The Instrumental Voter Goes to the News-Agent: Demand For Information, Election Closeness and the Media

Valentino Larcinese
Department of Government and STICERD
London School of Economics and Political Science

November 18, 2001
(Preliminary draft)

Abstract

This paper has two main purposes: first, it studies the impact of election closeness on political information demand and supply; second, it provides evidence on mass media behaviour during electoral campaigns. The theoretical part presents a model of information demand and supply. If voters act instrumentally then information demand should increase with the election closeness. Mass media are modeled as profit-maximizing firms that take into account information demand, the value of customers to advertisers and the marginal cost of customers. Information supply should be larger in constituencies where the contest is expected to be closer, there is an higher population density, and population is on average more valuable to advertisers (for example richer, better educated etc.). The impact of electorate size is theoretically undetermined. These conclusions are then tested with comfortable results on data for the 1997 general election in Britain.

1 Introduction and related literature

The literature explaining voting and elections in terms of rational choice is very vast. A common feature of different rational choice models is that voting is considered a mean to achieve public policy ends: this behavior is called instrumental voting. Citizens care about public policies and voting is seen as an instrument to reach those, or at least to increase the probability to get what is preferred.

This theory poses several problems, including the fact that the probability to be pivotal in large elections is normally so low that it could be considered negligibile in optimization processes. However, this criticism can be overcome if we are ready to compromise on what we intend by a rational act. In a weak sense, agents behave rationally according to their perception of the reality, that could be different from the "objective" state of facts. In the case of voting, the probability to be pivotal in a large election is clearly very low, but it is not zero. The subjective perception of the probability of casting a decisive vote does not necessarily coincide with the infinitesimal numbers that appropriate but cumbersome calculations would deliver. Moreover, voting has been described as a "low cost-low benefit" activity (see Aldrich, 1993): it is therefore possible that even small changes in this probability might have an effect on incentives to participate in election.

If we accept these reasonings then turnout should be larger in closer elections. Unfortunately empirical analysis does not deliver any firm conclusion [quote and examples...].

In this paper we will consider another implication of instrumental theories of voting: when elections are closer then information on candidates and platforms should be more valuable as the probability for a vote to matter is higher. Thus, if the suppliers of political information (mass media) are aware of this then we should expect their behavior to be influenced by election closeness. In this sense, mass media behaviour under different circumstances will give us the possibility to provide a different kind of test of instrumental theories of voting and of the role of election closeness as an incentive for participation. This clearly allows us to exploit information not used so far for this purpose.

Indeed, most people seem to believe that mass media have a vast impact on citizens' electoral choices. Politicians appear to struggle for media attention and tend to complain when they do not receive enough space on newspapers or television. Some politicians blame the media for bad electoral performances. In some countries access to television and electoral advertising during electoral campaigns are publicly regulated. All this must be based on the presumption that media are effective in influencing voters' behaviour.

Studies in this sense have not delivered any conclusive evidence, both because of an objective lack of data and because of the difficulty to identify the media effects in reality. In particular, the effects of media could potentially be of several different types. At the very minimum, media can be seen as informing the citizens about the different available options on the political market. Some theories, however, give to mass media more power than this and see them as capable of persuading the people by shaping their policy preferences. Finally, on the other side of the spectrum, it is conceivable that even the general values and principles of a community can be affected by media choices about what is worth to report, which aspects to highlight and the way news are delivered.

Rational choice theory, at least as long as the assumption of given preferences is mantained, seems clearly better compatible with the first hypothesis, that can be dated bach to the so-called theory of "uses and gratifications" (Blumler-McQuail (1968)). After a period of intense studies about "the media effects", this theory called for a change of perspective by asking rather what the people do with the media. According to this theory, citizens can be seen as active users of media, with preferences, expectations and demands. Following this approach a number of studies have tried to understand why and how agents acquire and process political information, and what is the role of personal characteristics (education, prior knowledge of political matters, interest etc.) in this process (Lau-Sears (1986), Ferejohn-Kuklinski (1990), Zaller (1992)).

From an economist's standpoint this calls for a formal analysis of the political information market, with a demand for information that comes from people's preferences and a supply of information provided, among others, by media firms. Economic studies of political information demand include Matsusaka (1995) and Larcinese (2000). Specific characteristics of the media industry (like economies of scale and concentration) can also be expected to matter for the way people are informed about politics. Works in this direction are Spence-Owen (1977) and Noam (1987).

The degree of citizens' awareness of political matters may in turn be expected to affect public policy making, both on efficiency and redistributive grounds. Examples of works in this sense are Grossman-Helpman (1996), Lohmann (1998) and Larcinese (1999). Media's impact in elections and implications for public policies have been studied by Stromberg (2001a and 2001b), Besley-Burgess (2001) and Besley-Prat (2001). [maybe extend this part]

The theoretical model presented in this paper builds on Stromberg (2001a) model of mass media competition. Stromberg argues that "the increasing-return-to-scale technology and advertising financing of media firms induce them to provide more news to large groups, such as tax payers and dispersed consumer interests, and groups that are valuable to advertisers". Eventually, this information bias will be taken into account by politicians when proposing electoral platforms and will therefore translate into a policy bias.

By identifying groups with electoral constituencies we will be able to implement a test of Stromberg's conclusions. One of his results concerns the effect of groups size on news supply. Larger groups should receive more media attention as they provide more readership and therefore more revenue. The same can be said of groups that are more valuable to advertisers (for example wealthier groups). However, in the context of our model a countervailing effect can be identified: in larger groups we should expect a more severe collective action problem. Thus, in larger constituencies the probability to cast a decisive vote is smaller and such will be the demand for information. This "collective action effect" can potentially offset the "group size effect"; only empirical investigation can shed further light and allow us to accept or reject Stromberg's theoretical result.

We will also consider the newspapers production function and the role of delivery costs: we can argue that more densely populated areas will receive more news coverage (other things equal) simply because the cost of the marginal reader is lower in such ares.

The paper can be summarized as follows. In the next section we will present the theoretical model of information demand and supply. Political information can be demanded for a number of reasons, including instrumental voting. Thus, it will be higher in marginal constituencies. This higher demand will induce a larger supply by profit-maximizing media firms. Media's revenue per reader is represented by the price payed for the newspaper plus the amount payed by advertisers per reader. This amount is not the same for all customers and can be expected to be higher for customers that are more valuable to advertisers (e.g. richer, better educated etc.). The cost of producing newspapers is fixed but there is a variable delivery cost. Thus, in equilibrium information supply is higher in marginal constituencies as well as in constituencies with richer and more concentrated electorate. About the size of the electorate we identify the "group size effect" and the "collective action effect". In section 3 these predictions are tested for the 1997 British general election, using data collected from a national newspaper during the electoral campaign, as well as electoral data and the 1991 Census. The results suggest a high degree of compatibility between our theory and the data. Section 4 concludes.

2 The model

Let's consider a polity divided into two electoral constituencies M and Q. Each constituency elects a member of parliament (MP). There are two competing parties L and R each presenting one candidate in all constituencies. MPs are elected in a first past the post system. With obvious notation we will indicate the candidates in each costituency with L_M, R_M, L_Q, R_Q.

Suppose the two candidates in each constituency are chosen independently by parties through a process that is unknown to citizens. This process can be represented for both parties by respective distribution functions $F_R(\theta)$ and $F_L(\theta)$ (with densities $f_L(\theta)$ and $f_R(\theta)$) over the support Θ of candidates types.

For simplicity we will assume that the policy space is the same as the candidates' space and, with some notational abuse, that utility from policy θ is θ. Policies are formed at the central level by the parliament of the two MPs and affect both constituencies. If θ_M is the candidate elected in constituency M and θ_Q is elected in constituency Q, then the central policy will be $\theta^* = \frac{1}{2}\theta_M + \frac{1}{2}\theta_Q$.

Let's consider the decision problem faced by citizens in constituency M. Given two candidates θ_{M1} and θ_{M2} the net benefit from electing the preferred one rather than the other is given by

$$\begin{aligned} B(\theta_{M1}, \theta_{M2}|\theta_Q) &= |(\frac{1}{2}\theta_Q + \frac{1}{2}\theta_{M1}) - (\frac{1}{2}\theta_Q + \frac{1}{2}\theta_{M2})| \\ &= \frac{1}{2}|\theta_{M1} - \theta_{M2}|. \end{aligned}$$

Analogously

$$B(\theta_{Q1}, \theta_{Q2}|\theta_M) = \frac{1}{2}|\theta_{Q1} - \theta_{Q2}|$$

We also assume that M is marginal and this is common knowledge; i.e., if we indicate with $P_i (i = M, Q)$ the (common) prior probability that a vote will result decisive, each agent believes that $P_M > P_Q$. We can think of this probabilities as coming from different prior beliefs about the candidates in the two constituencies. For example in constituency M the distribution functions $F_R(\theta)$ and $F_L(\theta)$ are "more similar" than in Q. However, also the population size in each constituency will clearly play a role as a larger electorate, with given priors, will reduce the probability of each single vote to be pivotal. With only two constituencies and given electorate this effect cannot be fully appreciated in the model. However, this "collective action effect" can be expected to play a role and will be considered in the empirical investigation. A trivial way to consider this effect is to write $P_i = P_i(S_i)$ where S_i is the size of constituency i.

2.1 Information demand

To avoid cumbersome notation we will focus on a generic constituency. Citizens utility from voting when types are known is then $V(\theta_L, \theta_R) = P_i B(\theta_L, \theta_R)$. However, the expected utility from an informed voting choice before candidates are selected is given by

$$V^* = P_i \int \int B(\theta_L, \theta_R) dF_L(\theta) dF_R(\theta). \; i = M, Q$$

For simplicity, and without loss of generality, here we will assume that there is no cost of voting.

We assume voters are ex ante uninformed about candidates. We will indicate the expected utility from uninformed voting as $\widetilde{V}$. We can then define the ex post utility of an informed vote versus an uninformed one as

$$\Delta V = V(\theta_L, \theta_R) - \widetilde{V}$$

Before gathering information, however, the candidate types are unknown. Thus, the ex ante utility of gathering information is

$$\Delta v = \int \int [V(\theta_L, \theta_R) - \widetilde{V}] dF_L(\theta) dF_R(\theta)$$

Lemma 1 $\Delta v = V^* - \widetilde{V} \geq 0$.

Proof.
Let's introduce the following notation

$$\begin{aligned} \mathcal{A} &= \left\{ \theta_L, \theta_R \text{ s.t. } P_i \int \int B(\theta_L, \theta_R) dF_L(\theta) dF_R(\theta) > 0 \right\}, \; i = M, Q \\ \overline{\mathcal{A}} &= \left\{ \theta_L, \theta_R \text{ s.t. } P_i \int \int B(\theta_L, \theta_R) dF_L(\theta) dF_R(\theta) < 0 \right\}, \; i = M, Q \end{aligned}$$

Suppose now that $F_L(\theta)$ and $F_R(\theta)$ are s.t. candidate L is preferred, i.e.

$$\int \int B(\theta_L, \theta_R) dF_L(\theta) dF_R(\theta) > 0$$

An uninformed voter in this case votes for candidate L. However, her ex ante utility is

$$\begin{aligned} \widetilde{V} &= P_i \int\int_{\mathcal{A}} B(\theta_L, \theta_R) dF_L(\theta) dF_R(\theta) - \\ P_i \int\int_{\overline{\mathcal{A}}} B(\theta_L, \theta_R) dF_L(\theta) dF_R(\theta),\ i &= M, Q \end{aligned}$$

The ex ante (i.e. before knowing the realization of candidates) utility of an informed voter is instead

$$\begin{aligned} V^* &= P_i \int\int_{\mathcal{A}} B(\theta_L, \theta_R) dF_L(\theta) dF_R(\theta) + \\ P_i \int\int_{\overline{\mathcal{A}}} B(\theta_L, \theta_R) dF_L(\theta) dF_R(\theta),\ i &= M, Q \end{aligned}$$

$$\begin{aligned} & P_i \int\int_{\overline{\mathcal{A}}} B(\theta_L, \theta_R) dF_L(\theta) dF_R(\theta) \\ = \ & \frac{1}{2} P_i \int\int_{\overline{\mathcal{A}}} |\theta_L - \theta_R| dF_L(\theta) dF_R(\theta) \geq 0 \\ \Rightarrow \ & V^* - \widetilde{V} \geq 0. \blacksquare \end{aligned}$$

Political information can be demanded for a number of different purposes. Instrumental voting is just one possibility. A sense of civic duty might also play a role as this can be seen as part of being a "good citizen". Political information can also be demanded to understand or forecast public policies and this in turn can be useful for better private decision-making[1]. Finally, information can be enjoyed as a consumption good and therefore be directly included in the utility function. We represent all this "exogenous" utility from information with I and say that total utility from information is

$$W = I + \beta \Delta v \tag{1}$$

Here β is a parameter we introduce for our convenience. If instrumental voting theory is valid then we should simply have $\beta = 1$. Otherwise we should expect $\beta = 0$, i.e. no demand for political information arising from voting decision-making. In general, thanks to the following result, we will be able to test $\beta > 0$ versus an alternative of $\beta = 0$.

Proposition 1 If $\beta > 0$ then W is higher in constituency M.

Proof. Straightforward from the (1), as $\Delta v = \Delta v(P)$ with $\frac{\partial \Delta v(P)}{\partial P} > 0$ and P is inversely related to expected margins of victory.

[1] See Larcinese (1999)

2.2 Information supply and mass media

We consider two newspapers A and B. They supply political news about both constituencies. We assume they have a fixed space $\overline{s}$ to devote to these news and indicate with $s_M^A \in [0, \overline{s}\ ]$ the space devoted by newspaper A to news about M; analogously we can define s_Q^A, s_M^B, s_Q^B. Each citizen buy one newspaper. The probability of a citizen buying newspaper j to get informed about constituency i is $q(s_i^j)$, with $q^{'} \geq 0$ and $q^{''} \leq 0$. We will assume each citizen will only care about her own constituency, thus simply ignoring news about the other constituency.

We then have $s_M^j + s_Q^j = \overline{s}$ $(j = A, B)$ and define a newspaper news profile as $\left\{ s_M^j, s_Q^j \right\}$. A citizen living in constituency M gets from newspaper A a utility from news equal to

$$w(s_M^A) = q(s_M^A) W_M.$$

Newspapers also report about other things apart from politics. Culture, sport, and other events are also covered as well as enjoyed by readers. Each paper has its own mix over these different forms of entertainment and also its own way of dealing with them. Also, the way politics in itself can be reported is not unique. The depth and the focus of news, as well as possible partizanship, all matter for the reader. We will therefore indicate the expected utility from newspaper A to citizen k in constituency M with $w(s_M^A) + a_k$, where a_k (b_k) is a fixed characteristic of newspaper A (B) that makes it different from B (A). Analogously for the other newspaper and constituency. We are then assuming that editorial choices, entertainment content, partizanship etc. are fixed characteristics of each newspaper: this is not an unrealistic assumption in the short run and certainly within the space of an electoral campaign.

Then we say that citizen k in constituency M buys newspaper A if

$$w(s_M^A) + a_k \geq w(s_M^B) + b_k \tag{2}$$

and buys newspaper B otherwise. Let's indicate with Δw_M the difference $w(s_M^A) - w(s_M^B)$ and with η_k the difference $a_k - b_k$.

Newspapers are uncertain about individual preferences, in particular preferences about the entertainment component. We assume η_k is distributed according to a distribution function G_i $(i = M, Q)$, which is common knowledge. The corresponding density function is g_i. Thus, the probability that citizen k in constituency M buys newspaper A is given by $Pr[\eta_k \leq \Delta w_M] = G_i(\Delta w_M)$.

Assumption 1 $\frac{|g_i^{'}(\Delta w_i)|}{g_i(\Delta w_i)} \leq \frac{|q_i^{''}(s_i^j)|}{W_i(q_i^{'}(s_i^j))^2}$, $i = M, Q$; $j = A, B$

Newspapers maximize expected profits. Each reader provides the newspaper with a revenue R which is the sum of the price directly paid by readers to buy the paper and the amount paid by advertisers per reader. Therefore total profits in the industry are given by $\pi = nR - 2C$, where n is the total number of citizens in the polity and C the fixed cost to produce each newspaper. For the moment, we only consider fixed costs and assume marginal costs are zero. Of course there are marginal costs represented by the printing and delivery costs but the "cost of the first copy" is normally by far the biggest cost.

Since we are interested in the share of the market newspapers have in each costituency we will rewrite the expected profit equation for newspaper A as

$$E(\pi^A) = R[E(n_M^A) + E(n_Q^A)] - C$$

where n_i^j is the number of readers newspaper j has in constituency i. For newspaper B we have $\pi^B = \pi - \pi^A$. Since costs are sunk, newspapers are only interested in maximizing revenue: in our model this simply implies that newspapers will maximize the expected number of readers ρ. Indicating with n_i the total number of voters in constituency i, we will have

$$E(\rho^A) = n_M G_M(\Delta w_M) + n_Q G_Q(\Delta w_Q)$$

A strategy for newspaper j is given by $s^j = [s_M^j, s_Q^j]$. We will indicate the set of feasible strategies for newspaper j with $S_j = \{s_M^j, s_Q^j | \ s_M^j + s_Q^j = \overline{s}\}$.

This is zero-sum game. Therefore a Nash equilibrium of the maximizing readership game is given by a feasible strategy profile $\{s_A^*, s_B^*\}$ s.t.

$$E(\rho^A | s_A^*, s_B) \leq E(\rho^A | s_A^*, s_B^*) \leq E(\rho^A | s_A, s_B^*)$$

Proposition 2 *Suppose Assumption 1 is satisfied, $\beta > 0$ and $n_M = n_Q$. Then an equilibrium strategy profile $\{s_A^*, s_B^*\}$ must satisfy $s_M^A = s_M^B > s_Q^A = s_Q^B$.*

Proof. The best reply function for newspaper j is defined implicitly by the first order conditions

$$\begin{aligned} n_M g_M(\Delta w_M) W_M q^{'}(s_M^j) &= 0 \\ n_Q g_Q(\Delta w_Q) W_Q q^{'}(s_Q^j) &= 0, \ j = A, B \end{aligned}$$

that imply

$$\begin{aligned} q^{'}(s_M^A) &= q^{'}(s_M^B) \\ q^{'}(s_Q^A) &= q^{'}(s_Q^B) \end{aligned}$$

and therefore

$$\begin{aligned} s_M^A &= s_M^B \\ s_Q^A &= s_Q^B \end{aligned}$$

Now remember that

$$\begin{aligned} W_M &= I + \beta \Delta v(P_M) \\ W_Q &= I + \beta \Delta v(P_Q) \\ \frac{\partial \Delta v(P_i)}{\partial P_i} &> 0, \ i = M, Q. \end{aligned}$$

Being $W_M > W_Q$ from the first order conditions we get that $s^j_M > s^j_Q$, $j = A, B$.

Assumption 1 guarantees that the second order conditions are satisfied (see Lindbeck-Weibull, 1987).■

So far we only focussed on the implications of marginality on information demand and supply. There are a number of other factors that can have an influence on information demand and supply and therefore should be used as control variables when trying to assess the effects of election closeness. On the media revenue side it is quite realistic to assume that not everyone has the same value to advertisers and that newspapers are capable of discriminating among different readers. The extent of this discrimination depends on the knowledge newpapers and advertisers have of market conditions and people's characteristics. Thus, we should expect this type of discrimination to become more and more relevant as new technologies improve the amount and quality of information on customers. Stromberg (2001a) relates the value to advertisers to an interest in specific aspects of public policy: for example, being interested in a particular public service rather than others reveals something about people's income, and readers' income is important for advertisers. At the same time in practice we do not observe any price discrimination across different readers. This means that discrimination will mainly occur through information supply.

Another consideration concerns costs. So far we assumed that the marginal cost of producing and delivering papers was zero. Although, as we said, marginal costs have only a minor part in the production of newspapers, for our purposes delivery costs could be important. We are considering possible spatial discrimination by newspaper and in this sense delivery costs could be quite different. In particular, in areas which are densily populated marginal delivery costs are probably negligible while they could be sizeable if our newspapers wanted to reach readers in remote parts of the country.

By modifying our assumptions and introducing differentiated constituencies we will therefore obtain a rationale for control variables that will make our test more reliable. At the same time in this way we will also be able to implement a direct test of some of the main Stromberg's results.

Heterogeneity here enters at the constituency level. In other terms newspapers are not able to discriminate readers according to any other individual characteristics apart from the constituency they come from, and we now assume constituencies are statistically different. This is actually the strategy that will be used to implement the empirical analsysis .

Assumption 2 $R_M \neq R_Q$.

Advertisers will induce from the constituency a number of other characteristics of interests and therefore will be willing to pay differently for marginal readers. Also the cost function is now different.

Assumption 3 *The newspaper cost function is* $TC = C + n_M G_M(\Delta w_M)c_M + n_Q G_Q(\Delta w_Q)c_Q$, *where* c_M *and* c_Q *are the cost of the marginal reader.*

For empirical purposes we will mainly identify c_M and c_Q with delivery costs. Now we can define the net marginal revenue per-reader as

$$r_i = R_i - c_i,\ i = M, Q$$

The profit equation for newspaper i can be re-written as

$$E(\pi^j) = r_M E(n_M^A) + r_Q E(n_Q^A) - C,\ j = A, B$$

To ensure that every citizen buys one newspaper and newspapers have an interest in reaching all citizens we assume the following:

Assumption 4 $r_i > 0\ \forall i$.

Now the problem will not simply be to maximize expected readership, as each reader must be weighted by her "net value". The next proposition provides the Nash equilibrium condition for this game.

Proposition 3 Suppose Assumptions 1-4 are satisfied and $\beta > 0$. Then an equilibrium strategy profile $\{s_A^*, s_B^*\}$ must satisfy $s_M^A = s_M^B, s_Q^A = s_Q^B$ and $\frac{q^{'}(s_M^*)}{q^{'}(s_Q^*)} = \frac{n_Q r_Q W_Q}{n_M r_M W_M}$.

Proof. The profit equation for newspaper j can be expressed as

$$E(\pi^i) = r_M n_M G_M(\Delta w_M) + r_Q n_Q G_Q(\Delta w_Q) - C,\ j = A, B.$$

The result follows from the first order conditions

$$\begin{aligned} r_M n_M g_M(\Delta w_M) W_M q^{'}(s_M^j) &= 0 \\ r_Q n_Q g_Q(\Delta w_Q) W_Q q^{'}(s_Q^j) &= 0,\ j = A, B. \end{aligned}$$

Assumption 1 guarantees that the second order conditions are satisfied (see Lindbeck-Weibull, 1987).■

Proposition 4 tells us that now newspapers can discriminate across constituencies also on the basis of further information they may have. Other things equal, information supply will be higher in the constituency with larger $r_i = R_i(Y, A, E...) - c_i$. On the revenue side we can relate the readers' value for advertisers to factors as income, age, education etc. The net value of readers for newspapers will then consider their location and be lower where delivery costs are higher; we will use population density to capture this element.

Finally, also the total size of constituencies n (in terms of absolute population, or absolute electorate) should play a role. However, as we noticed at the start of this section, we can have both a "groups size effect" (like in Stromberg) as well as a "collective action effect" and we will approach empirical investigation with no prior about the sign of this variable.

We can therefore summarize our findings in the following testable result:

Theoretical Results *Other things equal, information supply is higher in constituencies with a closer electoral race, more densely populated, and where citizens are on average more valuable to advertisers. The effect of the size of electorate is uncertain.*

3 Evidence

We will proceed now to verify the compatibility of our theoretical results with data. Empirical investigation will concern the 1997 general election in the United Kingdom. For the purpose of this analysis we will use data on England, Scotland and Wales. The political situation in Northern Ireland is substantially different from the rest of the country as the main divide is between Catholic and Protestants rather than on the traditional left-right dimension.

In the U.K. members of parliament (MPs) are elected one in each constituency in a first past the post system. The election is won by the part which obtains the larger number of MPs as support of the parliament is necessary to become prime minister. Party leaders are intended to be candidate prime minister, but they still need to win in their own costituency to become MP. There are two major parties, Conservative and Labour, although other parties manage to win in some constituencies. In particular the Liberal-democratic party is well established nationally as a "third party". In the 1997 the Labour party obtained a neat victory after 17 year of Conservative ruling.

3.1 The Data

We will use three main surces of data. The unit of observation is the electoral constituency. There were 641 such constituencies in Britain (England, Scotland and Wales) in 1997.

First of all we need data about information supply by newspapers. For this purpose we will use a a major national newspaper, "The Guardian". We will define information supply for each costituency as the number of articles that mention such constituency or one of its candidates during the last 30 days of the electoral campaign. This variable is indicated as N, for news.

We will then use information about electoral results[2]. In particular, we will use this information to measure the marginality of a constituency. A first possibility is to focus on percentage differences and therefore use the following formula:

$$\frac{W - R}{W + R} \tag{3}$$

where W is the percentage of votes for the winning candidate and R the percentage for the runner up[3]. The larger such indicator the lower the degree of marginality of the constituency. However, to capture the idea of marginality as the probability of casting a decisive vote, the absolute difference in votes between candidates might be a more appropriate indicator. We will consider both possibilities.

One problem with such indicators is that they measure election closeness ex post. This is not necessarily a good measure of expected marginality, altough a

[2]Quote source....

[3]This measure has been suggested by...

rational expectations assumption would work in our favour. One possibility to overcome this problem would be to use past election results. This is also quite problematic in our case. The main obstacle is that in between 1992 (year of the previous general election) and 1997 most constituency borders were changed, making therefore hard any comparison, even where the denominations were left unchanged. Moreover, in 1997 there were expectations of a large swing from the ruling party (Conservatives) to opposition parties (mainly the Labour): thus, previous election closeness could not represent a good measure of expected election closeness as this last would crucially depend on who held the constituency.

For these reasons, to capture expected closeness we will use the BBC "100 contested costituencies". Those are constituencies on which the BBC decided to focus its attention on the night of the results: they were the expected closest 100 Conservative-held constituencies (and therefore the decisive ones). This should capture information from polls and the general feeling about election closeness of BBC journalists.

We will also use information about the total number of electors in each constituency and the turnout percentage. With the first variable we try to test if the "group size effect" can actually prevail on the "collective action effect". The percentage of turnout indicates the extent of political participation (at least in the form of voting) and therefore can be broadly intended as a measure of interest in politics by the citizens of a given area.

Information about other possibly relevant characteristics of the constituency will be derived from the 1991 Census. These data are considered at the level of districts, local administration entities with no direct link with electoral constituencies. Some constituencies are contained within the borders of a single district and these posed no problems. Some others span over parts of different districts and in such cases data referred to districts have been weighted in order to get approximated constituency data.

First of all we include the population density, one of the key variables in our theoretical analysis, entering into the newspapers' cost function. Then we consider variables that can possibly give a representation of the social and economic conditions of the districts. Information on income is not avilable but several proxies have been used, like the unemployment rate and the percentage of citizenship with high qualifications (degree and higher). Age can have an influence on propensity to consume and consumption patterns (thus affecting how valuable a reader is to advertisers) and therefore has been included. Also, the percentage of inactive population (mainly retired, but also including students and permanently sick) is used. There are reasons (as well as anecdotical evidence[4]) to think that inactive population, in particular old or sick individuals, should be less valuable to advertisers, in particular as they tend to consume less than average.

One possible concern might derive from the fact that The Guardian, like most national newapapers in the U.K., is London based. This could bias the news in favour of London constituencies both because of a lower cost of news

[4]See Stromberg (2001)

collection and, more generally, because of a larger sensitivity to a nearer environment. This could be particularly relevant for our results about population density, being this variable generally higher in London than elsewhere. To overcome this potential problem we include a Greater London control dummy, equal to 1 for the Greater London constituencies.

Finally we also include a "big-shot" control. As some candidates have naturally a prominent position and bigger visibility during the electoral campaign, it seems necessary to be able to single out this effect from what we want to test. Therefore we introduce a dummy variable equal to 1 for constituencies where "big-shot" are candidates. By big-shot we intend all the candidates who have been ministers in any past government plus the members of the current "shadow-cabinet"[5].

Summary statistics are reported in table 1.

3.2 Empirical Specification

Using the dataset described above we intend to test the theoretical results reported at the end of section 2. Preliminary data analysis seems to suggest that a very limited number of constituencies get a disproportionate attention from media (see Tab. 2). For example almost 90% of constituencies have $news \leq 5$ while only 3 constituencies have $news > 100$. This suggests that the relationship we want to estimate could be highly non-linear. A linear regression would indeed deliver quite poor results. We will instead present estimates for the following equation:

$$\ln(N_i) = \alpha + \beta D_i + \gamma^{'} POP + \delta^{'} Z_i + u_i,\ i = 1, ..., N \tag{4}$$

where:

$N = news$, D is a measure of the distance between candidates (winner and runner up) or expected closeness (BBC100), POP is a two-dimensional vector of population density and size of the electorate and Z represents a set of control variables from the 1991 Census, plus the "big-shot" dummy. u_i as usual represents independent disturbance terms that have means of zero and are uncorrelated with the exogenous variables of the model. Estimation will be by OLS and will take into account potential heteroskedasticity. Specification tests will also be conducted, in particular to assess the log-linear functional form.

Almost all the parameters have an expected sign in terms of our model. However, our main parameter of interest is β. In general, we want to assess if β is significantly different from zero. As discussed previously, we will consider several possible measures of constituencies marginality.

The other variables serves as controls with respect to this aim but at the same time the γs and the δs are of interest for their own sake as we can use

[5] A "shadow-cabinet" is put in place by the opposition party and is composed by those who, in case of victory, most likely will become ministers. A shadow foreign secretary, for example, follows very closely the government foreign policy and is supposed to be able to control and propose alternatives. This makes shadow-cabinet members quite popular on the media.

their estimation to assess the validity of our model of information demand and supply.

3.3 Results

OLS estimates of equation 4 are reported in table 3. We start by considering the role of election closeness. In column 1 and 2 we use ex post indicators of election closeness (based on percentage distance in column 1 and absolute distance in column 2). In both cases ex post distance has the expected sign and is signicant at 5% level. When we use expected closeness (as captured by the dummy BBC100 in column 3) the significance level increases substantially, reaching a nil p-value for the hypothesis of β being zero. As discussed previously, the 1997 general election witnessed a large generalized shift of votes away from the Conservative party. This was to some extent expected and therefore the most interesting contituencies were especially the previously Conservative-held ones. In a sense, the final outcome was mainly decided in such constituencies and this should have increased the demand for information[6]. Thus, it is not surprising that when we focus on contested conservative constituencies results get sharper Actually, some constituencies may have been ex post very close just because the swing of votes has probably been larger than expected, making the labour candidates winning (marginally) also in constituencies that never were marginals or Labour-held before. Therefore, ex post marginality is clearly an imperfect measure of expected salience.

In column 4 we also get rid of another ex post indicator, turnout, but results remain substantially unchanged. Thus, from this analysis we can safely conclude that expected marginality is important for information acquisition and supply and therefore for instrumental voting.

Other variables also show a high compatibility of our model with facts. Population density has the expected sign and is always significant at the 5% level, except in column 3, where it is significant at the 7.5% level.

The sign of other control variables also show good support for some of the Stromberg-type conclusions. In particular, and differently from Stromberg, we saw that the effect of group's magnitude is not so uncontroversial. However, empirical evidence seems to suggest that the effect of the group size should overcome the potential collective action problem that size ingenerates> In particular, in column 2 we use the absolute distance between candidates and therefore we isolate the potential "group size effect". However, we do not notice particularly relevant differences between this and the other cases. Although the magnitude and significance of the electorate size are larger in column 2, these remain always positive and comfortably significant at the 5% or 1% level.

The same is not true for Turnout, that also should serve as a signal to newspapers about the degree of attention to political matters. The sign is

[6]In terms of our model, in general elections citizens care mainly about final policies: thus, marginality in one constituency is more relevant when it matters for the whole outcome of the election.

always negative, although we should better say that such parameter is never significant.

Good support also comes from the unemployment rate, that we use as a proxi for the level of well-being in a given constituency. On the other hand there is little evidence in favour of the relevance of other factors that the literature seems to have identified as determining information supply. Anecdotical evidence is reported of television programmes that have been suspended because watched mainly by the elderly, who where judged not valuable by advertisers. However, we find that constituencies with larger inactive population (mainly represented by retired) receive more attention from newspapers. It should be remembered that inactive people might have more time to devote to information gathering and when we come to election times retired people might also have all the incentives to put a disproportionate attention to political platforms.

Finally, average age and the percentage of people with high degrees do not seem to have significant effects, while there is evidence of some "Greater London effect", however not particularly large or significant. A pure control variable is big-shot. Both the magnitude and the significance of big-shot are very relevant but this does not come as a surprise nor is the consequence of any theoretical advance made in this paper.

4 Conclusion

The purpose of this paper is twofold. On one side it is to study the impact of election closeness on political information demand and supply. The relationship between closeness on turnout has been widely investigated in the literature, being regarded as an indicator of instrumental behaviour by voters. We focus on information acquisition as another consequence of instrumental voting and show that it should be higher when elections are expected to be closer. Profit maximizing mass media should therefore discriminate between different constituencies. We do not observe newspapers price discrimination in reality. However we show, both theoretically and empirically, that the media have a different way to discriminate, namely targeting their attention (in terms of reported news) to marginal constituencies.

The other objective of the paper was to test newspaper behavior during electoral campaigns in a direct way. The first such test was cleraly related to the role of election closeness. However, recent theoretical literature, and in particular Stromberg (2001a), has pointed out that the media should tend to target customers who are more valuable to advertisers, i.e. welthier, better educated, younger. By using the electoral constituencies in the 1997 British election as units of obervation, we can test these conclusions. Also, by modeling the newspapers' cost function and taking into account delivery costs, we derive that information supply should be higher in more densely populated areas. This too finds confirmation in our empirical investigation. We also provide empirical support for the idea that larger groups should receive more attention from the media, although we have shown that this conclusion does not necessarily follow from the theory.

Thus, this analysis seems to suggest a comfortable compatibility between instrumental voting behaviour and observed facts for what concerns information demand and supply. Also, if the media introduce a bias in the way people are informed about politics and if this bias is in turn exploited by politicians then we can speak of a "media-driven-bias" in public policy-making. At the same time there is an "attention-bias" that should substantially drive politicians to target marginal constituencies.

This analysis is by no means conclusive and there are several margins for improvements and questions that further research should try to address.

On the theoretical side, the model of media competition is still quite simple. New insights could come from explicitly considering the advertising market and the possibility for newspapers to select the combination of political information, advertsing and other news in the paper. Also, considering the possibility of entry and, more in general, different industry structures, could deliver interesting results as well as normative implications for regulating the media market. An analysis of this type with respect to politicians' accountability can be found in Besley-Prat (2001). Further research could help us understanding redistributive implications.

On the empirical side improvements are also in order. Furher data collection, about different papers, different elections and, finally, different countries, could help us to understand what is the robustness and the generality of our results and may be to isolate the relevant institutional characteristics that induce different behaviours. For what concerns our estimates, the relationship between news and closeness (as well as news and other variables) is clearly non-linear. We have chosen a log-linear specification and shown that it fits our data quite well. It would be useful however to resort to non-parametric estimation to get a better understanding of newspapers behaviour and motivations.

5 Appendix

5.1 Description of variables

- $N = news$. It is the number of articles appeared on the newspaper "The Guardian" during the last 30 days before the poll date and containing either a reference to the electoral contituency or the name of one of its candidates.

- $Dist\_P$ = percentage distance between the winning candidate and the runner-up in the 1997 election, given by the formula (W-R)/(W+R), where W=percentage of votes for the winning candidate, R=percentage of voter for the runner up.

- $Dist\_N$ = absolute distance between the winning candidate and the runner-up in the 1997 election, given by the formula (W-R)/(W+R), where

W=percentage of votes for the winning candidate, R=percentage of voter for the runner up.

- $BBC100$ = dummy variable equal to 1 if the constituency has been included by the BBC among the "100 contested constituencies".
- $Density$ = population density expressed as the number of residents per square kilometer multiplied by 1000.
- $Electorate/1000$ = total electorate in the constituency divided by 1000.
- BS = dummy variable equal to 1 if one of the candidates in the constituency has been classified as a "big-shot". This means when one of the candidates is a current or former minister or a current member of the "shadow cabinet".
- $Unemployment\%$ = unemployment rate, expressed as total unemployed over active population, multiplied by 100.
- $Inactive\%$= percentage of inactive population. This means the total of retired, students, permanently sick and other inactive over total residents multiplied by 100.
- Age = average age;
- $HIGHD$ = percentage of residents with high qualitifications, defined as the number of residents with degree or higher title over the total residents, multiplied by 100;
- $GLondon$ = dummy variable equal to 1 for the greater London constituencies.

References

[1] Besley, T. - Burgess, R. (2001): The political economy of government responsiveness: theory and evidence from India, mimeo, LSE.

[2] Besley, T. - Prat, A. (2001): Handcuffs for the grabbing hand? Media capture and government accountability, mimeo, LSE

[3] Blumler, J. - McQuail, D. (1968): Television in Politics: Its Uses and Influences, London, Faber.

[4] Ferejohn, J. - Kuklinski, J. (eds.) (1990): Information and Democratic Processes, Urbana & Chicago, University of Illinois Press.

[5] Grossman, G. - Helpman, E. (1996): Electoral Competition and Special Interest Politics, The Review of Economic Studies, 63, 265-286.

[6] Larcinese, V. (1999): Rational Ignorance and the Public Choice of Redistribution, mimeo, London School of Economics.

[7] Larcinese, V. (2000): Information Acquisition, Ideology and Turnout: Theory and Evidence from Britain, mimeo, London School of Economics.

[8] Lau, R. - Sears, D. (eds.) (1986): Political Cognition, Hillsdale, Erlbaum.

[9] Lindbeck, A. - Weibull, J.W. (1987): Balanced-budget redistribution as the outcome of political competition, Public Choice, 52, 273-97.

[10] Lohmann, S. (1998): An Information rationale for the Power of Special Interests, Americal Political Science Review, 92-4, 809-27.

[11] Matsusaka, J.G. (1995): Explaining Voter Turnout Patterns: an Information Theory, Public Choice, 84, 91-117.

[12] Noam, E.M (1987): A Public and Private Choice Model of Broadcasting, Public Choice, 55, 163-187.

[13] Spence,M.- Owen, B. (1977): Television programming, monopolistic competition and welfare, Quarterly Journal of Economics, 91, 103-26.

[14] Stromberg, D. (2001a): Mass Media Competition, Political Competition and Public Policy, mimeo, Stockholm University.

[15] Stromberg, D. (2001b): Radio's impact on New Deal Spending, mimeo, Princeton University.

[16] Zaller, J. (1991): The Nature and Origins of Mass Opinion, New York, Cambridge University Press.

Table 1: Summary statistics

	Obs	Mean	Std. Dev.	Min	Max
News	641	3.8658	18.4735	0 (app. to 0.01)	388
Log(News)	641	-1.3847	2.8319	-4.6052	5.961
Dist_P	641	0.2972	0.2006	0.0012	0.8219
Dist_N	641	10859	6906	53	33759
BBC100	641	0.156	0.3631	0	1
Density	641	1.6524	1.9092	0.0088	11.6298
Electorate/1000	641	66.5437	8.0574	22.983	101.68
Turnout%	641	71.3165	5.6359	51.4	82.2
BS	641	0.078	0.2684	0	1
Age	641	37.8866	1.881	32.8793	46.533
Inactive%	641	51.32	3.2413	38.858	61.0908
Unemployment%	641	9.45	3.8	2.868	22.4896
HighD%	641	7.0963	3.743	1.4891	25.084
GLondon	641	0.1154	0.3198	0	1

Table 2: The variable "News"

News	Frequency	Percent	Cumulate
0	265	41.34	41,34
1	165	25.74	67.08
2	74	11.54	78.63
3	28	4.37	83
4	18	2.81	85.80
5	24	3.74	89.55
6-10	25	3.9	93.45
11-20	19	2.96	96.41
21-30	10	1.56	97.97
31-40	3	047	98.44
41-50	3	047	98.91
51-100	4	0.62	99.53
>100	3	0.47	100

Table 3: Regression results (OLS)

Dependent Variable = Log(News)

	1	2	3	4
Dist_P	-1.4808 (-1.960)			
Dist_Nx1000		-0.0366 (-1.924)		
BBC100			1.1889 (4.306)	1.0415 (3.84)
Density	0.2448 (2.076)	0.2521 (2.153)	0.2081 (1.783)	0.2535 (2.227)
Electorate/1000	0.0321 (2.311)	0.3891 (2.884)	0.0334 (2.463)	0.336 (2.46)
Turnout	-0.04 (-1.406)	-0.0305 (-1.138)	-0.0472 (-1.801)	
BS	3.4017 (8.815)	3.3891 (8.726)	3.4925 (8.955)	3.4918 (8.899)
Age	-0.1074 (-1.268)	-0.1071 (-1.264)	-0.0701 (-0.83)	-0.07 (-0.836)
Inactive	0.1449 (2.022)	0.1423 (1.989)	0.1257 (1.753)	0.1209 (1.704)
Unemployment	-0.1789 (-2.544)	-0.179 (-2.555)	-0.1985 (-2.847)	-0.1625 (-2.414)
HighD	0.0327 (0.857)	0.0284 (0.742)	0.039 (1.036)	0.0322 (0.862)
GLondon	0.7046 (1.589)	0.7094 (1.605)	0.8574 (1.971)	0.8016 (1.855)
Constant	-2.8828 (-0.764)	-3.919 (-1.082)	-3.3395 (-0.948)	-6.8138 (-2.368)
Obs	641	641	641	641
R-squared	0.1662	0.166	0.1818	0.1779

Note: robust standard errors. T-statistics in parenthesis