - c \end{Bmatrix} & \text{if } \mu(n) \geq \hat{\mu}_i \end{cases} \tag{9}$$

[11] For example, the families of Uniform, Normal, and Cauchy distribution functions satisfy such property.

where $v(n) = \mu(n)\theta + (1-\mu(n))(1-\theta)$. In other words, if after m draws editor i has a posterior $\mu(n) < \hat{\mu}_i$ she will choose between alternative L with an expected payoff of $(1-\mu(n))(\delta - x_i) + \mu(n)(-\delta - x_i)$ or paying c and getting another signal. In this case, with probability v the editor will get signal σ_r in which case the value function becomes $V_i(n+1)$ and with probability $(1-v)$ she will get signal σ_l in which case the value function becomes $V_i(n-1)$. *Viceversa*, if after m draws editor i has a posterior $\mu(n) \geq \hat{\mu}_i$ she will choose between alternative R with an expected payoff of $(1-\mu(n))(x_i - \delta - 1) + \mu(n)(x_i + \delta - 1)$ or paying c and getting another signal. In this case, with probability v the editor will get signal σ_r in which case the value function becomes $V_i(n+1)$ and with probability $(1-v)$ she will get signal σ_l in which case the value function becomes $V_i(n-1)$. Notice also that the value function of editor i does not depend on how many draws she has already done (i.e., m), since the only relevant variable for her decision is the current difference of signals in favor of r (i.e., n).

The following Proposition characterizes the properties of the optimal information acquisition strategy by an editor.

Proposition 1 *For all $c > 0$, there exist $(\underline{n}_i^*, \bar{n}_i^*)$ such that for $\forall m, \forall i$:*

1. $\tau_{i,m}(n) = L$ *if* $n < \underline{n}_i^*$, $\tau_{i,m}(n) = R$ *if* $n > \bar{n}_i^*$ *and* $\tau_{i,m}(n) = d$ *if* $n \in (\underline{n}_i^*, \bar{n}_i^*)$.
2. $\frac{\partial \underline{n}_i^*}{\partial x_i} < 0$, $\frac{\partial \underline{n}_i^*}{\partial \delta} < 0$ *and* $\frac{\partial \underline{n}_i^*}{\partial c} > 0$
3. $\frac{\partial \bar{n}_i^*}{\partial x_i} < 0$, $\frac{\partial \bar{n}_i^*}{\partial \delta} > 0$ *and* $\frac{\partial \bar{n}_i^*}{\partial c} < 0$

Moreover

$$\left|\frac{\partial \bar{n}_i^*}{\partial x_i}\right| \begin{cases} < \left|\frac{\partial \underline{n}_i^*}{\partial x_i}\right| & \text{for } x_i < \frac{1}{2} \\ = \left|\frac{\partial \underline{n}_i^*}{\partial x_i}\right| & \text{for } x_i = \frac{1}{2} \\ > \left|\frac{\partial \underline{n}_i^*}{\partial x_i}\right| & \text{for } x_i > \frac{1}{2} \end{cases}$$

and

$$\left|\frac{\partial \bar{n}_i^*}{\partial \delta}\right| \begin{cases} < \left|\frac{\partial \underline{n}_i^*}{\partial \delta}\right| & \text{for } x_i < \frac{1}{2} \\ = \left|\frac{\partial \underline{n}_i^*}{\partial \delta}\right| & \text{for } x_i = \frac{1}{2} \\ > \left|\frac{\partial \underline{n}_i^*}{\partial \delta}\right| & \text{for } x_i > \frac{1}{2} \end{cases} ; \qquad \left|\frac{\partial \bar{n}_i^*}{\partial c}\right| \begin{cases} < \left|\frac{\partial \underline{n}_i^*}{\partial c}\right| & \text{for } x_i < \frac{1}{2} \\ = \left|\frac{\partial \underline{n}_i^*}{\partial c}\right| & \text{for } x_i = \frac{1}{2} \\ > \left|\frac{\partial \underline{n}_i^*}{\partial c}\right| & \text{for } x_i > \frac{1}{2} \end{cases}$$

Proof. See Appendix. ■

The following graph illustrates the optimal strategy of editor i after m draws, given a current difference of signals in favor of r equal to n:

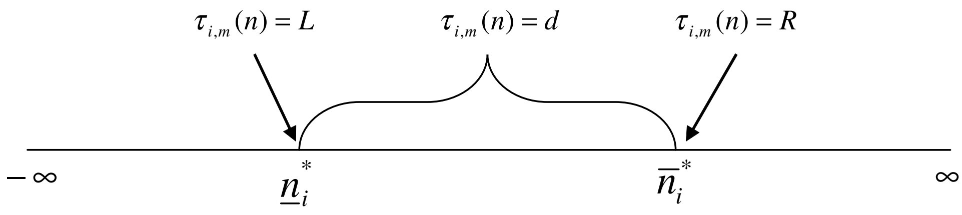

Figure 4. Optimal Strategy of editor i

In other words, $\underline{n}_i^*$ is the threshold below which editor i does not sample anymore and reports $|\underline{n}_i^*|$ more signals in favor of candidate L. Similarly, $\bar{n}_i^*$ is the threshold above which editor i does not sample anymore and reports $\bar{n}_i^*$ more signals in favor of candidate R.

For any given n a more "rightist" editor is always more likely to produce a report in favor of candidate R than in favor of L, with respect to a more "leftist" editor. Thus $x_i > x_j$ implies that $\underline{n}_i^* < \underline{n}_j^*$ and $\bar{n}_i^* < \bar{n}_j^*$. Moreover, given editors j and i with $x_j < x_i < \frac{1}{2}$, then $\bar{n}_j^* - \underline{n}_j^* < \bar{n}_i^* - \underline{n}_i^*$. That is, a leftist editor requires even less signal in favor of L than more in favor of R to stop sampling, with respect to a moderate editor. Similarly, given editors j and i with $x_j > x_i > \frac{1}{2}$, then $\bar{n}_j^* - \underline{n}_j^* < \bar{n}_i^* - \underline{n}_i^*$. That is, a rightist editor requires even less signal in favor of R than more in favor of L with respect to a moderate editor. Hence, the more moderate an editor is, the larger is her "information acquisition set" $N_i = \{n|\tau_{i,m}(n) = d\}$ (i.e., the set of the difference in the number of signals in favor of r (or in favor of l) such that editor i will keep sampling).[12] At the same time, an increase in the importance of the valence component of the editor's utility function (δ) makes an editor sample more in both directions (i.e., N_i becomes larger). Moreover, an increase in δ induces a leftist editor to increase her "leftist" stopping rule more than her "rightist" stopping rule (i.e., $|\underline{n}_i^*|$ increases more than $\bar{n}_i^*$). The opposite is true for a rightist editor. That is, a higher δ is associated with more sampling in both directions *and* a more symmetric stopping rule for all types of editors. Finally, when the cost of each signal is higher, it is optimal for each editor to "make her mind" sooner (i.e., N_i shrinks). Moreover, when c is larger the stopping rule of each editor is also more asymmetric. Hence, when information acquisition is more costly, each editor finds optimal to devote more resources in acquiring information in the direction that could change her "*ex-ante* decision" (i.e., in the direction of persuading her not to vote for the ideologically closer candidate).

Therefore, proposition 1 suggests that when δ is higher and c is lower we should expect any type of editor: *i)* to acquire more information; *ii)* to behave *as if* she were more moderate (i.e., to have more symmetric stopping rules).

Notice that, for $x_i = \frac{1}{2}$, $\bar{n}_i^* - \hat{n}_i = \hat{n}_i - \underline{n}_i^*$ and thus $\mu(\bar{n}_i^*) = 1 - \mu(\underline{n}_i^*)$. Moreover for

[12] Notice that it is always the case that either $N_i \equiv \varnothing$ or $N_i \equiv \{\underline{n}_i^*, \underline{n}_i^* + 1, ..., \bar{n}_i^* - 1, \bar{n}_i^*\} \supseteq \{0\}$.

$x_i > x_j$:

$$\mu(\underline{n}_i^*) < \mu(\underline{n}_j^*) < q < \mu(\bar{n}_i^*) < \mu(\bar{n}_j^*) \tag{10}$$

Moreover, given the comparative statics results of Proposition 1, we can directly derive some comparative statics results on the probability of choosing the "wrong" candidate.

Corollary 1

$\frac{\partial \Pr(\tau_i(n)=L|s=r)}{\partial c} > 0$ *and* $\frac{\partial \Pr(\tau_i(n)=R|s=l)}{\partial c} > 0$

$\frac{\partial \Pr(\tau_i(n)=L|s=r)}{\partial \delta} < 0$ *and* $\frac{\partial \Pr(\tau_i(n)=R|s=l)}{\partial \delta} < 0$

$\frac{\partial \Pr(\tau_i(n)=L|s=r)}{\partial x_i} < 0$ *and* $\frac{\partial \Pr(\tau_i(n)=R|s=l)}{\partial x_i} > 0$

Moreover, the more moderate an editor is, the lower is her overall probability of making errors.

Thus as expected, when the cost of sampling is higher editors will make more "errors" in the sense that they would be less likely to choose the high valence candidate. *Viceversa*, when editors care more about the quality of candidates their probability of mistakenly choosing the low quality candidate decreases (since as shown by proposition 1, when δ is higher editors acquire more information). On the other hand, more "rightist" editors are less likely to choose candidate L when the high quality one is R and are instead more likely to choose candidate R when the high quality one is L. However, overall, moderate editors are less likely to make a report in favor of the low quality candidate. This is due to the fact that, as shown by proposition 1, the more moderate an editor is, the more symmetric her sampling strategy is and also the more information she acquires before making a decision. Therefore, by taking on average a "more informed" decision, moderate editors are less likely to choose the low quality candidate.

For ease of notation, from now on we will denote the idiosyncratic preferences of an editor x_e and, thus, her optimal stopping rules as $(\underline{n}_e^*, \bar{n}_e^*)$. Next section analyzes the demand by citizens for the news reports of a media outlet as a function of the optimal stopping rules of its editor.

4 The Demand for News

In the previous section we have derived the optimal sampling strategy of an editor as a function of her idiosyncratic preferences. Moreover, given the idiosyncratic preferences of a media outlet's editor, each citizen i can infer the set of possible reports of a media outlet (i.e., citizen i knows that the editor will either stop acquiring information after having collected $\underline{n}_e^*$ signals in favor of L or $\bar{n}_e^*$ in favor of R). Let the citizens' action space be $A = \{W, NW\}$ where W stands for watch the news reports and NW for not watching the news reports. Then, the expected utility of citizen i from not getting any news report from the media outlet is:

$$U_i(NW) = \max\left\{U_i(L|q); U_i(R|q)\right\}$$

that is for $q = \frac{1}{2}$:

$$U_i(NW) = \begin{cases} U_i\left(L|\frac{1}{2}\right) & \text{for } x_i < \frac{1}{2} \\ U_i\left(R|\frac{1}{2}\right) & \text{for } x_i > \frac{1}{2} \end{cases}$$

If instead citizen i decides to pay a cost C to access the news reports, her expected utility will be:

$$\begin{aligned} U_i(W) = \Pr(n = \underline{n}_e^*) \max\left\{U_i\left(L|\mu(\underline{n}_e^*)\right); U_i\left(R|\mu(\underline{n}_e^*)\right)\right\} \\ + \Pr(n = \bar{n}_e^*) \max\left\{U_i\left(L|\mu(\bar{n}_e^*)\right); U_i\left(R|\mu(\bar{n}_e^*)\right)\right\} - C \end{aligned} \tag{11}$$

Where:[13]

$$\Pr(n = \underline{n}_e^*) = \frac{2\mu(\bar{n}_e^*) - 1}{2\left[\mu(\bar{n}_e^*) - \mu(\underline{n}_e^*)\right]} \tag{12}$$

and

$$\Pr(n = \bar{n}_e^*) = \frac{1 - 2\mu(\underline{n}_e^*)}{2\left[\mu(\bar{n}_e^*) - \mu(\underline{n}_e^*)\right]} \tag{13}$$

Let's now focus on the marginal viewer. That is, the viewer who is indifferent between watching and not watching the media outlet's reports. More specifically, there will be two marginal viewers. One representing the most rightist citizen willing to watch news reports from a media outlet having an editor with idiosyncratic preferences x_e. The other one representing the most leftist citizen willing to watch such news reports. That is, there will be a $\hat{x}_e = \hat{x}_e(x_e)$ and a $\tilde{x}_e = \tilde{x}_e(x_e)$ with $\hat{x}_e < \tilde{x}_e$ such that only citizens with $x_i \in [\hat{x}_e, \tilde{x}_e]$ will watch the news reports.[14]

Let's start analyzing the marginal viewer for $x_i < \frac{1}{2}$. Then $U_i(NW) = U_i\left(L|\frac{1}{2}\right)$ and since by (10) $\underline{n}_e^* < 0 < \bar{n}_e^*$, it must be the case that:

$$U_i\left(L|\mu(\underline{n}_e^*)\right) > U_i\left(R|\mu(\underline{n}_e^*)\right)$$

Moreover, the following individual rationality constraint must be satisfied for leftist citizens:

$$U_i\left(L|\mu(\bar{n}_e^*)\right) < U_i\left(R|\mu(\bar{n}_e^*)\right) \tag{IR_L}$$

otherwise, if $U_i\left(L|\mu(\bar{n}_e^*)\right) > U_i\left(R|\mu(\bar{n}_e^*)\right)$ (i.e., if citizen i would always prefer alternative L regardless of watching or not the news reports) then watching the news reports would

[13] These are simply the probabilities of hitting the two stopping thresholds in a stochastic process with two absorbing states. See Brocas and Carrillo (2007) for an analogous derivation.

[14] Notice that it could also be the case that $\hat{x}_e > \frac{1}{2}$ or $\tilde{x}_e < \frac{1}{2}$ but not both.

not be *ex-post* rational given the cost C. Thus the marginal leftist viewer will be the one having idiosyncratic preferences $\hat{x}_e$ such that:

$$U_i\left(L|\frac{1}{2}\right) = \frac{2\mu(\bar{n}_e^*) - 1}{2\left[\mu(\bar{n}_e^*) - \mu(\underline{n}_e^*)\right]} U_i\left(L|\mu(\underline{n}_e^*)\right) + \frac{1 - 2\mu(\underline{n}_e^*)}{2\left[\mu(\bar{n}_e^*) - \mu(\underline{n}_e^*)\right]} U_i\left(R|\mu(\bar{n}_e^*)\right) - C$$

that is:

$$\hat{x}_e = \frac{1}{2} - \delta(2\mu(\bar{n}_e^*) - 1) + \frac{C}{2\Pr(n = \bar{n}_e^*)} \tag{14}$$

Notice also that the *ex-post* rationality constraint (IR_L) is satisfied as long as $x > \frac{1}{2} - \delta(2\mu(\bar{n}_e^*) - 1) = x^{Min}$. Hence, since $\hat{x}_e > x^{Min}$, such constraint is automatically satisfied for any citizen willing to watch the news reports.

Let's now focus on the marginal viewer for $x_i > \frac{1}{2}$. Then $U_i(NW) = U_i\left(R|\frac{1}{2}\right)$ and since by (10) $\underline{n}_e^* < 0 < \bar{n}_e^*$, it must be the case that:

$$U_i\left(R|\mu(\bar{n}_e^*)\right) > U_i\left(R|\frac{1}{2}\right)$$

Moreover, the following individual rationality constraint must be satisfied for rightist citizens:

$$U_i\left(L|\mu(\underline{n}_e^*)\right) > U_i\left(R|\mu(\underline{n}_e^*)\right) \tag{IR_R}$$

otherwise, if $U_i\left(L|\mu(\underline{n}_e^*)\right) < U_i\left(R|\mu(\underline{n}_e^*)\right)$ (i.e., if citizen i would always prefer alternative R regardless of watching or not the news reports) then watching the news reports would not be *ex-post* rational given the cost C. Thus the marginal rightist viewer will be the one having idiosyncratic preferences $\tilde{x}_e$ such that:

$$U_i\left(R|\frac{1}{2}\right) = \frac{2\mu(\bar{n}_e^*) - 1}{2\left[\mu(\bar{n}_e^*) - \mu(\underline{n}_e^*)\right]} U_i\left(L|\mu(\underline{n}_e^*)\right) + \frac{1 - 2\mu(\underline{n}_e^*)}{2\left[\mu(\bar{n}_e^*) - \mu(\underline{n}_e^*)\right]} U_i\left(R|\mu(\bar{n}_e^*)\right) - C$$

that is:

$$\tilde{x}_e = \frac{1}{2} + \delta(1 - 2\mu(\underline{n}_e^*)) - \frac{C}{2\Pr(n = \underline{n}_e^*)} \tag{15}$$

Notice also that the *ex-post* rationality constraint (IR_R) is satisfied as long as $x < \frac{1}{2} + \delta(1 - 2\mu(\underline{n}_e^*)) = x^{Max}$. Hence, since $\tilde{x}_e < x^{Max}$, such constraint is automatically satisfied for any citizen willing to watch the news reports. We now introduce the following assumption:

Assumption 1

$$C < C^{MAX} = \delta\left(\frac{1 - \lambda^{\bar{n}_e^*|_{x_e = \frac{1}{2}}}}{1 + \lambda^{\bar{n}_e^*|_{x_e = \frac{1}{2}}}}\right)$$

It is easy to prove that when this assumption does not hold and $C > C^{MAX}$, there will never be any leftist or rightist citizen willing to watch any news report.[15] The following Lemma contains the main properties of the demand for news.

Lemma 1 *Let* $(\bar{n}_e^*, \underline{n}_e^*)$ *the optimal stopping rules of an editor with idiosyncratic preferences* x_e*. Then:*

i) $\frac{\partial \hat{x}_e}{\partial C} > 0$; $\frac{\partial \tilde{x}_e}{\partial C} < 0$

ii) $\frac{\partial \hat{x}_e}{\partial \delta} < 0$; $\frac{\partial \tilde{x}_e}{\partial \delta} > 0$

iii) $\frac{\partial \hat{x}_e}{\partial \Pr(n=\bar{n}_e^*)} < 0$; $\frac{\partial \tilde{x}_e}{\partial \Pr(n=\underline{n}_e^*)} > 0$

iv) $\frac{\partial \hat{x}_e}{\partial \underline{n}_e^*} > 0$; $\frac{\partial \tilde{x}_e}{\partial \underline{n}_e^*} < 0$

v) $\frac{\partial \hat{x}_e}{\partial \bar{n}_e^*} < 0$; $\frac{\partial \tilde{x}_e}{\partial \bar{n}_e^*} > 0$

vi) $\frac{\partial \Pr(n=\underline{n}_e^*)}{\partial x_e} = -\frac{\partial \Pr(n=\bar{n}_e^*)}{\partial x_e} < 0$

Let's now discuss the meaning of these results. Obviously, a higher opportunity cost of watching news reports decreases the number of leftist and rightist citizens willing to watch such reports. *Viceversa*, the higher is the valence component in the citizens utility function, the more leftist and rightist citizens will want to watch news. That is the more citizens care about knowing the state of the world, the more citizens will get informed. On the other hand, the number of leftist citizens watching news reports is an increasing function of the probability of an editor "hitting" the rightist threshold. That is, the more likely that the information collected by an editor will be pivotal in the leftist citizens' decision, the more leftist citizens will be willing to watch the media outlet's reports (a symmetric intuition holds for rightist citizens). Notice also that all citizens care about receiving the most accurate information. Indeed, the lower is $\underline{n}_e^*$ and the higher is $\bar{n}_e^*$, the more citizens will want to get informed. On the other hand, as expected, a more rightist editor has a lower probability of hitting the "leftist threshold" and a higher probability of hitting the "rightist threshold".

Hence, the above lemma summarizes the main features of the demand for news media by citizens. All citizens who value information (i.e., the ones whose *ex-post* ranking of candidates is not always the same as their *ex-ante* one) would like to watch a news media outlet having an editor who samples in both directions until infinity, since the more

[15] Notice that we have chosen $C^{MAX} = C^{MAX}\big|_{x_e=\frac{1}{2}}$ since $C^{MAX}\big|_{x_e\neq\frac{1}{2}} = \delta\frac{\left(\lambda^{\underline{n}_e^*}-1\right)\left(1-\lambda^{\bar{n}_e^*}\right)}{\lambda^{\underline{n}_e^*}-\lambda^{\bar{n}_e^*}}$ and $\frac{dC^{MAX}\big|_{x_e\neq\frac{1}{2}}}{dx_e}\begin{cases} >0 \text{ for } x_i<\frac{1}{2} \\ =0 \text{ for } x_i=\frac{1}{2} \\ <0 \text{ for } x_i>\frac{1}{2}\end{cases}$ hence assuming that $C < C^{MAX}\big|_{x_e=\frac{1}{2}}$ is the least restrictive assumption. Notice also that $\frac{\partial C^{Max}}{\partial \delta} > 0$ and $\frac{\partial C^{Max}}{\partial \bar{n}_e^*} > 0$.

information she gets, the higher the citizens' expected utility. However, given the cost of acquiring information for an editor and the opportunity cost that each citizen faces when accessing this information, when a citizen is choosing whether to watch a media outlet and/or choosing among alternative news media outlets, she takes into account two different components. That is, she considers how similar an editor's idiosyncratic preferences are to hers (i.e., how much "valuable" the information provided by an editor could be to her) but she also values the "intensity" of information acquisition by an editor (i.e., how much information an editor is acquiring and thus providing, on average).

More specifically, citizens can be divided into two categories depending on their idiosyncratic preferences. Citizens with preferences $x_i < \hat{x}_e|_{x_e=\frac{1}{2}}$ and $x_i > \tilde{x}_e|_{x_e=\frac{1}{2}}$ are "relatively extremists".[16] For these citizens only a media outlet with an editor with similar idiosyncratic preferences can be pivotal for their choice (i.e., they never find valuable the information coming from a moderate editor). Hence, either they will watch a media outlet with an editor with (sufficiently) similar preferences or they will not watch any media outlet at all.

On the other hand, citizens with preferences $\hat{x}_e|_{x_e=\frac{1}{2}} < x_i < \tilde{x}_e|_{x_e=\frac{1}{2}}$ are "relatively moderate" (i.e., liberal-moderates for $\hat{x}_e|_{x_e=\frac{1}{2}} < x_i < \frac{1}{2}$ and conservative-moderates for $\frac{1}{2} < x_i < \tilde{x}_e|_{x_e=\frac{1}{2}}$). These citizens find valuable the information coming from a moderate editor but they may find even more valuable the information coming from an editor with similar idiosyncratic preferences. More specifically, each citizen faces a basic trade-off between the "amount" and the "value" of information coming from different types of editors. A citizen can make two specular errors. She may choose L when L is the low quality candidate. Similarly, she may choose R when R is the low quality candidate. A moderate citizen (i.e., $x_i = \frac{1}{2}$) cares about these two errors equally. Hence, she always prefers to watch a media outlet having a moderate editor since such editor minimizes the overall probability of making errors (see Corollary 1). On the other hand, a liberal-moderate citizen cares more about not making the error of choosing R when $s = l$. As we have seen in Corollary 1, a liberal editor has a lower probability of making such error but a higher probability of making a report in favor of L when $s = r$ and a higher overall probability of making errors. Hence, when choosing between a media outlet with a moderate editor and one with a liberal editor, a liberal-moderate citizen will trade-off the *amount* and the *value* of information provided by different types of editors.[17]

The following section analyzes what are the implications of such demand for news for

[16] Notice that not all these citizens can be properly defined as "extremists" since not everyone of them is stubborn. Some of them may change their *ex-ante* ranking of preferences over candidates if they receive enough information in favor of the ideologically least preferred candidate (notice that for $\delta = \frac{1}{2}$ everyone would do so upon knowing the true state of the world).

[17] Durante and Knight (2009) analyze the demand of news in Italy. They show that, indeed, when the ideological position of a media outlet changes, viewers change their choice of news programs accordingly.

the optimal choice of editors by profit maximizing media outlets.

5 Optimal Choice of Editors by Media

5.1 Monopoly

We want now to analyze the implications of such citizens-editors model in a monopolistic market. The media outlet's owner wants to choose x_e to maximize viewership. Choosing an editor from the population of citizens is analogous to choosing a "product" location on the $[0, 1]$ line. Suppose the media outlet's owner chooses an editor with idiosyncratic preferences x_e. Then, the profit function is:

$$\Pi(x_e, \hat{x}_e, \tilde{x}_e) = F(\tilde{x}_e) - F(\hat{x}_e)$$

Hence, the media outlet owner will choose an editor with preferences x_e^{Mon} such that:

$$\left.\frac{\partial D(x_e)}{\partial x_e}\right|_{x_e = x_e^{Mon}} = 0$$

where $F(\tilde{x}_e)$ and $F(\hat{x}_e)$ are increasing functions of x_e. The following proposition characterizes under which conditions a profit-maximizing media outlet will hire a moderate editor and under which conditions it will hire a non-moderate one.

Proposition 2 *Suppose there is just a monopolist media outlet in the market for news. Then,*

1. *If*

$$\frac{\partial f(x_i)}{\partial x_i} \begin{cases} \geq 0 \text{ for } x_i < \frac{1}{2} \\ \leq 0 \text{ for } x_i > \frac{1}{2} \end{cases} \tag{16}$$

then the media outlet will always hire a moderate editor (i.e., $x_e^{Mon} = \frac{1}{2}$).

2. *If*

$$\frac{\partial f(x_i)}{\partial x_i} \begin{cases} < 0 \text{ for } x_i < \frac{1}{2} \\ > 0 \text{ for } x_i > \frac{1}{2} \end{cases} \tag{17}$$

then the media outlet will always hire a non moderate editor with preferences $x_{e_R}^{Mon} > \frac{1}{2}$ (or, equivalently, one with preferences $x_{e_L}^{Mon} = 1 - x_{e_R}^{Mon} < \frac{1}{2}$). However, even in this case there is an upper bound on the "extremism" of the optimal editor. That is

$\exists \tilde{C} < C^{Max}$ *with* $\frac{\partial \tilde{C}}{\partial x_e} \begin{cases} > 0 \text{ for } x_e < \frac{1}{2} \\ < 0 \text{ for } x_e > \frac{1}{2} \end{cases}$ *such that* $x_{e_R}^{Mon} \in (\frac{1}{2}, x_R]$ *where* x_R *is such that:*

$$\tilde{x}(x_{e_R}) = \tilde{x}\left(\tilde{C}\Big|_{x_R} = C\right)$$

The above proposition is showing that a monopolist media outlet will always choose a moderate editor when citizens are distributed uniformly or when the mass of moderate citizens is higher than the one of "extremists" ones. Instead, if the number of moderate citizens is lower than the one of "extremists", the media outlet will prefer to hire a non-moderate editor. Hence, even though citizens do not derive any exogenous utility from acquiring biased information and the media outlet is just maximizing profits, the endogenous acquisition of costly information may induce a media outlet to choose an editor whose optimal information acquisition strategy is "slanted" in favor of the alternative *ex-ante* preferred by a subset of citizens (e.g., the rightists one). Moreover, non-moderate citizens may prefer to face a media outlet having an editor with this kind of slant in information acquisition rather than facing a moderate one (even when all citizens share the same *ex-post* preferences over candidates, i.e., when $\delta = \frac{1}{2}$).

However, even in this case the optimal editor will not be too extremist. Less moderate citizens will indeed trade-off the benefit of having an editor with similar preferences and the cost of having an editor who will sample relatively less. Hence choosing a less moderate editor, after some point, will decrease the number of "rightist" citizens willing to watch the media outlet (i.e., for $x_e > x_R$).

5.2 Duopoly

Suppose now that $K = 2$. That is, the market for news is composed by two profit maximizing media outlets. The following proposition summarizes the possible Nash equilibria that can arise in this case depending on the distribution of citizens' preferences.

Proposition 3 *Suppose there are two media outlets in the market for news. Then:*

1. *If (16) is satisfied, then both media outlet will hire moderate editors (i.e.,* $x_{e_1} = x_{e_2} = \frac{1}{2}$*).*

2. *If (17) is satisfied then* $\exists C^{Dev} < C^{Max}$ *such that:*

   (a) If $C > C^{Dev}$*,then both media outlet will hire moderate editors (i.e.,* $x_{e_1} = x_{e_2} = \frac{1}{2}$*)*

(b) If $C < C^{Dev}$, then the two media outlets will hire non-moderate editors having symmetric idiosyncratic preferences (i.e., $x_{e_1} = 1 - x_{e_2} \neq \frac{1}{2}$). Moreover, the lower is C, the less moderate editors will be hired in equilibrium.

When (16) holds, despite the fact that by choosing, for example, a rightist editor a media outlet would increase the number of rightist citizens willing to watch its news (i.e., higher marginal rightist viewer), the net effect on the demand of choosing this editor rather than a moderate one would be always negative. Since choosing a less moderate editor also implies choosing an editor who will sample relatively less with respect to a more moderate one, the negative effect on moderate citizens' viewership would be higher than the positive effect on rightist citizens' viewership. Hence, while in a standard Hotelling model of product differentiation, the two firms would locate symmetrically so to capture all the market, in our setting the fact that different editors have different sampling strategies in terms of the degree of asymmetry between them (which would correspond to the product location on the classical Hotelling line) but also with regard to the "intensity" of sampling, prevents such equilibrium to emerge. Hence, the portion of market covered in a duopoly will still be the same as the one of a monopoly.

Moreover, even when (17) holds, if the opportunity cost of acquiring information is high with respect to the mass of non moderate citizens, the two media outlets will both choose moderate editors. This is the only case where a media outlet may not find convenient to choose a non-moderate editor in a duopoly while it would so in a monopoly. The reason behind this difference is that in the monopoly case choosing, for example, a rightist editor instead of a moderate one will decrease the demand of news by leftist citizens in a lower proportion with respect to the increase in the demand of news by rightist citizens. Instead, in the duopoly case, when the opportunity cost of acquiring information is high, by choosing a rightist editor, a media outlet may face a reduction in the demand for its news by *moderate* citizens larger than the increase in the demand by rightist citizens.

On the other hand, when the opportunity cost is low and/or the mass of extremists citizens is large, the demand of news by extremist citizens will be bigger. Hence, the two media outlets will choose specular types of non-moderate editors. That is, while in the monopolistic case there was only a rightist (or leftist) editor, in presence of two media outlets there will be also a leftist (or rightist) editor.

5.3 Multiple Media Outlets

We now analyze the case where there are multiple media outlets, i.e., $K > 2$. The above analysis has shown that when citizens' preferences are distributed uniformly in the policy space, in a duopoly both media outlets will hire moderate editors. The following proposition shows that this is not always the case when there are multiple media outlets in the market for news. More specifically, as the number of media outlets present in the market increases, the equilibrium where every media outlet chooses a moderate editor is not sustainable anymore. Indeed, any media outlet would have an incentive to differentiate its "news product" by choosing a non-moderate editor.

Proposition 4 *Suppose that citizen's idiosyncratic preferences are distributed uniformly in* $[0,1]$. *Then,* $\exists K^* \in (2,\infty)$ *such that for* $K > K^*$ *the set* $\left\{x_{e_j} = \frac{1}{2}, \forall j = 1,...,K\right\}$ *is not anymore an equilibrium. In such case, it will still exist a symmetric mixed-strategy Nash equilibrium. Moreover, if* $\theta > \frac{2}{3}$*:*

i) $\frac{dK^*}{d\delta} < 0$

ii) $\frac{dK^*}{dc} > 0$

iii) $\frac{dK^*}{dC} > 0$

The above proposition shows that when the market for "moderate news" gets crowded, media outlets will prefer to choose a different location for their news product.[18] When citizens care more about the quality of candidates, it is more likely that media outlet will hire non-moderate editors. This result, which may appear counter-intuitive, is due to the fact that a higher δ is associated with more extremists citizens willing to acquire information. Hence, since in presence of a higher δ there is a higher demand for news by extremists citizens media outlets have higher incentives to choose non-moderate editors.

Moreover, since when c is low $\bar{n}_e^*$ and $|\underline{n}_e^*|$ are high, when all media outlets choose a moderate editor (i.e., $x_e^j = \frac{1}{2}, \forall j$) the overall demand for news is high. Hence, there would also be a larger market to be "stolen" by choosing a non-moderate editor. Hence K^* is low when c is low. Instead, when c is high, a less moderate editor may not collect enough information to convince extremists citizen that is worth spending the opportunity cost of watching the media outlet. Hence, choosing a less moderate editor may lead to a small gain in the viewership of extremists citizens and a big loss in the viewership of moderate ones.

On the other hand, since the higher the opportunity cost of acquiring information, the less extremists citizens will find optimal to acquire information, as such cost increases the likelihood of media outlets choosing non-moderate editors decreases.[19] That is, we can also reinterpret the above proposition with respect to C. That is, for a given $K > 2$,

[18] Chan and Suen (2008) also consider a model of rational consumers and profit-maximizing media outlets. However, in their model more competition does not produce different editorial positions.

[19] Indeed $\lim\limits_{C \to C^{Max}} K^* \to \infty$.

there will exist a $C^*(K)$ such that for $C > C^*(K)$, all media outlets will hire a moderate editor from the population of citizens. Instead, for $C < C^*(K)$, media outlets will hire non-moderate editors. This result, along with the ones of propositions 2 and 3, suggests that we should expect more moderate editors to prevail in a news market where the opportunity cost is high. A clear application of such result is thus given by the market for news in the broadcast media sector with respect to the press. The opportunity cost of watching a report from a broadcast media is arguably lower than the one of reading a newspaper. Our analysis thus suggests that we should expect to find more moderate editors in the press than in the broadcast media sector. At the same time, we should expect more extremist citizens watching broadcast media and a higher overall demand for broadcast media with respect to the one faced by the press.

6 Conclusions

We have analyzed a market for news in which profit maximizing media outlets hire their editors from the population of citizens. We have shown that when information acquisition by editors is costly and when citizens have to incur in an opportunity cost to access information, citizens may find optimal to acquire information from a like-minded source of information (i.e., from a media outlet having an editor with similar idiosyncratic preferences). Consequently, a profit maximizing media outlet may prefer to hire a non-moderate editor in order to capture the demand for news of non-moderate citizens.

Hence, even though citizens do not derive any exogenous utility from acquiring biased information and the media outlet is just maximizing profits, the endogenous acquisition of costly information may induce a media outlet to choose an editor whose optimal information acquisition strategy is “slanted” in favor of the alternative *ex-ante* preferred by a subset of citizens. Moreover, the higher the degree of competition in the market for news, the more likely that media outlets may hire non-moderate editors. That is, when the market for news gets crowded, rather than sharing the demand for news of moderate citizens with the other media outlets, a media outlet may prefer to differentiate its news product by choosing a different location in the policy space (i.e., hire an editor with different idiosyncratic characteristics). Hence, our model provides a rationale for the presence of media bias purely based on the citizens' demand for the most valuable source of information. Thus, even though competition brings more media bias in our model (i.e., non-moderate editors), it still has a positive effect on citizens' welfare since it allows a higher portion of population to get informed.[20]

[20] Notice, however, that in a repeated game the effect of competition on welfare is more subtle. The short run polarization of beliefs is going to reinforce the demand for news coming from like-minded sources

Our results also point out that in a market where there is a high opportunity cost of acquiring information, there will be a lower demand for news by non-moderate citizens. Thus, we should expect more moderate editors to be hired by media outlets in such market with respect to a market where the opportunity cost of acquiring information is low. We think that a natural application of this result lies in the differences between the broadcast media industry and the press. According to our model, we should observe more moderate editors in the press sector with respect to the broadcast media sector. Moreover, broadcast media outlets should face a higher demand from extremist citizens (and a higher demand overall) with respect to the one faced by the press.

References

[1] Anderson, S., P., and McLaren, J., (2007) "Media Mergers and Media Bias with Rational Consumers", *mimeo,* University of Virginia.

[2] Aragones E., and Palfrey T., R., (2002) "Mixed Equilibrium in a Downsian Model with a Favored Candidate", *Journal of Economic Theory,* 103: 131-161.

[3] Baron, D., P., (2006) "Persistent Media Bias", *Journal of Public Economics*, 90(1): 1-36.

[4] Bernhardt, D., Krasa, S., and Polborn, M., K., (2008) "Political Polarization and the Electoral Effects of Media Bias", *Journal of Public Economics*, 92: 1092–1104.

[5] Besley, T., and Coate, S., (1997) "An Economic Model of Representative Democracy", *The Quarterly Journal of Economics*, 112(1): 85-114.

[6] Besley T., and Prat A., (2006) "Handcuffs for the grabbing hand? Media capture and government accountability", *American Economic Review*, 96(3): 720-736.

[7] Brocas, I. and Carrillo J.D., (2007) "Influence through Ignorance", *RAND Journal of Economics,* 38(4): 931–947.

[8] Brocas, I. and Carrillo J.D., (2008) "Information acquisition and choice under uncertainty", *Journal of Economics and Management Strategy,* forthcoming.

[9] Burke, J., (2008) "Primetime Spin: Media Bias and Belief Confirming Information", *Journal of Economics and Management Strategy,* 17(3): 633-665.

(see Gentzkow and Shapiro 2006 and Burke 2008). Hence, we may observe a long run polarization of beliefs and hence of choices by different citizens.

[10] Calvert, R. L., (1985) "The Value of Information: A Rational Choice Model of Political Advice." *Journal of Politics,* 47: 530–55.

[11] Chan, J., and Suen, W., (2008) "A Spatial Theory of News Consumption and Electoral Competition", *Review of Economic Studies,* 75(3): 699-728.

[12] Cukierman, A., and Tommasi, M., (1998) "When Does it Take a Nixon to Go to China?", *American Economic Review,* 88(1): 180-197.

[13] Dasgupta, P. and Maskin, E., (1986a) "The Existence of Equilibrium in Discontinuous Economic Games I: Theory", *The Review of Economic Studies,* 53(1): 1-26.

[14] Dasgupta, P. and Maskin, E., (1986b) "The Existence of Equilibrium in Discontinuous Economic Games II: Applications", *The Review of Economic Studies*, 53(1): 27-41.

[15] Della Vigna, S. and Kaplan, E., (2007) "The Fox News Effect: Media Bias and Voting", *The Quarterly Journal of Economics,* 122(3): 1187-1234.

[16] Djankov S., McLiesh C., Nenova T., and Shleifer A., (2003) "Who Owns The Media?" *Journal of Law and Economics*, 46: 341-382.

[17] Durante R. and Knight B., (2009) "Partisan Control, Media Bias, and Viewer Responses: Evidence from Berlusconi's Italy", *NBER Working Paper* 14762.

[18] Ellman, M. and Germano, F., (2008) "What Do the Papers Sell?", *Economic Journal,* forthcoming.

[19] Gentzkow M., and Shapiro J., (2006), "Media Bias and Reputation", *Journal of Political Economy*, 114(2): 280-316.

[20] Gentzkow M., and Shapiro J., (2008), "Competition and Truth in the Market for News", *Journal of Economic Perspectives,* 22(2): 133-154.

[21] Gentzkow M., and Shapiro J., (2009), "What Drives Media Slant? Evidence from U.S. Daily Newspapers", *Econometrica*, forthcoming.

[22] Gerber, A., Karlan, D. S. and Bergan, D., (2008) "Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions", *American Economic Journal: Applied Economics,* forthcoming.

[23] Gilligan, T. W. and Krehbiel K., (1987) "Collective Decisionmaking and Standing Committees: An Informational Rationale for Restrictive Amendment Procedures", *Journal of Law, Economics and Organization* 3(2): 287–335.

[24] Groseclose, T. and Milyo, J., (2005). "A Measure of Media Bias," *The Quarterly Journal of Economics,* 120(4): 1191-1237.

[25] Ho, D. E., and Quinn, K. M., (2008) "Measuring Explicit Political Positions of Media", forthcoming, *Quarterly Journal of Political Science.*

[26] Milgrom, P. and Roberts, J., (1986) "Relying on the Information of Interested Parties", *RAND Journal of Economics,* 17: 18-32.

[27] Mullainathan S., and Shleifer A., (2005), "The Market for News", *American Economic Review*, 95(4): 1031-1053.

[28] Osborne M., J., and Slivinsky A., (1996) "A Model of Political Competition with Citizen-Candidates", *The Quarterly Journal of Economics*, 111(1): 65-96.

[29] Suen, W., (2004). "The Self-Perpetuation of Biased Beliefs", *The Economic Journal,* 114: 337-396

7 Appendix

Proof of Proposition 1

The problem involves analyzing a stochastic process with two absorbing state. More specifically we want to determine the equations characterizing these two absorbing states (i.e., $\underline{n}_i^*$ and $\bar{n}_i^*$). After m draws, given that a current difference in signals in favor of r equal to n, the value function of editor i is given by (9).

First, suppose that the state of the world $s = r$. Then at a given point in time, given a difference in signals in favor of r equal to n, the value function of an editor with idiosyncratic preferences x_i will satisfy the following second order difference equation:

$$V_i^r(n) = \theta V_i^r(n+1) + (1-\theta)V_i^r(n-1) - c$$

where the associated homogenous equation is:

$$\theta y^2 - y + (1-\theta) = 0$$

whose solutions are:

$$y_1 = 1, y_2 = \frac{1-\theta}{\theta}$$

Moreover, since the difference equation is non-homogenous it has also a specific solution of the form $V_i^r(n) = Hn$, thus we should also find a solution of:

$$[\theta H(n+1) - H(n) + (1-\theta)H(n-1)] = c$$

Thus $H = \frac{c}{2\theta-1}$. Hence, the generic solution to this second order equation is:

$$V_i^r(n) = a + b\lambda^n + Hn$$

where $\lambda = \frac{1-\theta}{\theta}$. In order to find the values of a and b we should consider the two terminal conditions given stopping rule $\bar{n}_i$ and $\underline{n}_i$:

$$V_i^r(\bar{n}_i) = \delta(2\mu(\bar{n}_i) - 1) - (1 - x_i) \tag{18}$$

$$V_i^r(\underline{n}_i) = \delta(1 - 2\mu(\underline{n}_i)) - x_i \tag{19}$$

that is (18) represents the utility of editor i when reaching $\bar{n}_i$ signals in favor of state r (where she chooses alternative R). Similarly (19) represents the utility of editor i when reaching $|\underline{n}_i|$ signals in favor of state l (where she chooses alternative L). Thus, given these two terminal conditions we have that:

$$a + b\lambda^{\bar{n}_i} + H\bar{n}_i = \delta\frac{1-\lambda^{\bar{n}_i}}{1+\lambda^{\bar{n}_i}} - (1-x_i)$$

$$a + b\lambda^{\underline{n}_i} + Hn_i = \delta\frac{\lambda^{\underline{n}_i} - 1}{1+\lambda^{\underline{n}_i}} - x_i$$

where $\mu(n) = \frac{1}{1+\lambda^n}$. Thus:

$$\begin{aligned} V_i^r(n, \bar{n}_i, \underline{n}_i) \quad = \quad & \delta\left(\frac{1-\lambda^{\bar{n}_i}}{1+\lambda^{\bar{n}_i}}\right) - (1-x_i) + \frac{(\lambda^n - \lambda^{\bar{n}_i})}{(\lambda^{\underline{n}_i} - \lambda^{\bar{n}_i})} \\ & \left[2\delta\left(\frac{\lambda^{\underline{n}_i}\lambda^{\bar{n}_i} - 1}{(1+\lambda^{\underline{n}_i})(1+\lambda^{\bar{n}_i})}\right) + H(\bar{n}_i - \underline{n}_i) + 1 - 2x_i\right] - H(\bar{n}_i - n) \end{aligned}$$

Similarly, supposing that the state of the world $s = l$, we can derive $V_i^l(n, \bar{n}_i, \underline{n}_i)$:

$$\begin{aligned} V_i^l(n, \bar{n}_i, \underline{n}_i) &= \delta \frac{1-\lambda^{\bar{n}_i}}{1+\lambda^{\bar{n}_i}} - (1 - x_i) + \frac{\lambda^{\underline{n}_i}\left(\lambda^{\bar{n}_i} - \lambda^n\right)}{\lambda^n\left(\lambda^{\underline{n}_i} - \lambda^{\bar{n}_i}\right)} \\ & \left[2\delta \frac{1-\lambda^{\bar{n}_i}\lambda^{\underline{n}_i}}{\left(1+\lambda^{\bar{n}_i}\right)\left(1+\lambda^{\underline{n}_i}\right)} + H(\bar{n}_i - \underline{n}_i) - 1 + 2x_i\right] + H(\bar{n}_i - n) \end{aligned}$$

Thus the expected value of editor i givena difference of signals in favor of r equal to n is:

$$V_i(n, \bar{n}_i, \underline{n}_i) = \frac{1}{1+\lambda^n} V_i^r(n, \bar{n}_i, \underline{n}_i) + \frac{\lambda^n}{1+\lambda^n} V_i^l(n, \bar{n}_i, \underline{n}_i) \tag{20}$$

Therefore, the optimal "rightist" stopping rule $\bar{n}_i^*$ will be the value such that $\frac{\partial V_i}{\partial \bar{n}_i}|_{\bar{n}_i^*} = 0$, that is:

$$\frac{\partial V_i}{\partial \bar{n}_i}|_{\bar{n}_i^*} = \frac{1}{1+\lambda^n}\frac{\partial V_i^r}{\partial \bar{n}_i}|_{\bar{n}_i^*} + \frac{\lambda^n}{1+\lambda^n}\frac{\partial V_i^l}{\partial \bar{n}_i}|_{\bar{n}_i^*} = 0$$

similarly, the optimal "leftist" stopping rule $\underline{n}_i^*$ will be the value such that $\frac{\partial V_i}{\partial \underline{n}_i}|_{\underline{n}_i^*} = 0$, that is:

$$\frac{\partial V_i}{\partial \underline{n}_i}|_{n=\underline{n}_i^*} = \frac{1}{1+\lambda^n}\frac{\partial V_i^r}{\partial \underline{n}_i}|_{\underline{n}_i^*} + \frac{\lambda^n}{1+\lambda^n}\frac{\partial V_i^l}{\partial \underline{n}_i}|_{\underline{n}_i^*} = 0$$

Where it must be always the case that $\underline{n}_i^* < 0$ and $\bar{n}_i^* > 0$.[21] Hence $\bar{n}_i^*$ is defined implicitly by the following equation:

$$\frac{(\ln\lambda)\,\lambda^{\bar{n}_i^*}}{\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}}\left[(2x-1)\left(1+\lambda^{\underline{n}_i^*}\right) - 2\delta(\lambda^{\underline{n}_i^*} - 1) + H\left(\bar{n}_i^* - \underline{n}_i^*\right)\left(\lambda^{\underline{n}_i^*} - 1\right)\right] = H\left(1 - \lambda^{\bar{n}_i^*}\right) \tag{21}$$

similarly $\underline{n}_i^*$ is defined implicitly by the following equation:

$$\frac{(\ln\lambda)\,\lambda^{\underline{n}_i^*}}{\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}}\left[(2x-1)\left(\lambda^{\bar{n}_i^*} + 1\right) + 2\delta\left(1 - \lambda^{\bar{n}_i^*}\right) - H\left(\bar{n}_i^* - \underline{n}_i^*\right)\left(1 - \lambda^{\bar{n}_i^*}\right)\right] = -H\left(\lambda^{\underline{n}_i^*} - 1\right) \tag{22}$$

Notice that the optimal stopping rule $\bar{n}_i^*$ and $\underline{n}_i^*$ *do not* depend on n. That is the optimal stopping rule do not change depending on the realization of the signals.[22] Let's now analyze the comparative statics:

$$\frac{\partial \bar{n}_i^*}{\partial x_i} \propto \left.\frac{\partial^2 V_i}{\partial \bar{n}_i \partial x_i}\right|_{\bar{n}_i^*, \underline{n}_i^*} = 2(\ln\lambda)\,\lambda^{\bar{n}_i^*}\frac{\lambda^{\underline{n}_i^*} + 1}{\left(\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}\right)\left(1 - \lambda^{\bar{n}_i^*}\right)} < 0 \tag{23}$$

Since $\left(\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}\right) > 0$ and $\ln\lambda < 0$. Similarly:

$$\frac{\partial \underline{n}_i^*}{\partial x_i} \propto \left.\frac{\partial^2 V_i}{\partial \underline{n}_i \partial x_i}\right|_{\bar{n}_i^*, \underline{n}_i^*} = 2(\ln\lambda)\frac{\lambda^{\underline{n}_i^*}}{\lambda^{\underline{n}_i^*} - 1}\frac{\lambda^{\bar{n}_i^*} + 1}{\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}} < 0 \tag{24}$$

[21] Suppose not. That is $\underline{n}_i^* > 0$. Thus $\mu(\underline{n}_i^*) > \mu(n = 0) = p$. If $x_i > \frac{1}{2}$, this would imply that $\mu(\underline{n}_i^*) > \hat{\mu}_i$ and thus $\tau_i(\underline{n}_i^*) = R$ which contradicts the definition of $\underline{n}_i^*$. If $x_i < \frac{1}{2}$, then since $n = 0 < \underline{n}_i^*$, this implies that $\tau_i(n = 0) = L$ and thus the voter would never start sampling. A similar proof applies to show that $\bar{n}_i^* > 0$.

[22] A detailed formal derivation of the second order conditions, ensuring that $(\bar{n}_i^*, \underline{n}_i^*)$ is a global maximum, is available upon request to the author.

Therefore $\left|\frac{\partial \bar{n}_i^*}{\partial x_i}\right| > \left|\frac{\partial \underline{n}_i^*}{\partial x_i}\right|$ if and only if:

$$-\frac{2\left(\ln \lambda\right)}{\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}} \left(\frac{(\lambda^{\underline{n}_i^*}\lambda^{\bar{n}_i^*} - 1)(\lambda^{\bar{n}_i^*} + \lambda^{\underline{n}_i^*})}{\left(\lambda^{\underline{n}_i^*} - 1\right)\left(1 - \lambda^{\bar{n}_i^*}\right)} \right) > 0$$

Therefore since:

$$(\lambda^{\underline{n}_i^*}\lambda^{\bar{n}_i^*} - 1) \begin{cases} < 0 \text{ for } x_i < \frac{1}{2} \\ = 0 \text{ for } x_i = \frac{1}{2} \\ > 0 \text{ for } x_i > \frac{1}{2} \end{cases} \tag{25}$$

Then:

$$\left|\frac{\partial \bar{n}_i^*}{\partial x_i}\right| \begin{cases} < \left|\frac{\partial \underline{n}_i^*}{\partial x_i}\right| & \text{for } x_i < \frac{1}{2} \\ = \left|\frac{\partial \underline{n}_i^*}{\partial x_i}\right| & \text{for } x_i = \frac{1}{2} \\ > \left|\frac{\partial \underline{n}_i^*}{\partial x_i}\right| & \text{for } x_i > \frac{1}{2} \end{cases} \tag{26}$$

since for $x_i < \frac{1}{2}$, $\bar{n}_i^* > |\underline{n}_i^*|$ and thus in our notation $\bar{n}_i^* > |\underline{n}_i^*|$ and hence $\lambda^{\underline{n}_i^*}\lambda^{\bar{n}_i^*} < 1$, instead for $x_i = \frac{1}{2}$, $\bar{n}_i^* = |\underline{n}_i^*|$ and thus in our notation $\bar{n}_i^* = |\underline{n}_i^*|$ and hence $\lambda^{\underline{n}_i^*}\lambda^{\bar{n}_i^*} = 1$. Similarly for $x_i > \frac{1}{2}$, $\bar{n}_i^* < |\underline{n}_i^*|$ and thus in our notation $\bar{n}_i^* < |\underline{n}_i^*|$ and hence $\lambda^{\underline{n}_i^*}\lambda^{\bar{n}_i^*} > 1$. Let's now analyze the comparative statics w.r.t. δ

$$\frac{\partial \bar{n}_i^*}{\partial \delta} \propto \left.\frac{\partial^2 V_i}{\partial \bar{n}_i \partial \delta}\right|_{\bar{n}_i^*, \underline{n}_i^*} = -2\left(\ln \lambda\right)\lambda^{\bar{n}_i^*} \frac{\lambda^{\underline{n}_i^*} - 1}{\left(\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}\right)\left(1 - \lambda^{\bar{n}_i^*}\right)} > 0 \tag{27}$$

Similarly:

$$\frac{\partial \underline{n}_i^*}{\partial \delta} \propto \left.\frac{\partial^2 V_i}{\partial \underline{n}_i \partial \delta}\right|_{\bar{n}_i^*, \underline{n}_i^*} = -2\left(\ln \lambda\right)\lambda^{\underline{n}_i^*} \frac{1 - \lambda^{\bar{n}_i^*}}{(\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*})(1 - \lambda^{\underline{n}_i^*})} < 0 \tag{28}$$

That is the higher is δ, the more citizens care about knowing the true state of the world and the more they will sample in both directions. Therefore $\left.\frac{\partial^2 V_i}{\partial \bar{n}_i \partial \delta}\right|_{\bar{n}_i^*, \underline{n}_i^*} > \left.\frac{\partial^2 V_i}{\partial \underline{n}_i \partial \delta}\right|_{\bar{n}_i^*, \underline{n}_i^*}$ if and only if:

$$2\frac{-\ln \lambda}{\left(\lambda^{\underline{n}_i^*} - 1\right)\left(1 - \lambda^{\bar{n}_i^*}\right)}\left(\lambda^{\underline{n}_i^*}\lambda^{\bar{n}_i^*} - 1\right) > 0$$

Hence given (25) then:

$$\frac{\partial \bar{n}_i^*}{\partial \delta} \begin{cases} < \left|\frac{\partial \underline{n}_i^*}{\partial \delta}\right| & \text{for } x_i < \frac{1}{2} \\ = \left|\frac{\partial \underline{n}_i^*}{\partial \delta}\right| & \text{for } x_i = \frac{1}{2} \\ > \left|\frac{\partial \underline{n}_i^*}{\partial \delta}\right| & \text{for } x_i > \frac{1}{2} \end{cases} \tag{29}$$

Let's now analyze the comparative statics w.r.t. c. That is, $\frac{\partial \bar{n}_i^*}{\partial c} \propto \left.\frac{\partial^2 V_i}{\partial \bar{n}_i \partial c}\right|_{\bar{n}_i^*, \underline{n}_i^*}$:

$$\left.\frac{\partial^2 V_i}{\partial \bar{n}_i \partial c}\right|_{\bar{n}_i^*, \underline{n}_i^*} = \frac{1}{2\theta - 1}\left[\left(\ln \lambda\right)\lambda^{\bar{n}_i^*}\left(\bar{n}_i^* - \underline{n}_i^*\right)\frac{\lambda^{\underline{n}_i^*} - 1}{\left(\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}\right)\left(1 - \lambda^{\bar{n}_i^*}\right)} - 1\right] < 0 \tag{30}$$

Similarly $\frac{\partial \underline{n}_i^*}{\partial c} \propto \left.\frac{\partial^2 V_i}{\partial \underline{n}_i \partial c}\right|_{\bar{n}_i^*, \underline{n}_i^*}$

$$\left.\frac{\partial^2 V_i}{\partial \underline{n}_i \partial c}\right|_{\bar{n}_i^*, \underline{n}_i^*} = \frac{1}{2\theta - 1}(1 + \left(\ln \lambda\right)\lambda^{\underline{n}_i^*}\frac{\bar{n}_i^* - \underline{n}_i^*}{\lambda^{\underline{n}_i^*} - 1}\frac{\lambda^{\bar{n}_i^*} - 1}{\left(\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}\right)}) > 0 \tag{31}$$

Therefore $\left.\frac{\partial^2 V_i}{\partial \underline{n}_i \partial c}\right|_{\bar{n}_i^*, \underline{n}_i^*} > \left.\frac{\partial^2 V_i}{\partial \bar{n}_i \partial c}\right|_{\bar{n}_i^*, \underline{n}_i^*}$ if and only if:

$$\frac{(\lambda^{\underline{n}_i^*}\lambda^{\bar{n}_i^*} - 1)\left(\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}\right)}{\left(\lambda^{\underline{n}_i^*} - 1\right)\left(1 - \lambda^{\bar{n}_i^*}\right)} < 0$$

Hence:

$$\left|\frac{\partial \bar{n}_i^*}{\partial c}\right| \begin{cases} < \left|\frac{\partial \underline{n}_i^*}{\partial c}\right| & \text{for } x_i < \frac{1}{2} \\ = \left|\frac{\partial \underline{n}_i^*}{\partial c}\right| & \text{for } x_i = \frac{1}{2} \\ > \left|\frac{\partial \underline{n}_i^*}{\partial c}\right| & \text{for } x_i > \frac{1}{2} \end{cases}$$

Q.E.D.

Proof of Corollary 1

$\frac{\partial \Pr(\tau_i(n)=L|s=r)}{\partial c} > 0, \frac{\partial \Pr(\tau_i(n)=R|s=l)}{\partial c} > 0, \frac{\partial \Pr(\tau_i(n)=L|s=r)}{\partial x_i} < 0, \frac{\partial \Pr(\tau_i(n)=R|s=l)}{\partial x_i} > 0, \frac{\partial \Pr(\tau_i(n)=L|s=r)}{\partial \delta} < 0$ and $\frac{\partial \Pr(\tau_i(n)=R|s=l)}{\partial \delta} < 0$, simply follows from the comparative statics results of Proposition 1. We want instead to show that more moderate citizens have a lower probability of making errors. Using the same methodology of Lemma 1 in Brocas and Carrillo (2007) we can derive the probability of choosing the wrong alternative for a given state of the world, that is:

$$\Pr(\tau_i(n) = L|s = r) = \Pr(\textit{hitting } \underline{n}_i^*|r) = \frac{2\mu(\bar{n}_i^*) - 1}{\mu(\bar{n}_i^*) - \mu(\underline{n}_i^*)}\mu(\underline{n}_i^*)$$

$$\Pr(\tau_i(n) = R|s = l) = \Pr(\textit{hitting } \bar{n}_i^*|l) = \frac{1 - 2\mu(\underline{n}_i^*)}{\mu(\bar{n}_i^*) - \mu(\underline{n}_i^*)}\left[1 - \mu(\bar{n}_i^*)\right]$$

Thus the *ex-ante* probability of making an error is:

$$\Pr(\textit{error}) = \Pr(s = r)\Pr(\tau_i(n) = L|s = r) + \Pr(s = l)\Pr(\tau_i(n) = R|s = l)$$

that is:

$$\Pr(\textit{error}) = \frac{\lambda^{\bar{n}_i^*}(\lambda^{\underline{n}_i^*} - 1) + (1 - \lambda^{\bar{n}_i^*})}{2\left(\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}\right)}$$

Hence:

$$\frac{d\Pr(\textit{error})}{dx_i} = \frac{\partial \Pr(\textit{error})}{\partial \bar{n}_i^*}\frac{\partial \bar{n}_i^*}{\partial x_i} + \frac{\partial \Pr(\textit{error})}{\partial \underline{n}_i^*}\frac{\partial \underline{n}_i^*}{\partial x_i}$$

where

$$\frac{\partial \Pr(\textit{error})}{\partial \bar{n}_i^*} = \frac{1}{2}(\ln\lambda)\,\lambda^{\bar{n}_i^*}\frac{\left(\lambda^{\underline{n}_i^*} - 1\right)^2}{\left(\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}\right)^2} < 0$$

$$\frac{\partial \Pr(\textit{error})}{\partial \underline{n}_i^*} = -\frac{1}{2}(\ln\lambda)\,\lambda^{\underline{n}_i^*}\frac{\left(1 - \lambda^{\bar{n}_i^*}\right)^2}{\left(\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}\right)^2} > 0$$

hence given (23) and (24):

$$\frac{\partial \Pr(\textit{error})}{\partial x_i} = \frac{(\ln\lambda)^2(\lambda^{\underline{n}_i^*}\lambda^{\bar{n}_i^*} - 1)\left[\lambda^{\bar{n}_i^*}\left(\lambda^{2\underline{n}_i^*} - \lambda^{\bar{n}_i^*}\right)(\lambda^{\underline{n}_i^*} - 1) + \lambda^{\underline{n}_i^*}\left(\lambda^{\underline{n}_i^*} - \lambda^{2\bar{n}_i^*}\right)(1 - \lambda^{\bar{n}_i^*})\right]}{\left(\lambda^{\underline{n}_i^*} - 1\right)\left(1 - \lambda^{\bar{n}_i^*}\right)\left(\lambda^{\underline{n}_i^*} - \lambda^{\bar{n}_i^*}\right)^3}$$

Hence given (25) then:

$$\frac{\partial \Pr(error)}{\partial x_i}\begin{cases} <0 \text{ for } x_i<\frac{1}{2} \\ =0 \text{ for } x_i=\frac{1}{2} \\ >0 \text{ for } x_i>\frac{1}{2} \end{cases}$$

Q.E.D.

Proof of Lemma 1

It is immediate to verify $i), ii)$ and $iii)$. Let's focus on $iv)$. Notice that $\frac{\partial \hat{x}_e}{\partial \mu(\underline{n}_e^*)} = C(\frac{(2\mu(\bar{n}_e^*)-1)}{(1-2\mu(\underline{n}_e^*))^2}) > 0$. On the other hand, $\frac{\partial \tilde{x}_e}{\partial \mu(\underline{n}_e^*)} = -2\delta + \frac{C}{2\mu(\bar{n}_e^*)-1} < 0$ if and only if $C < 2\delta\left(2\mu(\bar{n}_e^*)-1\right) = \tilde{C}$. Thus since $\tilde{C} > C^{MAX}$, when *Assumption 1* holds $\frac{\partial \tilde{x}_e}{\partial \mu(\underline{n}_e^*)} < 0$. Thus since $\frac{\partial \mu(\underline{n}_e^*)}{\partial \underline{n}_e^*} > 0$, the result follows. Similarly for $v)$ notice that $\frac{\partial \hat{x}_e}{\partial \mu(\bar{n}_e^*)} = -2\delta + \frac{C}{1-2\mu(\underline{n}_e^*)} < 0$ if and only if $C < \hat{C}$. Thus since $\hat{C} > C^{MAX}$, $\frac{\partial \hat{x}_e}{\partial \mu(\bar{n}_e^*)} < 0$. On the other hand, $\frac{\partial \tilde{x}_e}{\partial \mu(\bar{n}_e^*)} = C(\frac{1-2\mu(\underline{n}_e^*)}{(2\mu(\bar{n}_e^*)-1)^2}) > 0$. Thus, since $\frac{\partial \mu(\bar{n}_e^*)}{\partial \bar{n}_e^*} > 0$, the result follows. For $vi)$ given that:

$$\frac{\partial \Pr(n=\underline{n}_e^*)}{\partial \underline{n}_e^*} = -\left(\ln\lambda\right)\lambda^{\underline{n}_e^*}\frac{1-\lambda^{2\bar{n}_e^*}}{2\left(\lambda^{\underline{n}_e^*}-\lambda^{\bar{n}_e^*}\right)^2} > 0$$

$$\frac{\partial \Pr(n=\underline{n}_e^*)}{\partial \bar{n}_e^*} = -\left(\ln\lambda\right)\lambda^{\bar{n}_e^*}\frac{\lambda^{2\underline{n}_e^*}-1}{2\left(\lambda^{\underline{n}_e^*}-\lambda^{\bar{n}_e^*}\right)^2} > 0$$

Hence since $\frac{\partial \bar{n}_e^*}{\partial x_e} < 0$ and $\frac{\partial \underline{n}_e^*}{\partial x_e} < 0$ then $\frac{\partial \Pr(n=\underline{n}_e^*)}{\partial x_e} < 0$ and since $\Pr(n=\underline{n}_e^*) = 1-\Pr(n=\bar{n}_e^*)$ the result follows. **Q.E.D.**

Proof of Proposition 2

The optimal strategy for a profit maximizing monopolist media outlet is to choose an editor with idiosyncratic preference x_e such that the demand is maximized. That is x_e^{Mon} must be such that:

$$\frac{\partial D}{\partial x_e} = \frac{\partial D}{\partial \bar{n}_e^*}\frac{\partial \bar{n}_e^*}{\partial x_e} + \frac{\partial D}{\partial \underline{n}_e^*}\frac{\partial \underline{n}_e^*}{\partial x_e} = 0$$

Where:

$$\frac{\partial D}{\partial \bar{n}_e^*} = \frac{\partial F(\tilde{x}_e)}{\partial \bar{n}_e^*} - \frac{\partial F(\hat{x}_e)}{\partial \bar{n}_e^*}$$

$$\frac{\partial D}{\partial \underline{n}_e^*} = \frac{\partial F(\tilde{x}_e)}{\partial \underline{n}_e^*} - \frac{\partial F(\hat{x}_e)}{\partial \underline{n}_e^*}$$

where $\frac{\partial F(\tilde{x}_e)}{\partial \bar{n}_e^*} = \frac{\partial}{\partial \bar{n}_e^*}\int_0^{\tilde{x}_e(\bar{n}_e^*)} f(x)dx$. Hence applying Leibniz's rule:

$$\frac{\partial F(\tilde{x}_e)}{\partial \bar{n}_e^*} = \frac{\partial}{\partial \bar{n}_e^*}\int_0^{\tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*)} f(x)dx = f(\tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*))\frac{\partial \tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \bar{n}_e^*}$$

thus,

$$\frac{\partial D}{\partial \bar{n}_e^*} = f(\tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*))\frac{\partial \tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \bar{n}_e^*} - f(\hat{x}_e(\bar{n}_e^*,\underline{n}_e^*))\frac{\partial \hat{x}_e(\bar{n}_e^*)}{\partial \bar{n}_e^*}$$

similarly

$$\frac{\partial D}{\partial \underline{n}_e^*} = f(\tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*))\frac{\partial \tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \underline{n}_e^*} - f(\hat{x}_e(\bar{n}_e^*,\underline{n}_e^*))\frac{\partial \hat{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \underline{n}_e^*}$$

Hence given (23) and (24) the first order condition reduces to:

$$\frac{\partial D}{\partial \bar{n}_e^*}\frac{\partial \bar{n}_e^*}{\partial x_e} = -\frac{\partial D}{\partial \underline{n}_e^*}\frac{\partial \underline{n}_e^*}{\partial x_e}$$

hence:

$$\frac{\left(\frac{\partial \tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \bar{n}_e^*}\frac{\partial \bar{n}_e^*}{\partial x_e}+\frac{\partial \tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \underline{n}_e^*}\frac{\partial \underline{n}_e^*}{\partial x_e}\right)}{\left(\frac{\partial \hat{x}_e(\bar{n}_e^*)}{\partial \bar{n}_e^*}\frac{\partial \bar{n}_e^*}{\partial x_e}+\frac{\partial \hat{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \underline{n}_e^*}\frac{\partial \underline{n}_e^*}{\partial x_e}\right)} = \frac{f(\hat{x}_e(\bar{n}_e^*,\underline{n}_e^*))}{f(\tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*))} \tag{32}$$

where

$$\tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*) = \left(\frac{1}{2}+\delta\frac{\lambda^{\underline{n}_e^*}-1}{\lambda^{\underline{n}_e^*}+1}-C\frac{1}{\lambda^{\underline{n}_e^*}+1}\frac{\lambda^{\underline{n}_e^*}-\lambda^{\bar{n}_e^*}}{1-\lambda^{\bar{n}_e^*}}\right)$$

$$\hat{x}_e(\bar{n}_e^*,\underline{n}_e^*) = \frac{1}{2}-\delta\frac{1}{\lambda^{\bar{n}_e^*}+1}\left(1-\lambda^{\bar{n}_e^*}\right)+C\frac{1}{\lambda^{\bar{n}_e^*}+1}\frac{\lambda^{\underline{n}_e^*}-\lambda^{\bar{n}_e^*}}{\lambda^{\underline{n}_e^*}-1}$$

First of all notice that for $\delta \leq \frac{1}{2}$ it is always the case that $\hat{x}_e(\bar{n}_e^*,\underline{n}_e^*) > 0$ and $\tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*) < 1$. Hence, both sides of the market for news will never be fully covered regardless of the type of editor chosen by a media outlet.[23] Moreover,

$$\frac{\partial \tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \bar{n}_e^*} = C(\ln\lambda)\frac{\lambda^{\bar{n}_e^*}}{\lambda^{\underline{n}_e^*}+1}\frac{1-\lambda^{\underline{n}_e^*}}{\left(1-\lambda^{\bar{n}_e^*}\right)^2} > 0$$

$$\frac{\partial \tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \underline{n}_e^*} = (\ln\lambda)\frac{\lambda^{\underline{n}_e^*}}{\left(\lambda^{\underline{n}_e^*}+1\right)^2}\left(2\delta-C\frac{\lambda^{\bar{n}_e^*}+1}{1-\lambda^{\bar{n}_e^*}}\right) < 0$$

similarly

$$\frac{\partial \hat{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \bar{n}_e^*} = (\ln\lambda)\frac{\lambda^{\bar{n}_e^*}}{\left(\lambda^{\bar{n}_e^*}+1\right)^2}\left(2\delta-C\frac{\lambda^{\underline{n}_e^*}+1}{\left(\lambda^{\underline{n}_e^*}-1\right)}\right) < 0$$

$$\frac{\partial \hat{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \underline{n}_e^*} = -C(\ln\lambda)\frac{\lambda^{\underline{n}_e^*}\left(1-\lambda^{\bar{n}_e^*}\right)}{\left(\lambda^{\underline{n}_e^*}-1\right)^2\left(\lambda^{\bar{n}_e^*}+1\right)} > 0$$

hence

$$\frac{\partial \tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial x_e}-\frac{\partial \hat{x}_e(\bar{n}_e^*)}{\partial x_e} = 2(\ln\lambda)^2\frac{(1-\lambda^{\underline{n}_e^*}\lambda^{\bar{n}_e^*})}{\left(\lambda^{\underline{n}_e^*}-\lambda^{\bar{n}_e^*}\right)}\left[\begin{array}{c} 2\delta\frac{\left(\lambda^{2\bar{n}_e^*}+\lambda^{2\underline{n}_e^*}\right)(1+\lambda^{\underline{n}_e^*}\lambda^{\bar{n}_e^*})+2\lambda^{\underline{n}_e^*}\lambda^{\bar{n}_e^*}(\lambda^{\bar{n}_e^*}+\lambda^{\underline{n}_e^*})}{\left(\lambda^{2\underline{n}_e^*}-1\right)\left(1-\lambda^{2\bar{n}_e^*}\right)} \\ +C\frac{4\left(\lambda^{2\underline{n}_e^*}(1-\lambda^{2\bar{n}_e^*})^2+\lambda^{2\bar{n}_e^*}(1-\lambda^{2\underline{n}_e^*})^2\right)(\lambda^{\underline{n}_e^*}-\lambda^{\bar{n}_e^*})}{\left(1+\lambda^{\bar{n}_e^*}\right)\left(\lambda^{\underline{n}_e^*}+1\right)\left(\lambda^{\underline{n}_e^*}-1\right)^3\left(1-\lambda^{\bar{n}_e^*}\right)^3}\end{array}\right] \tag{33}$$

In other words, for $x_e > \frac{1}{2}$ an increase in x_e increases $\hat{x}_e$ more than $\tilde{x}_e$ (and viceversa for $x_e < \frac{1}{2}$). Hence:

$$\frac{\frac{\partial \tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \bar{n}_e^*}\frac{\partial \bar{n}_e^*}{\partial x_e}+\frac{\partial \tilde{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \underline{n}_e^*}\frac{\partial \underline{n}_e^*}{\partial x_e}}{\frac{\partial \hat{x}_e(\bar{n}_e^*)}{\partial \bar{n}_e^*}\frac{\partial \bar{n}_e^*}{\partial x_e}+\frac{\partial \hat{x}_e(\bar{n}_e^*,\underline{n}_e^*)}{\partial \underline{n}_e^*}\frac{\partial \underline{n}_e^*}{\partial x_e}}\begin{cases} >1 \text{ for } x_e < \frac{1}{2} \\ =1 \text{ for } x_e = \frac{1}{2} \\ <1 \text{ for } x_e > \frac{1}{2}\end{cases}$$

Therefore $x_e = \frac{1}{2}$ is always a stationary point since in such case $\hat{x}_{\frac{1}{2}}(\bar{n}_e^*,\underline{n}_e^*) = 1-\tilde{x}_{\frac{1}{2}}(\bar{n}_e^*,\underline{n}_e^*)$ and thus by the symmetry of the distribution function, $f(\hat{x}_{\frac{1}{2}}(\bar{n}_e^*,\underline{n}_e^*)) = f(\tilde{x}_{\frac{1}{2}}(\bar{n}_e^*,\underline{n}_e^*))$.[24] On the other hand, for $x_{e_R} > \frac{1}{2}$ to be a stationary point it must be the case that $\frac{f(\hat{x}_{e_R}(\bar{n}_e^*,\underline{n}_e^*))}{f(\tilde{x}_{e_R}(\bar{n}_e^*,\underline{n}_e^*))} < 1$. Hence

[23] The same result applies in a model where $\delta \in \mathbb{R}^+$ and *supp* $[f(x)] = \mathbb{R}$.

[24] Notice that by the second order conditions $\left.\frac{\partial^2 D}{\partial^2 x_e}\right|_{x_e=\frac{1}{2}} = 0$.

$f(\hat{x}_{e_R}(\bar{n}_e^*, \underline{n}_e^*)) < f(\tilde{x}_{e_R}(\bar{n}_e^*, \underline{n}_e^*))$. Thus since $\left.\frac{\partial \tilde{x}_e}{\partial x_e}\right|_{x_e=1/2} > 0$, this implies that $\tilde{x}_{e_R}(\bar{n}_e^*, \underline{n}_e^*) > 1 - \hat{x}_{e_R}(\bar{n}_e^*, \underline{n}_e^*)$. Hence it must be the case that f is such that $\left.\frac{\partial f(\tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*))}{\partial x_e}\right|_{x_e=1/2} > 0$. Notice also that since the distribution function f is symmetric around $\frac{1}{2}$, so it must be the demand function. Hence, by symmetry it must exist also a stationary point $x_{e_L} = 1 - x_{e_R}$ such that $\frac{f(\hat{x}_{e_L}(\bar{n}_e^*, \underline{n}_e^*))}{f(\tilde{x}_{e_L}(\bar{n}_e^*, \underline{n}_e^*))} > 1$, that is $\left.\frac{\partial f(\hat{x}(\bar{n}_e^*, \underline{n}_e^*))}{\partial x_e}\right|_{x_e=1/2} < 0$. That is, if $F(x)$ is such that (16) holds there is a unique stationary point at $x_e = \frac{1}{2}$. *Viceversa*, if $F(x)$ is such that (17) holds, we may also have stationary points at $x_e \neq \frac{1}{2}$.

Let's study the nature of such stationary points. Let's focus on $x_e > \frac{1}{2}$. Then $\frac{\partial D}{\partial x_e} > 0$ if and only if:

$$f(\tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*))\frac{\partial \tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial x_e} > f(\hat{x}_e(\bar{n}_e^*, \underline{n}_e^*))\frac{\partial \hat{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial x_e}$$

Where $\frac{\partial \tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial x_e} > 0$ if and only if:

$$C < 2\delta \frac{\lambda^{2\underline{n}_e^*}\left(1-\lambda^{\bar{n}_e^*}\right)^3\left(\lambda^{\bar{n}_e^*}+1\right)}{\left(\lambda^{2\bar{n}_e^*}\left(\lambda^{2\underline{n}_e^*}-1\right)^2+\lambda^{2\underline{n}_e^*}\left(1-\lambda^{2\bar{n}_e^*}\right)^2\right)} = \tilde{C} \tag{34}$$

Moreover:

$$\frac{\partial \tilde{C}}{\partial \bar{n}_e^*} = -(\ln\lambda)\,4\delta \frac{\lambda^{2\underline{n}_e^*+\bar{n}_e^*}\left(1-\lambda^{\bar{n}_e^*}\right)^2\left(\left(2\lambda^{\bar{n}_e^*}+1\right)+\frac{\left(1-\lambda^{2\bar{n}_e^*}\right)\lambda^{\bar{n}_e^*}\left(\left(\lambda^{2\underline{n}_e^*}-1\right)^2+2\lambda^{2\underline{n}_e^*}\left(\lambda^{2\bar{n}_e^*}-1\right)\right)}{\lambda^{2\bar{n}_e^*}\left(\lambda^{2\underline{n}_e^*}-1\right)^2+\lambda^{2\underline{n}_e^*}\left(1-\lambda^{2\bar{n}_e^*}\right)^2}\right)}{\lambda^{2\underline{n}_e^*}\left(\lambda^{2\bar{n}_e^*}-1\right)^2+\lambda^{2\bar{n}_e^*}\left(\lambda^{2\underline{n}_e^*}-1\right)^2} > 0$$

$$\frac{\partial \tilde{C}}{\partial \underline{n}_e^*} = -(\ln\lambda)\,4\delta \frac{\lambda^{2\underline{n}_e^*}\left(\lambda^{\bar{n}_e^*}+1\right)\left(1-\lambda^{\bar{n}_e^*}\right)^3\left(\lambda^{2\bar{n}_e^*}\right)\left(\lambda^{4\underline{n}_e^*}-1\right)}{\left(\lambda^{2\underline{n}_e^*}\left(\lambda^{2\bar{n}_e^*}-1\right)^2+\lambda^{2\bar{n}_e^*}\left(\lambda^{2\underline{n}_e^*}-1\right)^2\right)^2} > 0$$

hence

$$\frac{\partial \tilde{C}}{\partial x_e} = \frac{\partial \tilde{C}}{\partial \bar{n}_e^*}\frac{\partial \bar{n}_e^*}{\partial x_e} + \frac{\partial \tilde{C}}{\partial \underline{n}_e^*}\frac{\partial \underline{n}_e^*}{\partial x_e} < 0 \tag{35}$$

that is:

$$\tilde{C} < \left.\tilde{C}\right|_{x_e=\frac{1}{2}} = \delta\left[\frac{1-\lambda^{\left(\left.\bar{n}_e^*\right|_{x_e=\frac{1}{2}}\right)}}{\lambda^{\left(\left.\bar{n}_e^*\right|_{x_e=\frac{1}{2}}\right)}+1}\right] = \left.C^{Max}\right|_{x_e=\frac{1}{2}}$$

hence $\tilde{C}$ is always lower than $\left.C^{Max}\right|_{x_e=\frac{1}{2}}$. On the other hand:

$$\frac{\partial \hat{x}_e(\bar{n}_e^*)}{\partial x_e} = \frac{2(\ln\lambda)^2}{\left(\lambda^{\underline{n}_e^*}-\lambda^{\bar{n}_e^*}\right)}\left(\left(2\delta - C\frac{\lambda^{\underline{n}_e^*}+1}{\left(\lambda^{\underline{n}_e^*}-1\right)}\right)\left(\lambda^{2\bar{n}_e^*}\frac{\lambda^{\underline{n}_e^*}+1}{\left(1-\lambda^{\bar{n}_e^*}\right)\left(\lambda^{\bar{n}_e^*}+1\right)^2}\right) - C\frac{\lambda^{2\underline{n}_e^*}\left(1-\lambda^{\bar{n}_e^*}\right)}{\left(\lambda^{\underline{n}_e^*}-1\right)^3}\right)$$

therefore $\left(\frac{\partial \hat{x}_e(\bar{n}_e^*)}{\partial \bar{n}_e^*}\frac{\partial \bar{n}_e^*}{\partial x_e} + \frac{\partial \hat{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial \underline{n}_e^*}\frac{\partial \underline{n}_e^*}{\partial x_e}\right) > 0$ if and only if:

$$C < 2\delta\left(\lambda^{2\bar{n}_e^*}\frac{\left(\lambda^{\underline{n}_e^*}+1\right)\left(\lambda^{\underline{n}_e^*}-1\right)^3}{\lambda^{2\bar{n}_e^*}\left(\lambda^{2\underline{n}_e^*}-1\right)^2+\lambda^{2\underline{n}_e^*}\left(1-\lambda^{2\bar{n}_e^*}\right)^2}\right) = \hat{C}$$

where:

$$\frac{\partial \hat{C}}{\partial \bar{n}_e^*} = 4\delta \left(\ln \lambda\right) \frac{\lambda^{2\bar{n}_e^*} \left(\lambda^{\underline{n}_e^*} - 1\right)^3 \left(\lambda^{\underline{n}_e^*} + 1\right) \left(\lambda^{2\underline{n}_e^*} \left(\lambda^{2\bar{n}_e^*} - 1\right)^2 + 2\lambda^{2\underline{n}_e^* + 2\bar{n}_e^*} \left(1 - \lambda^{2\bar{n}_e^*}\right)\right)}{\lambda^{2\underline{n}_e^*} \left(\lambda^{2\bar{n}_e^*} - 1\right)^2 + \lambda^{2\bar{n}_e^*} \left(\lambda^{2\underline{n}_e^*} - 1\right)^2} < 0$$

moreover:

$$\frac{\partial \hat{C}}{\partial \underline{n}_e^*} = 2\delta \left(\ln \lambda\right) \left(\lambda^{\underline{n}_e^*} - 1\right)^2 \frac{\lambda^{\underline{n}_e^* + 2\bar{n}_e^*} \left(2\lambda^{\underline{n}_e^*} \left(\lambda^{2\bar{n}_e^*} - 1\right)^2 \left(\lambda^{\underline{n}_e^*}(1 + \lambda^{\underline{n}_e^*}) + 1\right) + 2\left(\lambda^{2\underline{n}_e^*} - 1\right)^2 \lambda^{2\bar{n}_e^*}\right)}{\left(\lambda^{2\underline{n}_e^*} \left(\lambda^{2\bar{n}_e^*} - 1\right)^2 + \lambda^{2\bar{n}_e^*} \left(\lambda^{2\underline{n}_e^*} - 1\right)^2\right)^2} < 0$$

hence $\frac{\partial \hat{C}}{\partial x_e} = \frac{\partial \hat{C}}{\partial \bar{n}_e^*} \frac{\partial \bar{n}_e^*}{\partial x_e} + \frac{\partial \hat{C}}{\partial \underline{n}_e^*} \frac{\partial \underline{n}_e^*}{\partial x_e} > 0$. Therefore:

$$\hat{C} = 2\delta \left(\lambda^{2\bar{n}_e^*} \frac{\left(\lambda^{\underline{n}_e^*} + 1\right) \left(\lambda^{\underline{n}_e^*} - 1\right)^3}{\lambda^{2\bar{n}_e^*} \left(\lambda^{2\underline{n}_e^*} - 1\right)^2 + \lambda^{2\underline{n}_e^*} \left(1 - \lambda^{2\bar{n}_e^*}\right)^2} \right) > \hat{C}\Big|_{x_e = \frac{1}{2}} = \delta \frac{\left(1 - \lambda^{\bar{n}_e^*}\right)}{\left(1 + \lambda^{\bar{n}_e^*}\right)} = C^{Max}\big|_{x_e = \frac{1}{2}}$$

therefore it is always the case that for $x_e > \frac{1}{2}$, $\frac{\partial \hat{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial x_e} > 0$. Thus since for $C > \tilde{C}$, when $x_e \geq \frac{1}{2}$ an increase in x_e increases $\hat{x}_e$ and decreases $\tilde{x}_e$. Hence, for $C > \tilde{C}$ an increase in x_e decreases the demand for news. Let's now focus on $C < \tilde{C}$ where $\frac{\partial \hat{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial x_e} > 0$ and $\frac{\partial \tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial x_e} > 0$ and analyze the two different cases. Let's now analyze the two different cases separately.

1. Suppose $F(x)$ is such that (16) holds. We show by contradiction that in this case, $x_e = \frac{1}{2}$ is always the global maximum. Suppose not. Then for $x_e > \frac{1}{2}$ it must be the case that $\frac{\partial D}{\partial x_e} > 0$. Hence:

   $$f(\tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)) \left(\frac{\partial \tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial \underline{n}_e^*} \frac{\partial \underline{n}_e^*}{\partial x_e} + \frac{\partial \tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial \bar{n}_e^*} \frac{\partial \bar{n}_e^*}{\partial x_e} \right) > f(\hat{x}_e(\bar{n}_e^*, \underline{n}_e^*)) \left(\frac{\partial \hat{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial \underline{n}_e^*} \frac{\partial \underline{n}_e^*}{\partial x_e} + \frac{\partial \hat{x}_e(\bar{n}_e^*)}{\partial \bar{n}_e^*} \frac{\partial \bar{n}_e^*}{\partial x_e} \right)$$

   From condition (33) we know that $\frac{\partial \tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial x_e} < \frac{\partial \hat{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial x_e}$. Then it must be the case that $f(\tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)) > f(\hat{x}_e(\bar{n}_e^*, \underline{n}_e^*))$ which contradicts the initial assumption. Hence, in this first case it is always the case that $x_e = \frac{1}{2}$ is a global maximum since regardless whether $C < \tilde{C}$ or $C \geq \tilde{C}$, $\frac{\partial D}{\partial x_e} < 0$ for $x_e > \frac{1}{2}$.

2. Suppose $F(x)$ is such that (17) holds. In such case, $x_e = \frac{1}{2}$ cannot be a global maximum since $\frac{\partial D}{\partial x_e} > 0\Big|_{x_e = \frac{1}{2}}$ because:

   $$\left(\frac{\partial \tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial \underline{n}_e^*} \frac{\partial \underline{n}_e^*}{\partial x_e} + \frac{\partial \tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial \bar{n}_e^*} \frac{\partial \bar{n}_e^*}{\partial x_e} \right)\Bigg|_{x_e = \frac{1}{2}} = \left(\frac{\partial \hat{x}_e(\bar{n}_e^*, \underline{n}_e^*)}{\partial \underline{n}_e^*} \frac{\partial \underline{n}_e^*}{\partial x_e} + \frac{\partial \hat{x}_e(\bar{n}_e^*)}{\partial \bar{n}_e^*} \frac{\partial \bar{n}_e^*}{\partial x_e} \right)\Bigg|_{x_e = \frac{1}{2}}$$

   that is an increase in x_e at $\frac{1}{2}$ increases $\tilde{x}_e$ and $\hat{x}_e$ by the same amount. Hence, since for any $x_e > \frac{1}{2}$ in this case $f(\tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)) > f(\hat{x}_e(\bar{n}_e^*, \underline{n}_e^*))$, $\frac{\partial D}{\partial x_e}\Big|_{x_e = \frac{1}{2}} > 0$. Thus the stationary point $x_{e_R}^{Mon} > \frac{1}{2}$ such that (32) is satisfied will be a global maximum on $[\frac{1}{2}, 1]$.That is, in this second case we have that for $\frac{1}{2} \leq x_e < x_{e_R}^{Mon}$:

   $$\frac{\partial}{\partial x_e} \int_{\frac{1}{2}}^{\tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)} f(x)dx > \left| \frac{\partial}{\partial x_e} \int_{\hat{x}_e(\bar{n}_e^*, \underline{n}_e^*)}^{\frac{1}{2}} f(x)dx \right|$$

   However, the editor cannot be too "extremist". Indeed, $\tilde{x}_e$ is increasing in x_e as long as condition (34) is satisfied. Moreover since by (35) as x_e increases $\tilde{C}$ decreases, the demand of news may increase in x_e up to the point where $C = \tilde{C}$. That is the most "extremist" editor such that $\tilde{x}_e$ is increasing in x_e will be the one with preferences x_R satisfying the following condition:

$$\tilde{C}\left(\bar{n}_e^*(x_R), \underline{n}_e^*(x_R)\right) = C \tag{36}$$

or, equivalently, we can write the above condition as:

$$\tilde{x}(x_{e_R}) = \tilde{x}\left(\tilde{C}\Big|_{x_R} = C\right)$$

Hence $C < \tilde{C}\Big|_{x_R}$ is a necessary condition to have:

$$\frac{\partial}{\partial x_e}\int_{\frac{1}{2}}^{\tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)} f(x)dx > \left|\frac{\partial}{\partial x_e}\int_{\hat{x}_e(\bar{n}_e^*, \underline{n}_e^*)}^{\frac{1}{2}} f(x)dx\right|$$

That is, $C < \tilde{C}\Big|_{x_R}$ is a necessary condition that needs to be satisfied for the media outlet being able to increase its demand by choosing a less moderate editor.

By the symmetry of f, choosing an editor with symmetric preferences will also be profit-maximizing. That is, we have two global maxima in this case $x_{e_R}^{Mon}$ and $x_{e_L}^{Mon} = 1 - x_{e_R}^{Mon}$. Hence, an analogous proof applies to show that for $x_{e_L}^{Mon} < x_e \leq \frac{1}{2}$:

$$\left|\frac{\partial}{\partial x_e}\int_{\hat{x}_e(\bar{n}_e^*, \underline{n}_e^*)}^{\frac{1}{2}} f(x)dx\right| > \frac{\partial}{\partial x_e}\int_{\frac{1}{2}}^{\tilde{x}_e(\bar{n}_e^*, \underline{n}_e^*)} f(x)dx$$

Q.E.D.

Proof of Proposition 3

Let's start with the case where (16) holds. We show that in this case the unique equilibrium is such that $x_e^1 = x_e^2 = \frac{1}{2}$. Suppose that media outlet 1 deviates by choosing $x_e^1 > x_e^2 = \frac{1}{2}$. If media outlet one deviates, the indifferent viewer, i.e., the viewer who will be indifferent between watching media outlet 1 and media outlet 2 is the one having preferences x_I such that $U_I(W_1) = U_I(W_2)$. Hence:

$$x_I(\bar{n}_{e_1}^*, \underline{n}_{e_1}^*)\big|_{x_e = x_e^1, x_e^2 = \frac{1}{2}} = \frac{1}{2} + \frac{\delta}{\left(\lambda^{\underline{n}_{e_1}^*}\lambda^{\bar{n}_{e_1}^*} - 1\right)}\left(\frac{(1 - \lambda^{\bar{n}_{e_2}^*})\left(\lambda^{\underline{n}_{e_1}^*} - \lambda^{\bar{n}_{e_1}^*}\right)}{\left(\lambda^{\bar{n}_{e_2}^*} + 1\right)} - \left(\lambda^{\underline{n}_{e_1}^*} - 1\right)\left(1 - \lambda^{\bar{n}_{e_1}^*}\right)\right)$$

Let's now analyze the no-deviation condition. The no-deviation condition is such that $\nexists x_e > \frac{1}{2}$ such that the demand if deviating is higher than the demand if not deviating. The demand if not deviate

$$D^{NDev}(x_e^1) = D^{NDev}\left(x_e^2\right) = \frac{1}{2}\left[F(\tilde{x}|_{x_e = \frac{1}{2}}) - F(\hat{x}|_{x_e = \frac{1}{2}})\right] = \left[F(\tilde{x}|_{x_e = \frac{1}{2}}) - F\left(\frac{1}{2}\right)\right]$$

Instead the demand if deviates is:

$$D^{Dev}(x_e^1) = \left[F(\tilde{x}|_{x_e^1}) - F(x_I|_{x_e^1})\right]$$

Hence given a uniform distribution it must be the case that:

$$x_I(\bar{n}_{e_1}^*, \underline{n}_{e_1}^*)\big|_{x_e = x_e^1, x_e^2 = \frac{1}{2}} - \frac{1}{2} > \tilde{x}|_{x_e^1} - \tilde{x}|_{x_e = \frac{1}{2}}$$

hence no-deviation if and only if:

$$C > C^{THR} = \delta \frac{\left(\lambda^{2\underline{n}^*_{e1}} - 1\right)\left(1 - \lambda^{\bar{n}^*_{e1}}\right)}{\left(\lambda^{\underline{n}^*_{e1}}\lambda^{\bar{n}^*_{e1}} - 1\right)^2} \left(\frac{\left(\lambda^{\underline{n}^*_{e1}} - \lambda^{\bar{n}^*_{e1}}\right)}{\left(\lambda^{\underline{n}^*_{e1}} + 1\right)} - \frac{(1 - \lambda^{\bar{n}^*_{e2}})}{\left(\lambda^{\bar{n}^*_{e2}} + 1\right)}\left(\lambda^{\bar{n}^*_{e1}} + 1\right)\right)$$

where $C^{THR} > 0$ if and only if

$$\frac{\left(\lambda^{\underline{n}^*_{e1}} - \lambda^{\bar{n}^*_{e1}}\right)}{\left(\lambda^{\underline{n}^*_{e1}} + 1\right)\left(\lambda^{\bar{n}^*_{e1}} + 1\right)} > \frac{(1 - \lambda^{\bar{n}^*_{e2}})}{\left(\lambda^{\bar{n}^*_{e2}} + 1\right)}$$

where

$$\frac{\partial}{\partial \underline{n}^*_{e1}}\left(\frac{\left(\lambda^{\underline{n}^*_{e1}} - \lambda^{\bar{n}^*_{e1}}\right)}{\left(\lambda^{\underline{n}^*_{e1}} + 1\right)\left(\lambda^{\bar{n}^*_{e1}} + 1\right)}\right) = (\ln \lambda)\frac{\lambda^{\underline{n}^*_{e1}}}{\left(\lambda^{\underline{n}^*_{e1}} + 1\right)^2} < 0$$

$$\frac{\partial}{\partial \bar{n}^*_{e1}}\left(\frac{\left(\lambda^{\underline{n}^*_{e1}} - \lambda^{\bar{n}^*_{e1}}\right)}{\left(\lambda^{\underline{n}^*_{e1}} + 1\right)\left(\lambda^{\bar{n}^*_{e1}} + 1\right)}\right) = \left(-(\ln \lambda)\frac{\lambda^{\bar{n}^*_{e1}}}{\left(\lambda^{\bar{n}^*_{e1}} + 1\right)^2}\right) > 0$$

Hence $\frac{\partial}{\partial x_e}\left(\frac{\left(\lambda^{\underline{n}^*_{e1}} - \lambda^{\bar{n}^*_{e1}}\right)}{\left(\lambda^{\underline{n}^*_{e1}} + 1\right)\left(\lambda^{\bar{n}^*_{e1}} + 1\right)}\right) > 0$ if and only if:

$$(1 - \lambda^{\underline{n}^*_{e1}}\lambda^{\bar{n}^*_{e1}})\left(2\lambda^{\underline{n}^*_{e1}}\lambda^{\bar{n}^*_{e1}}\left(\lambda^{\bar{n}^*_{e1}} + \lambda^{\underline{n}^*_{e1}}\right) + \left(\lambda^{2\bar{n}^*_{e1}} + \lambda^{2\underline{n}^*_{e1}}\right)(1 + \lambda^{\underline{n}^*_{e1}}\lambda^{\bar{n}^*_{e1}})\right) > 0$$

which can never be the case since $(1 - \lambda^{\underline{n}^*_{e1}}\lambda^{\bar{n}^*_{e1}}) < 0$ for $x_e > \frac{1}{2}$. Hence:

$$\frac{\partial}{\partial x_e}\left(\frac{\left(\lambda^{\underline{n}^*_{e1}} - \lambda^{\bar{n}^*_{e1}}\right)}{\left(\lambda^{\underline{n}^*_{e1}} + 1\right)\left(\lambda^{\bar{n}^*_{e1}} + 1\right)}\right) < 0$$

which implies that:

$$\frac{\left(\lambda^{\underline{n}^*_{e1}} - \lambda^{\bar{n}^*_{e1}}\right)}{\left(\lambda^{\underline{n}^*_{e1}} + 1\right)\left(\lambda^{\bar{n}^*_{e1}} + 1\right)} < \left.\frac{\left(\lambda^{\underline{n}^*_{e1}} - \lambda^{\bar{n}^*_{e1}}\right)}{\left(\lambda^{\underline{n}^*_{e1}} + 1\right)\left(\lambda^{\bar{n}^*_{e1}} + 1\right)}\right|_{x_e = \frac{1}{2}} = \frac{(1 - \lambda^{\bar{n}^*_{e2}})}{\left(\lambda^{\bar{n}^*_{e2}} + 1\right)} \tag{37}$$

hence $C^{THR} < 0$. Therefore, in a duopoly when the distribution of citizens' idiosyncratic preferences is such that (16) holds (and where citizens watch at most one media report), there will never be an incentive to deviate from the equilibrium at $x^1_e = 1 - x^2_e = \frac{1}{2}$. Moreover, notice that this is the unique Nash equilibrium. If the two media outlets choose editors with different preferences, then each of them would clearly have an incentive to deviate by choosing a moderate editor.
Let's now analyze the case where (17) holds. First of all, in order to ensure that there is someone willing to watch media 1 the following condition must be satisfied

$$x_I(\bar{n}^*_{e1}, \underline{n}^*_{e1})\big|_{x_e = x^1_e, x^2_e = \frac{1}{2}} < \tilde{x}(x^1_e)$$

that is:

$$C < \bar{C} = 2\delta \frac{\left(1 - \lambda^{\bar{n}^*_{e_1}}\right)}{\left(\lambda^{\bar{n}^*_{e_2}} + 1\right)} \tag{38}$$

where obviously $\bar{C} > 0$.[25] Let's now analyze the no deviation condition for $C < \bar{C}$. Media outlet 1 will not hire a non-moderate editor as long as:

$$\frac{F(x_I|_{x_e^1})}{F(\tilde{x}|_{x_e^1}) - F(\tilde{x}|_{x_e = \frac{1}{2}})} > \frac{1}{2}$$

Let C^{Duop} be the opportunity cost solving such equation:

$$\frac{F\left(\frac{1}{2} + \frac{\delta}{\left(\lambda^{\underline{n}^*_{e_1}} \lambda^{\bar{n}^*_{e_1}} - 1\right)} \left(\frac{(1 - \lambda^{\bar{n}^*_{e_2}})\left(\lambda^{\underline{n}^*_{e_1}} - \lambda^{\bar{n}^*_{e_1}}\right)}{\left(\lambda^{\bar{n}^*_{e_2}} + 1\right)} - \left(\lambda^{\underline{n}^*_{e_1}} - 1\right)\left(1 - \lambda^{\bar{n}^*_{e_1}}\right)\right)\right)}{F\left(\frac{1}{2} + \delta \frac{\lambda^{\underline{n}^*_{e_1}} - 1}{\lambda^{\underline{n}^*_{e_1}} + 1} - C \frac{1}{\lambda^{\underline{n}^*_{e_1}} + 1} \frac{\lambda^{\underline{n}^*_{e_1}} - \lambda^{\bar{n}^*_{e_1}}}{1 - \lambda^{\bar{n}^*_{e_1}}}\right) - F\left(\frac{1}{2} + \delta \frac{\left(1 - \lambda^{\bar{n}^*_{e_2}}\right)}{\lambda^{\bar{n}^*_{e_2}} + 1} - C\right)} > \frac{1}{2} \tag{39}$$

Now let $C^{Dev} = \min\left\{\bar{C}, C^{Duop}, C^{Max}\right\}$ then for $C \in \left(0, C^{Dev}\right)$ media outlet 1 will have an incentive to deviate.[26] Hence, in such case there is no equilibrium where both media outlets choose a moderate editor. Moreover, C^{Dev} must be lower than C^{Max} since for $C = C^{Max}$ only citizens with $x_i = \frac{1}{2}$ watch news reports and thus firm 1 will never have an incentive to deviate. And for $C > C^{Max}$ none will watch news reports. Let's now show that it can never exist an equilibrium with $x_{e_1} = x_{e_2} \neq \frac{1}{2}$. Suppose the two media outlets choose the same type of non-moderate editors (e.g., $x_{e_1} = x_{e_2} > \frac{1}{2}$). By doing so their demand would be

$$D^1(x_{e_1} = x_{e_2}) = D^2(x_{e_1} = x_{e_2}) = \frac{F(\tilde{x}_{e_1}) - F(\hat{x}_{e_1})}{2}$$

while if media outlet 2 chooses an editor with preferences $x_{e_2} = 1 - x_{e_1}$ its demand would be:

$$D^2(x_{e_2} = 1 - x_{e_1}) = \frac{1}{2} - F(\hat{x}_{e_2})$$

where by symmetry $\hat{x}_{e_2} = 1 - \tilde{x}_{e_1}$. Hence $F(\hat{x}_{e_2}) = 1 - F(\tilde{x}_{e_1})$. Thus no-deviation if and only if:

$$\frac{F(\tilde{x}_{e_1}) - F(\hat{x}_{e_1})}{2} > F(\tilde{x}_{e_1}) - \frac{1}{2}$$

but since $x_{e_1} > \frac{1}{2}$, then $\tilde{x}_{e_1} > 1 - \hat{x}_{e_1}$ and given condition (17) the above condition cannot hold. An analogous proof applies for $x_{e_1} = x_{e_2} < \frac{1}{2}$. Hence for $C \in \left(0, C^{Dev}\right)$ the unique Nash equilibrium is such that $x_{e_1} = 1 - x_{e_2}$. Suppose $x_{e_1} = 1 - x_{e_2} > \frac{1}{2}$.[27] For this to be an equilibrium, the following conditions must be satisfied for media outlet 1, $\forall \varepsilon > 0$:[28]
1) No deviation to the right:

$$F(x_I|_{x_e = x_{e_1} + \varepsilon > 1 - x_{e_2}}) - \frac{1}{2} = F(\tilde{x}|_{x_e = x_{e_1} + \varepsilon > 1 - x_{e_2}}) - F(\tilde{x}|_{x_{e_1} = 1 - x_{e_2}})$$

[25] Notice that $\bar{C} > C^{Max}$ as long as $2\delta \frac{\left(1 - \lambda^{\bar{n}^*_{e_1}}\right)}{\left(\lambda^{\bar{n}^*_{e_2}} + 1\right)} - \delta \frac{\left(1 - \lambda^{\bar{n}^*_{e_2}}\right)}{\left(\lambda^{\bar{n}^*_{e_2}} + 1\right)} > 0$ where a sufficient condition for this to be true is $\theta > \frac{2}{3}$.

[26] Clearly, if $C^{Dev} < 0$, firm 1 will never have an incentive to deviate. We have seen in the previous case that when F is a uniform c.d.f. $C^{Dev} = C^{THR} < 0$.

[27] Obviously $x_{e_1} = 1 - x_{e_2} < \frac{1}{2}$ will also be an equilibrium when $x_{e_1} = 1 - x_{e_2} > \frac{1}{2}$ is an equilibrium.

[28] Symmetric conditions apply for media outlet 2.

2) No deviation to the left:

$$\frac{1}{2} - F(x_I|_{x_e=x_{e_1}-\varepsilon<1-x_{e_2}}) = F(\tilde{x}|_{x_{e_1}=1-x_{e_2}}) - F(\tilde{x}|_{x_e=x_{e_1}-\varepsilon<1-x_{e_2}})$$

hence the above two conditions together imply that for $x_{e_1} = 1 - x_{e_2} > \frac{1}{2}$ to be an equilibrium it must be the case that for $\varepsilon \to 0$:

$$\left.\frac{\partial F(x_I)}{\partial x_e}\right|_{x_{e_1}=1-x_{e_2}} = \left.\frac{\partial F(\tilde{x})}{\partial x_e}\right|_{x_{e_1}=1-x_{e_2}} \tag{40}$$

Given that for $C < C^{Dev}$ the opposite of (39) holds, $\left.\frac{\partial F(x_I)}{\partial x_e}\right|_{x_e=1/2} < \left.\frac{\partial F(\tilde{x})}{\partial x_e}\right|_{x_e=1/2}$. Moreover, taking the derivative of x_I with respect to x_e and then evaluating it in a symmetric equilibrium:

$$\left.\frac{\partial x_I(\bar{n}^*_{e_1}, \underline{n}^*_{e_1})}{\partial x_e}\right|_{x_e^1=1-x_e^2} = \frac{\delta (\ln\lambda)^2 \left[\lambda^{\underline{n}^*_{e_1}} \left(\lambda^{\underline{n}^*_{e_1}} - \lambda^{2\bar{n}^*_{e_1}}\right)(1-\lambda^{\bar{n}^*_{e_1}}) + \lambda^{\bar{n}^*_{e_1}}\left(\lambda^{2\underline{n}^*_{e_1}} - \lambda^{\bar{n}^*_{e_1}}\right)(\lambda^{\underline{n}^*_{e_1}} - 1)\right]}{\left(\lambda^{\underline{n}^*_{e_1}} - 1\right)\left(1-\lambda^{\bar{n}^*_{e_1}}\right)(\lambda^{\underline{n}^*_{e_1}} - \lambda^{\bar{n}^*_{e_1}})^2} > 0 \tag{41}$$

hence given (17), $\left.\frac{\partial F(x_I)}{\partial x_e}\right|_{x_{e_1}=1-x_{e_2}} > 0$. On the other hand, we know from the proof of proposition 2 that $\tilde{x}_e$ is increasing in x_e only as long as $x_e < x_R$ where x_R is the solution of (36). Hence, given (17):

$$\left.\frac{\partial F(\tilde{x})}{\partial x_e}\right|_{x_{e_1}=1-x_{e_2}} \begin{cases} > 0 \text{ for } x_e < x_R \\ < 0 \text{ for } x_e > x_R \end{cases}$$

Hence, it will always exist a $x^*_{e_1} = 1 - x^*_{e_2} < x_R$ such that condition (40) is satisfied. Moreover, since as C decreases $\frac{\partial \tilde{x}_e(\bar{n}^*_{e_1}, \underline{n}^*_{e_1})}{\partial x_e}$ increases, then when C is lower the RHS of (40) increases. Hence also the LHS of (40) must increase. Hence, given (41), to increase the LHS of (40) x_e must increase. That is, a lower C is associated with an equilibrium where the two media outlets choose less moderate editors.

Q.E.D.

Proof of Proposition 4

We have to analyze the no-deviation condition with K media outlets. Let $\bar{n}^*_e = -\underline{n}^*_e$ be the stopping thresholds chosen by a moderate editor. The demand media outlet 1 faces if it hires a moderate editor as all the other media outlets is $\forall j \in \{2, 3, ..., K\}$:

$$D^{NDev}(x_e^1) = D^{NDev}\left(x_e^j\right) = \frac{1}{K}\left[F(\tilde{x}|_{x_e=\frac{1}{2}}) - F(\hat{x}|_{x_e=\frac{1}{2}})\right] = \frac{2}{K}\left[F(\tilde{x}|_{x_e=\frac{1}{2}}) - F(\frac{1}{2})\right]$$

Instead the demand that media outlet 1 faces if it deviates from such position is:

$$D^{Dev}(x_e^1) = \left[F(\tilde{x}|_{x_e^1}) - F(x_I|_{x_e^1})\right]$$

Hence given a uniform distribution, media outlet 1 will prefer not to hire a moderate editor if and only if:

$$\tilde{x}|_{x_e^1} - x_I|_{x_e^1} > \frac{2}{K}\left[\tilde{x}|_{x_e=\frac{1}{2}} - \frac{1}{2}\right]$$

hence:

$$K > K^* = \frac{\frac{2}{\left(\lambda^{\underline{n}^*_{e_1}} - \lambda^{\bar{n}^*_{e_1}}\right)}\left[\delta\frac{\left(1-\lambda^{\bar{n}^*_e}\right)}{\lambda^{\bar{n}^*_e}+1} - C\right]}{\delta\frac{1}{\left(\lambda^{\underline{n}^*_{e_1}}\lambda^{\bar{n}^*_{e_1}}-1\right)}\left(\frac{\left(\lambda^{\underline{n}^*_{e_1}}-1\right)}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)} - \frac{(1-\lambda^{\bar{n}^*_e})}{\left(\lambda^{\bar{n}^*_e}+1\right)}\right) - C\frac{1}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(1-\lambda^{\bar{n}^*_{e_1}}\right)}}$$

where the denominator is positive as long as $C < \bar{C}$ where $\bar{C}$ is given by (38). Hence, since we have shown in the proof that $\bar{C} > C^{Max}$, then K^* is always positive. Moreover, our game satisfies the properties of Theorem 4 in Dasgupta and Maskin (1986b) for the existence of an equilibrium in a product competition game. Hence, the K^* media outlets game possesses a symmetric mixed-strategy Nash equilibrium. Let's perform some comparative statics.

$$\begin{aligned}\frac{\partial K^*}{\partial \bar{n}^*_{e_1}} &= \frac{2(\ln\lambda)\lambda^{\bar{n}^*_{e_1}}\left(\frac{\delta}{\lambda^{\bar{n}^*_e}+1}\left(1-\lambda^{\bar{n}^*_e}\right)-C\right)}{\left(\frac{\delta\left(\frac{\lambda^{\underline{n}^*_{e_1}}-1}{\lambda^{\underline{n}^*_{e_1}}+1}+\frac{\lambda^{\bar{n}^*_e}-1}{\lambda^{\bar{n}^*_e}+1}\right)}{\lambda^{\underline{n}^*_{e_1}+\bar{n}^*_{e_1}}-1} - \frac{C}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(1-\lambda^{\bar{n}^*_{e_1}}\right)}\right)\left(\lambda^{\underline{n}^*_{e_1}}-\lambda^{\bar{n}^*_{e_1}}\right)} \\ &\left(\frac{1}{\left(\lambda^{\underline{n}^*_{e_1}}-\lambda^{\bar{n}^*_{e_1}}\right)} + \frac{\left(\delta\frac{\lambda^{\underline{n}^*_{e_1}}\left(\frac{\lambda^{\underline{n}^*_{e_1}}-1}{\lambda^{\underline{n}^*_{e_1}}+1}+\frac{\lambda^{\bar{n}^*_e}-1}{\lambda^{\bar{n}^*_e}+1}\right)}{\left(\lambda^{\underline{n}^*_{e_1}+m}-1\right)^2} + \frac{C}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(\lambda^{\bar{n}^*_{e_1}}-1\right)^2}\right)}{\left(\delta\frac{\left(\frac{\left(\lambda^{\underline{n}^*_{e_1}}-1\right)}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)}-\frac{(1-\lambda^{\bar{n}^*_e})}{\left(\lambda^{\bar{n}^*_e}+1\right)}\right)}{\left(\lambda^{\underline{n}^*_{e_1}}\lambda^{\bar{n}^*_{e_1}}-1\right)} - \frac{C}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(1-\lambda^{\bar{n}^*_{e_1}}\right)}\right)}\right)\end{aligned}$$

hence $\frac{\partial K^*}{d\bar{n}^*_{e_1}} < 0$. On the other hand:

$$\begin{aligned}\frac{\partial K^*}{\partial \underline{n}^*_{e_1}} &= \frac{\left(\frac{-(\ln\lambda)2\lambda^{\underline{n}^*_{e_1}}}{\left(\lambda^{\underline{n}^*_{e_1}}-\lambda^{\bar{n}^*_{e_1}}\right)}\left(\delta\frac{\left(1-\lambda^{\bar{n}^*_e}\right)}{\lambda^{\bar{n}^*_e}+1}-C\right)\right)}{\left(\delta\frac{\left(\frac{\left(\lambda^{\underline{n}^*_{e_1}}-1\right)}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)}-\frac{(1-\lambda^{\bar{n}^*_e})}{\left(\lambda^{\bar{n}^*_e}+1\right)}\right)}{\left(\lambda^{\underline{n}^*_{e_1}}\lambda^{\bar{n}^*_{e_1}}-1\right)} - C\frac{1}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(1-\lambda^{\bar{n}^*_{e_1}}\right)}\right)^2} \\ &\left(\frac{\left(\lambda^{\bar{n}^*_e}\lambda^{\bar{n}^*_{e_1}}+\lambda^{\bar{n}^*_{e_1}}(1-\lambda^{\bar{n}^*_{e_1}})+\lambda^{2\underline{n}^*_{e_1}}\left(\lambda^{\bar{n}^*_{e_1}}-\lambda^{\bar{n}^*_e}\right)+\lambda^{\underline{n}^*_{e_1}}\lambda^{\bar{n}^*_e}\left(\lambda^{\bar{n}^*_{e_1}}\lambda^{\underline{n}^*_{e_1}}-2\right)+\lambda^{\underline{n}^*_{e_1}}\lambda^{2\bar{n}^*_{e_1}}\left(\lambda^{\bar{n}^*_e}\lambda^{\underline{n}^*_{e_1}}-2\right)+1\right)}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)^2\left(\lambda^{\bar{n}^*_e}+1\right)\left(\lambda^{\underline{n}^*_{e_1}}-\lambda^{\bar{n}^*_{e_1}}\right)\left(\lambda^{\underline{n}^*_{e_1}+\bar{n}^*_{e_1}}-1\right)} - \frac{C\left(\frac{\left(\lambda^{\bar{n}^*_{e_1}}+1\right)}{\left(\lambda^{\underline{n}^*_{e_1}}-\lambda^{\bar{n}^*_{e_1}}\right)\left(\lambda^{\underline{n}^*_{e_1}}+1\right)}\right)}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(1-\lambda^{\bar{n}^*_{e_1}}\right)}\right)\end{aligned}$$

hence $\frac{\partial K^*}{\partial \underline{n}^*_{e_1}} > 0$, $\forall C \in (0, C^{Max})$ and $\theta > 2/3$.[29] Moreover:

$$\frac{\partial K^*}{\partial \delta} = -\frac{2C\left(\frac{\left(\lambda^{\bar{n}^*_e}+\left(1-\lambda^{\bar{n}^*_{e_1}}\right)+\lambda^{\underline{n}^*_{e_1}}(\lambda^{\bar{n}^*_{e_1}}-2\lambda^{\bar{n}^*_e})+\lambda^{\bar{n}^*_{e_1}}\left(\lambda^{\underline{n}^*_{e_1}}\lambda^{\bar{n}^*_e}-1\right)\right)}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(1-\lambda^{\bar{n}^*_{e_1}}\right)}\right)}{\left(\lambda^{\underline{n}^*_{e_1}}-\lambda^{\bar{n}^*_{e_1}}\right)\left(\lambda^{\bar{n}^*_e}+1\right)\left(\lambda^{\underline{n}^*_{e_1}+\bar{n}^*_{e_1}}-1\right)\left(\frac{\delta\left(\frac{\lambda^{\underline{n}^*_{e_1}}-1}{\lambda^{\underline{n}^*_{e_1}}+1}+\frac{\lambda^{\bar{n}^*_e}-1}{\lambda^{\bar{n}^*_e}+1}\right)}{\lambda^{\underline{n}^*_{e_1}+\bar{n}^*_{e_1}}-1}-\frac{C}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(1-\lambda^{\bar{n}^*_{e_1}}\right)}\right)^2} < 0$$

since $\left(\lambda^{\bar{n}^*_e}+\left(1-\lambda^{\bar{n}^*_{e_1}}\right)+\lambda^{\underline{n}^*_{e_1}}(\lambda^{\bar{n}^*_{e_1}}-2\lambda^{\bar{n}^*_e})+\lambda^{\bar{n}^*_{e_1}}\left(\lambda^{\underline{n}^*_{e_1}}\lambda^{\bar{n}^*_e}-1\right)\right) > 0$.Hence

$$\frac{dK^*}{d\delta} = \frac{\partial K^*}{\partial \delta} + \frac{\partial K^*}{\partial \bar{n}^*_{e_1}}\frac{\partial \bar{n}^*_{e_1}}{\partial \delta} + \frac{\partial K^*}{\partial \underline{n}^*_{e_1}}\frac{\partial \underline{n}^*_{e_1}}{\partial \delta} < 0$$

since $\frac{\partial K^*}{\partial \delta} < 0$, $\frac{\partial K^*}{\partial \bar{n}^*_{e_1}} < 0$, $\frac{\partial K^*}{\partial \underline{n}^*_{e_1}} > 0$, $\frac{\partial \bar{n}^*_{e_1}}{\partial \delta} > 0$ and $\frac{\partial \underline{n}^*_{e_1}}{\partial \delta} < 0$.Hence as δ increases K^* decreases. Moreover:

$$\frac{dK^*}{dc} = \frac{\partial K^*}{\partial \bar{n}^*_{e_1}}\frac{\partial \bar{n}^*_{e_1}}{\partial c} + \frac{\partial K^*}{\partial \underline{n}^*_{e_1}}\frac{\partial \underline{n}^*_{e_1}}{\partial c}$$

hence since we have shown above that $\frac{\partial K^*}{\partial \bar{n}^*_{e_1}} < 0$, $\frac{\partial K^*}{\partial \underline{n}^*_{e_1}} > 0$ and since $\frac{\partial \bar{n}^*_{e_1}}{\partial c} < 0$ and $\frac{\partial \underline{n}^*_{e_1}}{\partial c} > 0$, then:

$$\frac{dK^*}{dc} > 0$$

Moreover,

$$\frac{dK^*}{dC} = \frac{2\left(\frac{\frac{\delta}{\lambda^{\bar{n}^*_e}+1}\left(1-\lambda^{\bar{n}^*_e}\right)-C}{\left(1-\lambda^{\bar{n}^*_{e_1}}\right)\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(\delta\frac{1}{\left(\lambda^{\underline{n}^*_{e_1}}\lambda^{\bar{n}^*_{e_1}}-1\right)}\left(\frac{\left(\lambda^{\underline{n}^*_{e_1}}-1\right)}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)}-\frac{(1-\lambda^{\bar{n}^*_e})}{\left(\lambda^{\bar{n}^*_e}+1\right)}\right)-C\frac{1}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(1-\lambda^{\bar{n}^*_{e_1}}\right)}\right)}-1\right)}{\left(\lambda^{\underline{n}^*_{e_1}}-\lambda^{\bar{n}^*_{e_1}}\right)\left(\frac{\delta}{\lambda^{\underline{n}^*_{e_1}+\bar{n}^*_{e_1}}-1}\left(\frac{\lambda^{\underline{n}^*_{e_1}}-1}{\lambda^{\underline{n}^*_{e_1}}+1}+\frac{\lambda^{\bar{n}^*_e}-1}{\lambda^{\bar{n}^*_e}+1}\right)-\frac{C}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(1-\lambda^{\bar{n}^*_{e_1}}\right)}\right)}$$

hence the above is positive if and only if:

$$\left(\frac{\left(1-\lambda^{\bar{n}^*_e}\right)}{\lambda^{\bar{n}^*_e}+1} - \frac{\left(1-\lambda^{\bar{n}^*_{e_1}}\right)\left(\lambda^{\underline{n}^*_{e_1}}-1\right)}{\left(\lambda^{\underline{n}^*_{e_1}}-\lambda^{\bar{n}^*_{e_1}}\right)}\right) > 0$$

which is always true since:

$$\frac{\left(1-\lambda^{\bar{n}^*_{e_1}}\right)\left(\lambda^{\underline{n}^*_{e_1}}-1\right)}{\left(\lambda^{\underline{n}^*_{e_1}}-\lambda^{\bar{n}^*_{e_1}}\right)} - \frac{\left(\lambda^{\underline{n}^*_{e_1}}-\lambda^{\bar{n}^*_{e_1}}\right)}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(\lambda^{\bar{n}^*_{e_1}}+1\right)} = -(1-\lambda^{\underline{n}^*_{e_1}}\lambda^{\bar{n}^*_{e_1}})^2 < 0$$

and $\frac{\left(\lambda^{\underline{n}^*_{e_1}}-\lambda^{\bar{n}^*_{e_1}}\right)}{\left(\lambda^{\underline{n}^*_{e_1}}+1\right)\left(\lambda^{\bar{n}^*_{e_1}}+1\right)} < \frac{(1-\lambda^{\bar{n}^*_e})}{(\lambda^{\bar{n}^*_e}+1)}$ by condition (37). Thus $\frac{dK^*}{dC} > 0$. **Q.E.D.**

[29] A detailed formal proof showing that $\theta > 2/3$ is a sufficient condition for $\frac{\partial K^*}{\partial \underline{n}^*_{e_1}} > 0$ is available upon request to the author.