



ZeMKI Working Papers

ZeMKI Working Paper | No. 51

ISSN 2510-9855

Shichao Zhang, Shinae Lee, Chankyung Pak

Uncertain Election Polls, Uncertain News Coverage: Two-Sided Heterogeneity in News Outlet-Pollster Relations



Universität
Bremen



ZeMKI Zentrum für
Medien-, Kommunikations- und
Informationsforschung

Universität Bremen | University of Bremen
ZeMKI, Zentrum für Medien-, Kommunikations- und Informationsforschung
Linzer Str. 4, 28359 Bremen, Germany, E-mail: zemki@uni-bremen.de
www.zemki.uni-bremen.de | www.zemki.uni-bremen.de/en/

Shichao Zhang (zhshch@knu.ac.kr)

Shichao Zhang is a doctoral student at Kyungpook National University in Daegu, South Korea. Her research interests include science communication, public opinion, and cross-cultural communication. She received her M.A. in Media and Communication from the same university in 2023. She has conducted research on polling coverage, AI and media perception, digital media literacy. Her dissertation focuses on science popularization on digital media, trying to examine audience responses to online science videos under multimodality and platform-mediated environments. Methodologically, she is looking forward to using computational approaches to investigate transformations in the digital public sphere.

Shinae Lee (leesa1257@naver.com)

Shinae Lee is a doctoral student at Kyungpook National University in Daegu, South Korea. Her research looks at how media technologies, cultural expectations, and institutional structures interact to influence patterns of communication and the exercise of power in society. In particular, she is interested in how media reconstructs social reality, with a particular focus on the marginalized groups that emerge as a result—and the forms of solidarity they build in response. Also she has contributed to multiple grant-funded research projects in the fields of digital journalism, data analysis, and diaspora studies, working with institutions such as the National Research Foundation of Korea, the Korea Press Foundation, and the News Communication Promotion Committee.

Chankyung Pak (chankyungpak@knu.ac.kr)

Chankyung Pak (Ph.D. Michigan State University) is an associate professor of the Department of Media and Communication at Kyungpook National University. His research interests lie in computational social sciences as applied to the study of transforming digital media ecosystems and their impact on the formation of public discourse. He also investigates how generative AI (re)creates disjunctures between technology ethics and legal models. In 2022, he was a ZeMKI Visiting Research Fellow.

Working Paper No. 51, October 2025

Published by ZeMKI, Centre for Media, Communication and Information Research, Linzer Str. 4, 28359 Bremen, Germany. The ZeMKI is a Central Research Unit of the University of Bremen.

Copyright in editorial matters, University of Bremen © 2025

ISSN: 2510-9855

Copyright, Electronic Working Paper (EWP) 51 - Uncertain Election Polls, Uncertain News Coverage: Two-Sided Heterogeneity in News Outlet-Pollster Relations, Shichao Zhang, Shinae Lee, Chankyung Pak, 2025

The authors have asserted his moral rights.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form of binding or cover other than that in which it is published. In the interests of providing a free flow of debate, views expressed in this EWP are not necessarily those of the editors or the ZeMKI/University of Bremen.

Uncertain Election Polls, Uncertain News Coverage: Two-Sided Heterogeneity in News Outlet-Pollster Relations

Abstract

Election polls and their news coverage function as a form of chained gatekeeping, shaping voters' perceptions of fellow citizens' preferences. The recent proliferation of polling firms equipped with low-cost, novel methods has not only multiplied competing results—confusing the perception of public opinion—but also compelled news outlets to selectively rely on trusted pollsters. To examine how outlets make these organizational choices, this study conceptualizes poll coverage as an interorganizational relationship between outlets and polling firms. By evaluating poll diversity at the outlet level, we find that highly institutionalized outlets adhere to journalistic norms by diversifying their coverage, yet still exhibit a preference for legacy pollsters. Further, under conditions of dual heterogeneity, we show that emergent outlets tend to increase the visibility of less credible pollsters, thereby contributing to a more fragmented news ecosystem.

Keywords

election polls, election coverage, polling diversity, interorganization relationship, dual heterogeneity

1 Introduction

Despite the long-standing criticism of election polls for reducing the various meanings and roles of public opinion in democratic deliberation (Brodie, Parmelee, Brackett & Altman, 2001), they continue to play an indispensable role in shaping citizens' awareness and beliefs about how their fellow citizens view public matters, and consequently how they themselves decide (Meffert, Huber, Gschwend & Pappi, 2011; Irwin, 2002). This signaling function has become even more important as the growing complexity of modern society makes genuine deliberation increasingly difficult (Lepore, 2015). Yet, the perception of public opinion depends on an additional layer of gatekeeping through news coverage, as most citizens learn about poll results via news stories (Beckers, 2021). Accordingly, the dual sources of potential bias – pollsters' measurement errors and news outlets' selective reporting of polls – have attracted scholarly attention to the processes by which public opinion is translated into collective decision-making (Hillygus, 2011).

In recent decades, the gatekeeping functions of both journalism and polling have grown markedly more complex, largely driven by technological innovation. The proliferation of low-cost platforms for personalized information dissemination, such as social media and web portals, has generated a multitude of news outlets – or pseudo-outlets – many of which do not adhere to conventional journalistic norms (Singer, 2005). The new media ecosystem, populated by emergent online actors, offers greater diversity of perspectives on the one hand, but is often blamed for misleading fragmented audiences with alternative facts and extreme claims on the other (Guo & Vargo, 2020; Miró-Llinares & Aguerri, 2023). A parallel development has taken place among pollsters, who increasingly exploit novel technological opportunities, such as mobile and web-based non-probability sampling, along with complex statistical weighting and modeling to correct for selection bias, rather than relying on traditional, costly face-to-face probability sampling (Prosser & Mellon, 2018). Some of these newcomers gained prominence by producing strikingly accurate forecasts during politically turbulent times, when many traditional polls failed. Yet

many later suffered embarrassing missteps in subsequent elections, further contributing to an information environment marked by competing and conflicting claims (Graham, Hillygus, & Trexler, 2024).

Despite this dual heterogeneity in the interconnected two-step information gatekeeping processes, the literature has largely treated this issue as a unilateral information-selection problem: as if a series of numbers simply becomes available during an election cycle, and news outlets decide whether or not to report them (Beckers, 2021). Although this selection is understood to be influenced by outlets' organizational traits — such as ideology, target range (national, local, etc.), and medium (Benson, 2004; Mathisen, 2023; Searles, Smith & Sui, 2018) — the framework has generally remained limited to choosing which available information to report. Within this perspective, selecting polls to cover is not fundamentally different from the broader process of gatekeeping available information.

However, when a growing number of pollsters produce divergent results within the same time window using different methods — placing both pollsters and outlets along a heterogeneous continuum from emergent to institutionalized — the selection problem itself is likely to become institutionalized. In other words, outlets increasingly choose among pollsters rather than among their products (Cho, Ryoo & Kim, 2017). This tendency may arise because outlets have limited time, resources, and expertise to evaluate numerous poll results; building inter-organizational relationships is thus more cost-efficient than making case-by-case decisions. In this context, outlets with substantial resources are more likely to engage in individualized decision-making, which underscores the relevance of examining how the degree of *choosing pollsters rather than polls* varies by outlet type.

Moreover, lasting relationships between news outlets and polling firms, potentially emerging from this dual heterogeneity, could foster a segregated news ecosystem composed of relatively disconnected clusters of outlets and pollsters. As organizational ecology scholars note, organizations often rely on homophilistic relationships driven by shared goals, concerns, and resources (O'Connor & Shumate, 2020; Wang, 2023). While such symbiotic ties can help market entrants gain visibility and legitimacy through alliances with peers (Rao, 2017), they may also “insulate an organization within its own population, or within relationships with other similar populations” (Weber, 2012). Segregated news ecosystems, evident in the links across legacy media, social media, and online communities (González-Bailón et al., 2023), can thus be interpreted as the product of these homophilistic connections in the news domain.

In this study, we approach the poll-selection problem through the lens of interorganizational linking. Specifically, we analyze the coverage of polls by online news outlets during the 2025 presidential election in South Korea. We model election news as a bipartite interorganizational network between 67 news outlets and 24 polling firms. Our findings show that legacy media with greater organizational resources tend to diversify their poll coverage while still prioritizing recognized legacy polling firms. In contrast, emergent outlets rely on a narrower set of pollsters. Moreover, although polls from emergent firms are significantly less covered than those from incumbents, they are nonetheless sustained by emergent outlets, suggesting homophilistic relationships between emergent outlets and emergent pollsters. By extending the analysis of poll coverage to these many-to-many relationships, we view poll reporting as an ecosystem, enabled by AI-based matching between news and poll data. We recommend this automated method as a way to generalize the monitoring of public opinion visibility, providing insights into the effects of linked mutual heterogeneity beyond the separate analysis of polls and news.

The rest of this paper is organized as follows. First, we develop our hypotheses on poll coverage by drawing on media sociology literature addressing information diversification

and the institutionalization of news organizations, as well as on interorganizational ecology research. Next, we describe our methods for automated data collection, AI-based news-poll matching, and statistical analysis to test the proposed hypotheses. We then present the main findings, followed by a discussion of their theoretical and practical implications.

2 Polling Source Diversification in Election Coverage

The transformation of the online user environment has reshaped the broader media landscape, giving rise to a proliferation of digital-native and specialized online news outlets (Anderson, 2013). This shift has created a more multidimensional news ecology, one that favors information resources that are accessible, verifiable, and compatible with accelerated production cycles (Petre, 2015; Tandoc, 2014). Amid intensified resource competition and pressures of production, news outlets tend to rely on sources that can provide timely and stable data in order to ensure both the immediacy and credibility of their reporting (Gans, 2004; Lewis, Williams & Franklin, 2008). Pollings inherently satisfy these conditions, offer quantifiable and timely indicators of public sentiment, carry the appearance of scientific objectivity, and are standardized for journalistic use (Searles, Ginn & Nickens, 2016). Thus, polls have become one of the most prominent informational sources in election coverage, particularly during election campaigns (Strömbäck & Kaid, 2009).

In the contemporary period, the polling industry has undergone a substantial expansion, reflected in an increase in the number of selectable products on the information-supply side. Evidence shows that in nearly half of the 157 countries, more than five different polling firms conduct pre-election polls, and in about twenty percent of cases, the number of such firms exceeds ten (WAPOR, 2023). Advances in methodology have further lowered the time and cost of conducting surveys, allowing election polls to be produced with increasing frequency and efficiency (Pew Research Center, 2023). In the United States, this expansion reached its peak in the month leading up to the 2024 presidential election, with more than one hundred organizations fielding nearly one thousand pre-election polls during that period (FiveThirtyEight, 2024). With the proliferation of polling firms and survey results contributing to an unprecedented information source of election coverage, it becomes necessary to critically address the increasing diversification in which polls are selected.

Although the polling industry has been steadily expanding (WAPOR, 2023), the growth of polling makes the selection of information sources increasingly diverse and complex for news outlets. Pollings have long been largely conceptualized as informational inputs into coverage, with research focusing primarily on what kind of polls are more likely to be selected (Patterson, 2005; Searles et al., 2016) and how poll results are presented in coverage (Strömbäck & Kaid, 2009). Limited scholarly attention has been given to the role of polling firms as organizational entities, particularly in their interactions with news outlets. Even though pollsters' distinctive resource capacities, funding and commissioning structures (Gabay, 2022; Tourangeau, Presser & Sun, 2014), methodological routines (Kennedy et al., 2018; Shirani-Mehr, Rothschild, Goel & Gelman, 2018), and professional norms (American Association for Public Opinion Research, 2021) constitute organizational-level elements influencing their function as information sources. Different polling firms exhibit significant variation in their organizational characteristics, giving rise to the so-called 'house effects,' which refer to systematic biases associated with particular firms (Druckman, Fein & Leeper, 2012; Jackman, 2005). This highlights the fundamentally heterogeneous nature of polling firms, significant systematic differences attributable to them, even under identical methodological conditions across repeated surveys over multiple years

(Schumann, Shamon & Hake 2020), constitutes a crucial dimension of the diversity of polling.

The selection and reporting of information from among numerous potential sources constitutes a routine and indispensable element of news production, carried out through a series of gatekeeping processes (O'Neill & Harcup, 2009; Shoemaker & Vos, 2009). Given the constraints of time and space, journalists must sift through vast streams of information and choose only a fraction for public dissemination. Source selection, however, is not a matter of randomness; rather, it reflects deeper issues of power, influence, and representation embedded in the journalistic process (Berkowitz, 2019). Research confirms that source selection continues to vary by outlets (Boehmer, Carpenter & Fico, 2019; Thomson et al., 2025; Mathisen, 2023). These differences are evident across national contexts, between for-profit and nonprofit outlets, and between local and non-local media. These patterns of source use are strongly shaped by organizational traits, including ownership structures, professional norms, and editorial routines (Berkowitz, 2019; Gans, 2004; Schudson, 1989).

Professional journalism standards have long emphasized the importance of source diversity as a safeguard for balanced and comprehensive coverage (Kurpius, 2002). Diversification of information sources also serves as a practical mechanism through which news outlets strengthen their credibility (Zerback & Schneiders, 2024). Many news outlets have adopted source diversification as part of their editorial practice to promote fairness in reporting (e.g., The Washington Post, 2025). The extent of institutionalization within outlets influences the degree to which professional norms are translated into newsroom routines. Highly institutionalized outlets typically demonstrate stronger professional cultures and more formalized editorial procedures, conditions that facilitate a broader and more systematic use of sources (Benson, 2004). Therefore, our first hypothesis is that *higher levels of institutionalization of news outlets are expected to diversify the sources used in polling coverage* (Hypothesis 1).

3 Relationship between News Outlets and Pollsters

Beyond the normative ideal of source diversification, this selection process is also constrained by the pre-established networks to which news outlets are connected with sources (Dimitrova & Strömbäck, 2009). In election coverage, the reporting of poll results often diverges from the aggregate outcomes of the polls, as not all results receive balanced attention. Some news firms rely disproportionately on a small number of highly institutionalized polling firms with established brand authority (Searles et al., 2016). Even as outlets demonstrate greater flexibility in agenda-setting or political storytelling, their sourcing of polls still reflects similar patterns of concentration, particularly favoring legacy or in-house polls (Searles, et al., 2018). Established firms often enjoy reputational authority and long-standing legitimacy, while emergent firms vary in quality and visibility, creating asymmetries in how their polls are received and utilized (Shirani-Mehr, et al., 2018; Toff, 2019). More credible sources tend to attract greater journalistic attention and achieve more prominent visibility in coverage.

Based on this reasoning, we argue that the recognition of polling firms serves as a key moderating variable in the relationship between the institutionalization of news outlets and the diversity of sources in poll reporting. Accordingly, our second hypothesis is that *the effect of institutionalization of news outlets on source diversity in poll coverage* (Hypothesis 1) *is moderated by the recognition of polling firms* (Hypothesis 2). Such that higher recognition of polling firms will attenuate the positive effect of institutionalization

on source diversity, as outlets gravitate toward legacy firms at the expense of diversification.

If Hypothesis 2 is taken, highly institutionalized news outlets tend to prioritize established and authoritative legacy pollsters, and emergent pollsters should have been marginalized over time. Yet such a perspective raises a paradox: if polls lacking authority are unlikely to be reported, why do these firms persist and continue to conduct surveys in the press-polling ecosystem? A similar question has been cast for emergent news outlets: why and how are alternative/hyperpartisan media able to survive and persist? While audience segmentation generates niches (Guess, Nyhan & Reifler, 2018), it is the dense, mutually linked structure (de León, Makhortykh & Adam, 2024; Zeng & Schäfer, 2021) that retains traffic inside the internal ecosystem and thereby underwrites continued survival and circulation. This observation indicates that beyond an organizational institutionalization perspective, it is essential to consider the homophilous ties of dual organizations, with particular attention to the relationally driven configurations that characterize the press-polling ecosystem.

We try to draw on the perspective of community ecology to describe the relations between news outlets and polling firms' population in the media ecosystem (Weber, 2012). Within traditional mainstream media, cooperation with legacy polling firms may develop symbiotic ties: polling firms obtain recurring exposure and institutional legitimacy through news coverage, while news outlets gain authoritative, data-driven content that enhances the perceived credibility of their reporting (Frankovic, 2005; Iyengar, Norpoth & Hahn, 2004). The notion of *assortativity* from network science also captures that empirical tendency of organizations with similar characteristics to connect preferentially with each other (Newman, 2003). Connections between emergent media and emergent pollsters, meanwhile, may constitute an alternative ecology, one that is sustained less by institutional authority than by the provision of more dramatic or unconventional results, which attract attention in the competitive media market. This leads to the following research question: *Do emergent outlets rely more heavily on emergent pollsters than legacy media do (RQ1)?*

4 Methods

Collecting News / Poll Data

The dataset used in this study consists of information about polling firms and news outlets, as well as the relationships between them derived from election reports. We first describe how the entity-level data was collected and then explain how relationship data are constructed by AI-assisted matching on election poll news.

Scraping Poll Data: To retrieve polls for the 21st Korean presidential election held on June 3rd, 2025, we scraped *Public Opinion M* website, co-developed by *Seoul National University* research team led by Jong Hee Park and *MBC*, one of Korea's public TV networks.¹ The website retrieves election poll data mandated to be registered with the National Election Survey Deliberation Commission (NESDC)² and visualizes both individual poll results and the estimation of latent trends in public opinion through statistical aggregation. To collect additional metadata, we also scraped the NESDC website itself. In total, we obtained 94 polls conducted by 24 firms between May 11th, 2025, when the two major parties (the

¹ <https://poll-mbc.co.kr/>

² <https://www.nesdc.go.kr/portal/main.do>

Democratic Party and the People Power Party) confirmed their nominees (Jae Myung Lee and Moon-soo Kim), and May 28th, 2025 (18 days), when Korean election law prohibits the release of poll results. This implies that an average of 5.2 polls were released per day during this period, rendering poll selection a substantial concern for news outlets. The campaign period was shorter than the usual electoral cycle because the election was held by an emergency process mandated by the Korean Constitution following the removal of President Suk Yeol Yoon on April 4th.

Poll Coverage Data: To obtain poll reporting data, we scraped news articles from *Daum*, the second most popular Internet portal in Korea.³ Our research team has been collecting news stories in real time from the ‘channels’ of 144 news outlets registered on *Daum*. The scraper checks each channel every five minutes and records news stories not yet stored in our database. A total of 701,893 news articles were collected during the 18 days. For each story, the scraper extracted the article text, topic categories assigned by *Daum*, bylines, and publication date and time. To identify poll-related coverage, we retained only stories mentioning at least one polling firm and one candidate from the two major parties, yielding 1,868 stories.

According to *Media Users in Korea 2023* report (Korea Press Foundation, 2023), the vast majority of South Korean news users access news through internet portals offering news aggregation. In 2023, the usage rate of online news aggregators was 69.6%, substantially higher than in many other countries. Among portal news users, 92.1% use *Naver* and 23.1% use *Daum*. Despite *Daum*’s lower usage rate, this study nonetheless relied on it for three reasons.

First, although the two portals differ somewhat in how they evaluate media partners for main-page placement, they are similar in the number of search-affiliated outlets. As of 2023, *Daum* had about 1,350 affiliated media outlets (Ji, 2023, November 24), compared to 981 for *Naver* (AD•PR & Media, 2023, May 15). Second, data from *Daum* may better approximate the full set of published stories. Interviews with media professionals indicated that, while articles are typically submitted to both portals, deletion and content management are concentrated on *Naver* because of its higher visibility. As a result, duplicated or corporate-sponsored pieces circulated through informal journalist networks (*kumi*) often remain available on *Daum* but are removed from *Naver*. Third, practical considerations also mattered. *Naver* allows greater editorial autonomy in formatting, which results in inconsistently structured metadata in article HTML. Although *Naver* provides a news search API with standardized metadata, it severely limits the number of retrievable articles.

Matching News and Polls by AI

As the volume of poll-related news coverage to be matched against the NESDC records increased nearly threefold, completing this task manually would have entailed prohibitive costs. To address this challenge, we relied on the Large Language Model-based artificial intelligence APIs, which have advanced rapidly in recent years, to support the matching process. After testing several services for their ability to extract poll references from news coverage and evaluating both output quality and projected costs at scale, we concluded that the *Claude 3.5 Haiku* model (version released October 22, 2024) proved sufficient for the task. Using the following prompt, the model was applied programmatically to the 1,868 news reports previously collected. The entire extraction process required

³ <https://www.daum.net/>

roughly two days and cost approximately \$50, yielding structured information about the polls reported in each article.

First prompt:

“You are an expert in accurately extracting polling information from news articles. Extract as many details as possible, while providing only information that can be verified as accurate.”

Second prompt:

“The following polling information has been extracted from the article:

{extracted_poll_info}

Please match the above extracted polling information with the complete list of polls provided below ({total_polls} in total):

{poll_reference_str}

For each poll, provide following informations:

- 1. The name of polling firms mentioned in the news;*
- 2. The poll period or dates mentioned in the news;*
- 3. The matched poll ID (from the ID column of poll data)*
- 4. The matching basis (e.g., name of poll firms, dates, candidate support, etc.)*
- 5. The level of matching reliability (high/medium/low)*

If the poll reported in news articles is not included in the registered poll records, indicate that it is a ‘new poll’ and provide all available information extracted from the news article.”

The next step involved automating the matching between the poll information extracted by Claude and the official NESDC records. A match was deemed valid when the reported poll dates overlapped with the dates recorded in the commission’s registry. When this condition was not met, supplementary details such as candidate support in polls, sample sizes, and margins of error were used to facilitate a match.

At the time of data collection, 144 news outlets were officially supplying content to *Daum*, but only 82 of them published election poll coverage. Of these, 67 news outlets published at least three reports on presidential election polls during the study period. The subsequent analyses focus exclusively on this subset of 67 outlets. All these processes yielded 1,728 news stories included in our final dataset.

Constructed Metrics

To measure the recognition level of polling firms, we propose a metric called *coverage centrality*, based on network science. Coverage centrality is essentially a weighted *in-degree centrality* in a bipartite graph between news outlets and polling firms. Simple in-degree centrality would count the number of news reports covering polls conducted by a given firm, indicating how often that firm’s election polls are reported. However, this measure ignores the frequency with which firms conduct polls. As a result, it may inflate the recognition level

of firms that conduct many polls rarely covered, while underestimating firms that conduct fewer polls but are consistently reported. Since we aim to capture the extent of market recognition, this bias is problematic.

We therefore define coverage as a weighted sum:

$$wtd_{ij} = n_{ij} / \# \text{ of polls}_j$$
$$\text{coverage centrality}_j = \sum_i wtd_{ij}$$

where i and j index news outlets and polling firms, respectively, and n_{ij} and $\# \text{ of polls}_j$ denote the frequency with which news outlet i reports results from firm j and the number of polls conducted by firm j during the data collection period.

Further, we estimate polling firms' *consistent* bias toward candidates using the method proposed by Jackman (2005) and Park (2013). This approach is a variant of a *state-space model* for time-series data, which assumes a *latent* true state and treats observed data as *biased realizations* of the truth. In our context, the latent state represents the underlying public opinion toward presidential candidates, and each poll result is modeled as the sum of this latent opinion and a firm-specific bias. Operationally, a firm's bias (also called 'house effect') is estimated as a systematic deviation from contemporaneous polls towards one of the two candidates from the largest two parties. Park (2013) applied this method in a public website jointly developed with MBC. Methodological details are also available on the *Public Opinion M* website and its accompanying Github page.⁴

To measure news outlets' level of institutionalization, we combined organizational data from the *Korea Press Yearbook 2022* (Korea Press Foundation Industry Analysis Team, 2022) with web-scraped sources. First, we used organizational age as a proxy for institutional legacy, and the number of international correspondents and press awards won as indicators of commitment to traditional reporting. Further, we calculated the number of unique reporter names from the byline data as a proxy for a news organization's resource scale. Because it is difficult to examine interaction effects between each institutional variable and polling firms' traits (H2), we conduct dimensionality reduction to create composite institutionalization scores, which we explain in the following sections.

To test how news outlets' organizational traits and polling firms' recognition affect diversity of poll coverage (H1 & H2), we constructed a *poll diversity* measure as the inverse of Herfindahl-Hirschman index (HHI) for each outlet:

$$\text{poll diversity}_i = (1 - \sum_j p_{ij}^2) \times 100$$

⁴ https://poll-mbc.co.kr/2025_여론M_매뉴얼.pdf & <https://github.com/jongheepark/poll-MBC>

where p_{ij} denotes the share of firm j 's polls in outlet i 's poll coverages. To examine the moderating effect of polling firm recognition (H2), we calculated *mean coverage centrality for each outlet*, where weights are given by the number of reports covering each firm.

Statistical Models

To test H1 and H2, we estimate outlet-level linear regressions with poll diversity as the dependent variable ($n=67$). For H2, however, including interaction terms with five institutional variables would substantially reduce degrees of freedom given the small sample size. We first use factor analysis to obtain institutionalization-level scores, and then generate interaction terms. For RQ1, which concerns assortativity between news outlets and polling firms, we estimate a dyadic regression model where the dependent variable is the number of polls from firm j covered by outlet i . Independent variables include firm-level traits; the number of polls conducted, house effects (i.e., consistent bias), and coverage centrality (i.e., recognition level).

The regression requires additional adjustments because the dependent variable contains many zeros – that is, many firms' polls are never reported by many outlets. To address this, we adopt a Tobit model for censored data (Tobin, 1958). Moreover, to allow for heterogeneity in outlets' responses to firms with varying house effects, we include a random coefficient on house effects. This specification is motivated by two considerations. First, unlike legacy media, many emergent outlets have uncertain ideological orientations. Second, even for outlets with relatively well-known stances, it is unclear whether they systematically prefer polls from firms whose biases align with their slant. Our previous work in the Korean context suggests that poll selection is not primarily driven by ideology (Lee, Zhang & Pak, 2024).

Because there is no clear consensus on estimating random-effects Tobit models in a frequentist framework, we use a Bayesian approach implemented in *Stan* (Stan Development Team, 2024) via the *brms* package in R (Bürkner, 2017) with uninformative priors.

5 Results

Description of Data

Before addressing hypotheses and research questions, we first present the overall profile of our Korean press-polling firm dataset. The data are inherently dyadic, involving two distinct sets of agents – news outlets and polling firms – and the relationships formed through outlets' coverage of election polls. In this section, we present the organizational traits of these two sets of agents and provide descriptive evidence that suggests patterns in their dyadic relationships.

Table 1 summarizes the characteristics of the polling firms. The 24 firms that conducted at least one election poll during the study period conducted an average of 3.76 polls, with the number ranging from one to nine. This heterogeneity is even more pronounced when coverage by news outlets is considered. While the average number of times a firm's polls were covered was 89.44, some firms' polls were never reported. Similarly, while the mean probability of a poll being covered was 0.13, the most covered firm – likely considered the most credible – had a coverage probability of 0.59, meaning that its polls were reported by more than half of the outlets. The wide range of coverage centrality

underscores the variation in the extent to which the polling firms are recognized and trusted by news outlets.

Table 1. Characteristics of polling firms (n=24)

	Mean	SD	Min	Max
# polls	3.76	2.18	1	9
# covers	89.44	187.81	0	800
Prop. covered	0.13	0.17	0	0.59
Coverage centrality	23.94	39.33	0.67	142.50

Second, there is also substantial heterogeneity in news outlets’ organizational traits (Table 2). Among the 67 organizations in our dataset, some are legacy outlets with long histories, large reporting staffs including international correspondents, and records of journalism awards. Others are emergent outlets less bound by journalistic conventions, with some not even including bylines in their stories. Variation is also evident in polling coverage. While the mean diversity score is 0.64, some outlets had a score of zero, indicating reliance on only one polling firm. This suggests that certain outlets may be uninterested in presenting a balanced view of public opinion even when competing assessments exist. Moreover, coverage centrality values ranged widely, from 0.89 to 626.68, indicating that some outlets rely on emergent polling firms not widely trusted by others.

Table 2. Characteristics of news outlets (n=67).

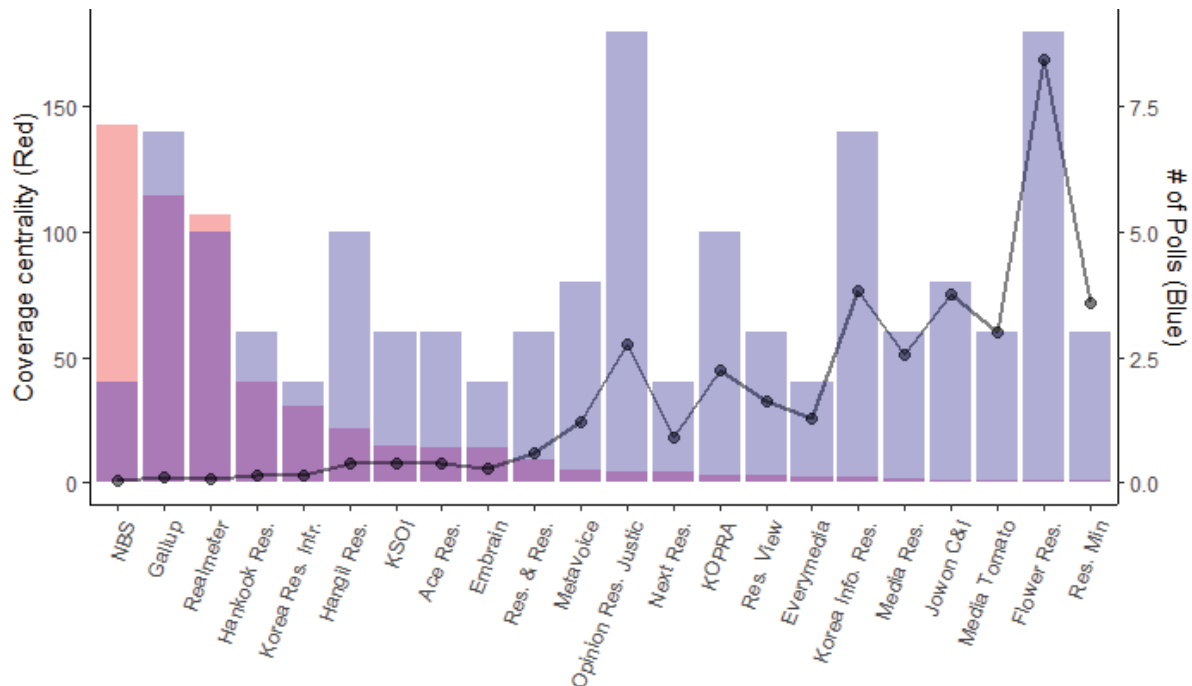
	Mean	SD	Min	Max
# articles	4956.21	2.18	1	9
# reporters	82.64	187.81	0	800
# correspondents	1.42	0.17	0	0.59
Awards	1.69	39.33	0.67	142.50
Age	30.92	21.02	3	121
Poll covers	33.56	28.33	1	143
Diversity	0.64	0.22	0	0.86
Centrality	126.174	111.25	0.89	626.68

The observations so far suggest substantial heterogeneity among *both* polling firms and news outlets. This indicates that treating poll coverage solely as a one-sided information selection problem is insufficient for understanding the role of the emergent pollsters. We instead conceptualize this as an interorganizational linking problem between the two distinct groups, which shapes how readers perceive public opinion depending on their media diet – that is, the mix of legacy and emergent outlets.

To examine which types of polling firms are more likely to be covered, Figure 1 compares firms’ coverage centrality (red) and the number of polls conducted (blue). As expected, well-established firms, such as *Gallop*, *Realmeter*, and *Korea Research*, received the most coverage. *The National Barometer Survey (NBS)* is jointly conducted by *Embrain*, *Kstat*, *Korea Research*, *Hankook Research*. Interestingly, coverage centrality is not proportional to – and may even be inversely related to – the number of polls conducted. This suggests

that emergent polling firms often conduct a substantial number of election polls, yet most of these are not reported by major news outlets.

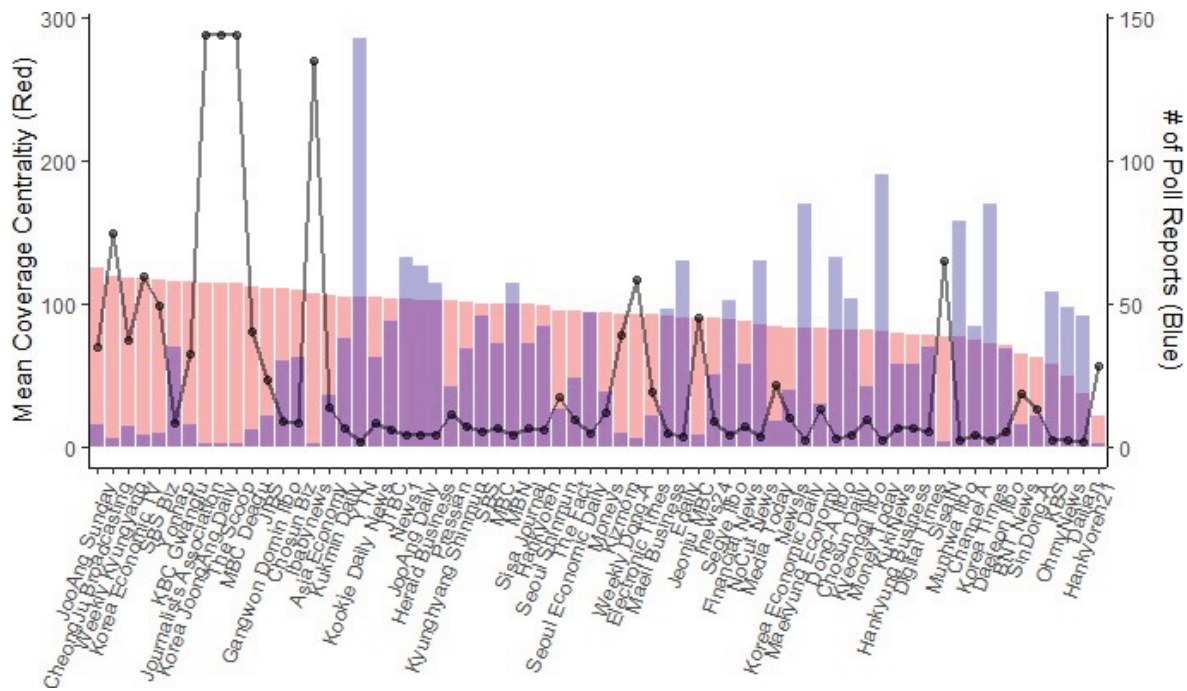
Figure 1. Number of polls (blue) and coverage centrality (red) of polling firms.



Note: The line graph shows the ratio of poll counts (blue) to coverage centrality (red).

Examining the types of polling firms covered by news outlets provides insights into *assortativity* between outlets and pollsters. Figure 2 plots the mean coverage centrality of polling firms covered by outlets (red; weighted by the number of polls covered) and the number of election poll reports (blue). Since outlets on the x-axis are ordered by the number of reporters – a proxy for organizational size – the frequency of election poll reporting does not appear associated with organizational size. This suggests that election polls are a popular news topic across outlets, regardless of organizational traits. By contrast, mean coverage centrality declines from left to right: while most outlets cover election polls to a similar extent, smaller outlets tend to rely more on less recognized polling firms. Emergent outlets, compared to legacy media, rely more heavily on emergent polling firms, a pattern consistent with assortativity (RQ1).

Figure 2. The number of polls covered (blue) and the mean coverage centrality of covered polling firms (red) for news outlets, sorted by the number of reporters.



Note: The line graph shows the ratio of mean coverage centrality (red) to news counts (blue).

Metrics for Institutionalization Level

Before conducting formal statistical tests, we reduce the dimensionality of organizational variables for news outlets. We first report an exploratory factor analysis (EFA) with two factors (Table 3). We focus on the $K=2$ model because it is the maximum number of factors feasible with only five observed variables, and the $K=1$ assumption produced low loadings for *age* and *awards*. The number of articles, reporters, and correspondents all loaded strongly on Factor 1 (all > 0.8), which reflects organizational resources. Age and the number of awards loaded more heavily on Factor 2 (0.553 and 0.653), but these loadings were weaker than those for Factor 1, and Factor 2 explained substantially less variance (0.209 $<$ 0.440).

Table 3. Exploratory Factor Analysis (K=2, varimax)

	Uniquess	Factor Loadings	
		Factor 1	Factor 2
# articles	0.142	0.919	-0.117
# reporters	0.175	0.827	0.376
# awards	0.542	0.175	0.653
# correspondents	0.179	0.801	0.423
Age	0.714		0.553
SS loadings		2.202	1.046
Proportion var.		0.440	0.209
Cumulative var.		0.440	0.650

To further assess whether the two-factor latent structure was supported by the data, we conducted a confirmatory factor analysis (CFA) with the two-factor specification (Table 4). However, the second factor, comprising *awards* and *age*, exhibited poor validity indicators: composite reliability (CR) was below 0.60 (Nunnally, 1978; Hair, Howard & Nitzl, 2020), and convergence validity score (average variance extracted; AVE) was below 0.5 (Fornell & Larcker, 1981).

Table 4. Confirmatory Factor Analysis (K=2)

Latent	Observed	Estimate	SE	CR(p)	AVE	CR
Resources	# Articles	0.753	0.107	7.066***	0.745	0.897
	# Reporters	0.916	0.098	9.388***		
	# Correspondents	0.891	0.099	8.989***		
Maturity	# Awards	0.760	0.212	3.583***	0.402	0.560
	Age	0.464	0.162	2.872**		

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Thus, we excluded this factor and conducted a CFA on a one-factor model including *the number of articles, reporters, and correspondents*. This final model (Table 5) demonstrated satisfactory reliability (AVE = 0.748, CR = 0.898) and excellent fit indices (SRMR=0, CFI=1, TLI=1, RMSEA=0), supporting a unidimensional construct of organizational size.

Table 5. Confirmatory Factor Analysis (K=1)

Latent	Observed	Estimate	SE	CR(p)	AVE	CR
Resources	# Articles	0.767	0.106	7.240***	0.748	0.898
	# Reporters	0.921	0.098	9.401***		
	# Correspondents	0.879	0.100	8.770***		

Outlets' Poll Diversification (H1/H2)

To test our hypotheses on outlets' diversification of polling firms as a function of their institutionalization level and the recognition level of polling firms (H1/H2), we estimated outlet-level Tobit regressions with poll diversity as the dependent variable. Model 1 and 2 included the raw institutional variables and the mean coverage centrality of firms whose polls were reported. The difference between the two models is whether the variables excluded from the constructed institutional maturity measure (i.e. the number of *awards* and *age*) were included. In both models, institutional variables were not statistically significant predictors of poll diversification. By contrast, mean coverage centrality had a statistically significant effect ($\beta = 0.08$, $p < 0.001$). This suggests that outlets placing greater value on recognized polling firms also tend to diversify their sources. It is noteworthy that the estimated effect size (0.08) and its standard error were consistent across all models, indicating that this result is robust across model specifications.

Table 6. Outlet-level OLS Regression Results (H1/H2)

	Model 1	Model 2	Model 3	Model 4
(Intercept)	52.38*** (3.89)	47.81*** (5.486)	53.65*** (3.887)	59.8*** (3.937)
# articles	0.00 (0.000)	0.00 (0.000)		
# reporters	0.03 (0.024)	0.02 (0.026)		
Correspondents	-0.71 (0.64)	-0.72 (0.712)		
Awards		0.02 (0.495)		
Age		0.13 (0.116)		
Org. Resources			0.39 (2.754)	25.66*** (7.380)
Mean Centrality (Polling firm)	0.08** (0.025)	0.08** (0.026)	0.08** (0.024)	0.08*** (0.022)
Resources × Mean Cent.				-0.16* (0.044)
R ²	0.21	0.23	0.19	0.33
N	67	67	67	67

The statistically insignificant results for the raw institutional variables may reflect high correlations among the variables and the limited sample size. When we instead included the unidimensional organizational resources, measure constructed by CFA (Model 3), the effect remained insignificant ($\beta = 0.39$, $p = 0.887$). However, when we added the interaction between resources and mean centrality in the final model testing H2 (Model 4), the resources measure became strongly significant ($\beta = 25.66$, $p < 0.001$). This result is unlikely to be due to chance, as R^2 increased substantially from 0.21 in Model 1 to 0.33 in

Model 4. In other words, incorporating the interaction term improves model fit more than including multiple observed variables. We therefore conclude that outlets with greater organizational resources diversify the polling firms whose election polls they report, consistent with H1. Nevertheless, this finding should be interpreted with caution, given the small sample size.

Further, the significant negative effect of the interaction term ($\beta = -0.16, p < 0.05$) supports H2. That is, when news outlets place greater emphasis on polling firms' recognition, even affluent outlets tend to diversify less. This is an intuitive result if one considers the trade-off between the value of diversification and the validity of poll results.

Taken together, the evidence for H1 and H2 shows that organizational resources matter for election poll reporting. Outlets with large reporting staffs and robust editorial processes value both validity and diversity, while navigating the trade-off between them. In contrast, prior work has noted that outlets with insufficient resources often resort to horse-race poll coverage as inexpensive filler content (Searles et al., 2016). Juxtaposed with this observation, our findings suggest that emergent online outlets may rely on results from emergent pollsters without regard for either diversity or validity. This leads us to RQ1, concerning assortativity between news outlets and polling firms.

Assortativity between Media and Poll Houses (RQ1)

To test assortativity (RQ1), we report dyadic Tobit regression results in Table 7. Assortativity would be indicated by a significant interaction between polling firms' coverage centrality and news outlets' organizational resources. Model 1 is a baseline specification without the interaction term or random coefficients. Model 2 adds the interaction term of interest, and Model 3 further includes random coefficients for the intercept and polling firms' bias (house effect), as described in the methods section.

For comparability, we report Bayesian R^2 values along with p-values based on the MAP estimate, which we interpret in a frequentist manner. Given the sample size and the use of uninformative priors, the results should be driven primarily by the data rather than prior assumptions.

Table 7. Dyadic tobit regression model results for RQ1

	Model 1	Model 2	Model 3
(Intercept)	-2.31*** (0.125)	-2.26*** (0.122)	-2.34*** (0.179)
House effect (<i>j</i>)	0.11** (0.035)	0.11** (0.034)	0.14*** (0.039)
Coverage centrality (<i>j</i>)	0.04*** (0.002)	0.04*** (0.002)	0.04*** (0.002)
Org. resources (<i>i</i>)	0.30*** (0.063)	0.22 (0.081)	0.11 (0.164)
Coverage centrality × Resources		0.01*** (0.001)	0.01*** (0.001)
$\log \sigma$	0.65*** (0.037)	0.63*** (0.036)	0.47*** (0.037)
Random effect	X	X	✓
R ²	0.189	0.193	0.607
N	1,608	1,608	1,608

Note 1: For Model 3, mean Monte Carlo samples and their standard deviations are reported.

Note 2: McFadden's pseudo R²s are reported for Model 1 and Model 2, and Bayesian R² is reported for Model 3.

Note 3: Statistical significance codes for Model 3 are based on a Bayesian p-value based on the density at the maximum a posteriori (MAP).

Overall, the estimated results are consistent across model specifications. First, polling firms favorable to Kim tend to be preferred on average. This result should be interpreted with caution, however, as it reflects only an aggregate tendency and individual outlets may behave differently. Still, the consistent result in Model 3 ($\beta = 0.14$, $p < 0.001$) suggests that a preference for Kim-favored firms is an overall pattern.

As expected, polling firms' recognition level had a positive effect on the likelihood of coverage ($\beta = 0.04$, $p < 0.001$ across all three models). In Model 1, organizational resources had a positive and statistically significant effect on the probability that polls were covered. However, consistent with our expectation in framing RQ1, this effect did not apply uniformly across all polling firms. When the interaction between outlets' organizational resources and firms' recognition level was included in Models 2 and 3, the interaction term was positive and statistically significant (e.g., $\beta = 0.01$, $p < 0.001$ in Model 3) while the main effect of organizational resources disappeared (e.g., $\beta = 0.11$, $p = 0.825$ in Model 3). This supports our assortativity hypothesis: well-resourced outlets disproportionately prefer recognized polling firms, whereas emergent outlets are more likely than legacy outlets to cover polls from emergent firms.

6 Discussion

With innovations in polling technologies and declining costs, many more polling firms exist today than a decade ago (Prosser & Mellon, 2018). While greater variety in assessments of public opinion could aid voter decision-making, competing results from different firms – and their recent high-profile failures – often generate more confusion about public

opinion than clarity, thereby complicating voting decisions (Hillygus, 2011). This situation resembles the proliferation of emergent news outlets at the onset of the online news era three decades ago: the diversity of perspectives was expected to enhance citizens' understanding of public affairs, but instead often generated confusion in the public sphere (Happer & Philo, 2013).

This creates a *dual uncertainty problem*: it is unclear both why many polls fail to reflect public opinion (while some succeed) and how news outlets choose among competing polls to report. Thus, uncertainty about what citizens think arises not only from outlets' one-sided selection of polls, as often assumed in the journalism literature, but also from the dyadic relationships between two sets of organizations, each with its own biases. We therefore adopt an organizational perspective that conceptualizes election poll reporting as inter-institutional relationships rather than a one-sided information selection problem. Overall, we find that legacy media diversify the polling firms whose results they cover, while still placing more weight on widely recognized ones. By contrast, emergent media rely more on a limited set of pollsters and are more likely than legacy media to prefer emergent pollsters.

To news outlets, polling firms are information sources with qualitative differences that provide competing facts. Although reporting poll results is a relatively inexpensive way to fill news space and thus economically appealing (Searles et al., 2016), legacy media are more cautious in reporting competing information, as the journalism literature suggests (Berkowitz, 2019; Kurpius, 2002). Accordingly, we expected that news outlets with a higher institutionalization level would diversify the polling firms they cover, as they diversify other information sources (H1). This expectation was supported. Further, well-resourced outlets not only diversify but also consider polling firms' recognition, producing the moderating effect (H2). Specifically, when outlets report more polls from highly recognized firms, they diversify less. This indicates that well-resourced outlets weigh the trade-off between diversity and credibility. The finding underscores that journalistic professionalism matters in an environment of poll overload.

However, outlet-level analysis alone does not reveal specific preferences for particular types of pollsters. We therefore examined dyadic relationships to assess how emergent pollsters' results are utilized. We proposed the *assortativity hypothesis*, namely that legacy media prefer legacy pollsters, whereas emergent media are more likely than legacy media to prefer emergent pollsters (RQ1). This expectation was supported. As shown in our descriptive analysis, many polls conducted by emergent pollsters were never covered by most outlets. This is unsurprising: the news value of polls depends largely on the accumulated reputation of polling firms, given the difficulty of assessing methodological validity under tight reporting schedules (Rosenstiel, 2005), and reporting inaccurate 'facts' may damage outlets' reputations (Gentzkow & Shapiro, 2006). Thus, it is natural to question the purpose served by emergent pollsters whose results are little known to the public. Relatedly, some argue that emergent pollsters belong to alternative news ecosystems that supply "alternative facts" to relatively segregated audiences, alongside emergent outlets including not only unconventional organizations but also short-form video platforms and YouTube channels (de León, Makhortykh & Adam, 2024; Miró-Llinares and Aguerri, 2023). Our assortativity findings partially support this suspicion: emergent polls are used almost exclusively by emergent outlets, though we did not directly test whether such pairs form disconnected information ecosystems. This presents an important avenue for future research.

At minimum, our results suggest that as emergent media with less conventional editorial processes gain prominence, audiences will increasingly encounter polls from emergent pollsters – and the reverse is also true. Although polls from emergent firms are not yet

widely reported, it is important to monitor their accuracy and patterns of coverage. Such monitoring is especially critical because economic incentives generated by technological change often override professional norms during periods of disruption in the media environment (Lowrey, 2012).

To continuously track diverse polls and their reporting, it is important to construct an automated workflow that matches polling data with news data and incorporates related metadata into real-time analysis and visualization. Whereas previous studies examined only a few traditional outlets and incumbent pollsters (Searles et al., 2016; Larsen and Fazekas, 2020), our focus on emergent pollsters in an online news system filled with emergent media raises the substantial cost of matching poll-news data pairs. Consider an environment with 10 outlets and 10 pollsters versus one with 100 outlets and 30 pollsters. An AI-powered automated system, as suggested in this study, offers a relatively inexpensive and robust solution for such real-time tasks. This approach could support the development of practical systems that provide voters with aggregated information, helping them navigate an era of competing truths exacerbated by the *dual uncertainty problem* we identified.

Viewing election poll reporting as a case of doubly uncertain gatekeeping was made possible by adopting a theoretical perspective that treats pollsters, like news outlets, as institutional actors. Whereas political science and communication research have often conceptualized polling firms as *measurement machines* producing numerical estimates with varying biases due to methodological defects (Cantrell, 1989), few attempts have captured them as institutional entities. This was understandable in an era dominated by a handful of established pollsters with long histories and recognized expertise. Today, however, with a proliferation of firms differing in institutional history, staffing, methods, and results, it is increasingly untenable to ignore organizational traits. Moreover, since poll results reach voters only through news outlets, perceptions of public opinion inevitably emerge from the combined biases of both institutions. Our findings underscore the value of incorporating an organizational perspective into the study of public opinion.

Like any study, this one has limitations. First, the number of outlets that reported polls is somewhat limited ($n = 67$). Our list is comprehensive in that it includes all outlets reporting election polls on the Daum news aggregation service. However, Naver — the more popular aggregator — maintains a different list of outlets that is not necessarily shorter. Greater generalizability would require combining data from both sources, but this was not feasible due to limited resources. Moreover, the outlet list is short in absolute terms, which may raise questions about the validity of our statistical analysis. We urge readers not to rely blindly on the statistical significance of our outlet-level models (H1/H2), but to interpret them alongside more robust descriptive and visual evidence. In addition, our measure of institutionalization level was essentially limited to organizational size (or resources), making our conclusions about journalistic institutionalization somewhat unidimensional. This issue, too, could be addressed with a larger sample. Finally, the election cycle under study was atypical. The 21st Korean presidential election followed a period of political turmoil marked by President Suk Yeol Yoon's declaration of martial law and subsequent impeachment. This peculiarity made the campaign relatively one-sided in favor of the Democratic Party and shortened by law. Because our analysis focused on organizational traits rather than partisan motivations, we do not believe this context drove our results. Nevertheless, consolidating data from other election cycles in future research would enhance generalizability.

Finally, the study's focus on South Korea — a specific polity and media environment — raises further generalizability concerns. For instance, given the two-party system, highly polarized opinion, and numerous new polling firms, our findings may be more comparable

to the U.S. context than to Europe or Japan. In addition, media environments with a larger number of legacy regional and local outlets, each with longer histories, are likely to exhibit greater heterogeneity in institutionalization level and thus may resemble our findings more closely than less. We therefore argue that our findings from South Korea are not unique to that country.

Not only news outlets but also polling firms now contribute to the multiplicity of information sources, and neither is immune to bias. This dual uncertainty may undermine voters' ability to gauge fellow citizens' opinions, making it more difficult to base political decisions on a clear estimation of their choices' value (Meffert & Gschwend, 2011; Moy & Rinke, 2012). By framing poll reporting as an inter-organizational linkage between outlets and pollsters, we situate this issue within organizational studies and examine how the changing news environment conditions the information available to citizens and shapes their perception of public opinion.

The assortative patterns we observed suggest that emergent actors' behavior may exacerbate the mobilization of disguised public opinion rather than treating polls as factual measurements. At the same time, our results imply that journalistic professionalism encourages outlets to balance diversity of coverage with the credibility of the polls they report. Because meticulous poll selection is costly in terms of budget and editorial capacity, journalist training and public resources that reduce the cost of identifying credible polls are needed.

7 References

- AD•PR & Media. (2023, May 15). Naver News partnered with 981 news outlets as of the first half of 2023, of which approximately 250 were independent media excluding major and institutional outlets. AP News. Retrieved from <https://www.apnews.kr/news/articleView.html?idxno=3010007> [AP신문 (2023, 05, 15) 2023년 상반기 네이버 뉴스 제휴 언론사 981개 메이저 매체, 기관지 등을 제외한 독립언론사는 250여개.]
- American Association for Public Opinion Research. (2021). Code of professional ethics and practices. AAPOR. Retrieved from <https://wapor.org/about-wapor/code-of-ethics/>
- Anderson, C. W. (2013). *Rebuilding the News: Metropolitan Journalism in the Digital Age*. Temple University Press.
- Beckers, K. (2021). "According to the polls..." opinion poll coverage in network evening news during the 2020 US Election Campaign. *International Journal of Public Opinion Research*, 33(4), 1050-1061.
- Benson, R. (2004). Bringing the sociology of media back in. *Political Communication*, 21(3), 275-292.
- Berkowitz, D. (2019). Reporters and their sources. In *The handbook of journalism studies* (pp. 165-179). Routledge.
- Boehmer, J., Carpenter, S., & Fico, F. (2019). More of the same? Influences on source use and source affiliation diversity in for-profit and nonprofit news organizational content. *Journalism Studies*, 20(2), 173-192.
- Brodie, M., Parmelee, L. F., Brackett, A., & Altman, D. E. (2001). Polling and democracy. *Public Perspective*, 12(4), 10-24.
- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80, 1-28.
- Cantrell, P. D. (1989). *Political polling in America: A study of institutional structures and processes*. New School for Social Research.
- Cho, B., Ryoo, S. Y., & Kim, K. K. (2017). Interorganizational dependence, information transparency in interorganizational information systems, and supply chain performance. *European Journal of Information Systems*, 26(2), 185-205.

- de León, E., Makhortykh, M., & Adam, S. (2024). Hyperpartisan, alternative, and conspiracy media users: An anti-establishment portrait. *Political Communication*, 41(6), 877-902.
- Dimitrova, D. V., & Strömbäck, J. (2009). Look who's talking: Use of sources in newspaper coverage in Sweden and the United States. *Journalism practice*, 3(1), 75-91.
- Druckman, J. N., Fein, J., & Leeper, T. J. (2012). A source of bias in public opinion stability. *American Political Science Review*, 106(2), 430-454.
- FiveThirtyEight. (2024, October 26). What the polls can't tell us about America's election. Retrieved from <https://www.ft.com/content/870dabd1-93d2-4f88-9d10-b576ec5713c9>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50.
- Frankovic, K. A. (2005). Reporting "The polls" in 2004. *Public Opinion Quarterly*, 69(5), 682-697.
- Gabay, N. (2022). The effect of news media political bias on pre-elections poll results: evidence from the 2019-20 Israeli elections. *Israel Affairs*, 28(6), 917-939.
- Gans, H. J. (2004). *Deciding what's news: A study of CBS Evening News, NBC Nightly News, Newsweek, and Time*. Northwestern University Press.
- Gentzkow, M., & Shapiro, J. M. (2006). Media bias and reputation. *Journal of Political Economy*, 114(2), 280-316.
- González-Bailón, S., Lazer, D., Barberá, P., Zhang, M., Allcott, H., Brown, T., ... & Tucker, J. A. (2023). Asymmetric ideological segregation in exposure to political news on Facebook. *Science*, 381(6656), 392-398.
- Graham, M. H., Hillygus, D. S., & Trexler, A. (2024). Misleading polls in the media: Does survey clickbait have social consequences?. *Public Opinion Quarterly*, 88(2), 315-336.
- Guess, A. M., Nyhan, B., & Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign. European Research Council. Retrieved from <http://www.ask-force.org/web/Fundamentalists/Guess-Selective-Exposure-to-Misinformation-Evidence-Presidential-Campaign-2018.pdf>
- Guo, L., & Vargo, C. (2020). "Fake news" and emerging online media ecosystem: An integrated intermedia agenda-setting analysis of the 2016 US presidential election. *Communication Research*, 47(2), 178-200.
- Hair Jr, J. F., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, 109, 101-110.
- Happer, C., & Philo, G. (2013). The role of the media in the construction of public belief and social change. *Journal of Social and Political Psychology*, 1(1), 321-336.
- Hillygus, D. S. (2011). The evolution of election polling in the United States. *Public opinion quarterly*, 75(5), 962-981.
- Irwin, G. A. (2002). According to the polls: The influence of opinion polls on expectations. *The Public Opinion Quarterly*, 66(1), 92-104.
- Iyengar, S., Norpoth, H., & Hahn, K. S. (2004). Consumer demand for election news: The horserace sells. *The Journal of Politics*, 66(1), 157-175.
- Jackman, S. (2005). Pooling the polls over an election campaign. *Australian Journal of Political Science*, 40(4), 499-517.
- Ji, M. (2023, November 24). Daum search displays articles exclusively from 150 partnered news outlets. Dong-A Ilbo. Retrieved from <https://www.donga.com/news/Society/article/all/20231124/122328912/1> [지민구 (2023, 11, 24). 다음, 150개 뉴스제휴 언론사 기사만 검색 노출. <동아일보>.]
- Korea Press Foundation (2023). *Media users in Korea 2023*. Korea Press Foundation. Korea Press Foundation Industry Analysis Team (2022). *The Korea press yearbook 2022*. Korea Press.
- Kurpius, D. D. (2002). Sources and civic journalism: Changing patterns of reporting? *Journalism & Mass Communication Quarterly*, 79(4), 853-866.
- Larsen, E. G., & Fazekas, Z. (2020). Transforming stability into change: How the media select and report opinion polls. *The International Journal of Press/Politics*, 25(1), 115-134.

- Lee, S., Zhang, S., & Pak, C. (2024). Do Korean news media selectively report election polls? A Bayesian analysis of house effects, disparities in poll outcomes, and their relationship to reporting frequency during the 20th presidential election. *Korean Journal of Journalism and Communication*, 68(6), 251-287.
[이신애·장스차오·박찬경 (2024). 한국 언론은 선거 여론조사 결과를 선택적으로 보도하는가? 제20대 대선에서 조사기관 편향 및 조사결과의 격차와 보도 빈도의 관계에 대한 베이지안 분석. <한국언론학보>, 68권6호, 251-287.]
- Lepore, J. (2015, November 8). Politics and the new machine. *The New Yorker*. Retrieved from <https://www.newyorker.com/magazine/2015/11/16/politics-and-the-new-machine>
- Lewis, J., Williams, A., & Franklin, B. (2008). A compromised fourth estate? UK news journalism, public relations and news sources. *Journalism Studies*, 9(1), 1-20.
- Lowrey, W. (2012). Journalism innovation and the ecology of news production: Institutional tendencies. *Journalism & Communication Monographs*, 14(4), 214-287.
- Mathisen, B. R. (2023). Sourcing practice in local media: Diversity and media shadows. *Journalism Practice*, 17(4), 647-663.
- Meffert, M. F., & Gschwend, T. (2011). Polls, coalition signals and strategic voting: An experimental investigation of perceptions and effects. *European Journal of Political Research*, 50(5), 636-667.
- Meffert, M. F., Huber, S., Gschwend, T., & Pappi, F. U. (2011). More than wishful thinking: Causes and consequences of voters' electoral expectations about parties and coalitions. *Electoral Studies*, 30(4), 804-815.
- Miró-Llinares, F., & Aguerri, J. C. (2023). Misinformation about fake news: A systematic critical review of empirical studies on the phenomenon and its status as a 'threat'. *European Journal of Criminology*, 20(1), 356-374.
- Moy, P., & Rinke, E. M. (2012). Attitudinal and behavioral consequences of published opinion polls. In *Opinion polls and the media: Reflecting and shaping public opinion* (pp. 225-245). London: Palgrave Macmillan UK.
- Newman, M. E. J. (2003). Mixing patterns in networks. *Physical Review E*, 67(2), 026126. Nunnally, J. C. (1978). *Psychometric theory: 2d Ed.* McGraw-Hill.
- O'Connor, A., & Shumate, M. (2020). A multidimensional network approach to strategic communication. In *Future directions of strategic communication* (pp. 71-88). Routledge.
- O'Neill, D., & Harcup, T. (2009). News values and selectivity. In *The handbook of journalism studies* (pp. 181-194). Routledge.
- Park, J. H. (2013). House bias in the 2012 presidential campaign polls in South Korea. *Survey Research*, 14(1), 1-30. [박종희 (2013). 제18대 대선 여론조사에서 나타난 조사기관 편향. <조사연구>, 14권 1호, 1-30.]
- Patterson, T. E. (2005). Of polls, mountains: U.S. journalists and their use of election surveys. *Public Opinion Quarterly*, 69(5), 716-724.
- Petre, C. (2015). *The traffic factories: Metrics at Chartbeat, Gawker Media, and The New York Times*. Tow Center for Digital Journalism, Columbia Journalism School.
- Pew Research Center. (2023, April 19). How public polling has changed in the 21st century. Retrieved from <https://www.pewresearch.org/methods/2023/04/19/how-public-polling-has-changed-in-the-21st-century/>
- Prosser, C., & Mellon, J. (2018). The twilight of the polls? A review of trends in polling accuracy and the causes of polling misses. *Government and Opposition*, 53(4), 757-790.
- Rao, H. (2017). Interorganizational ecology. In J. A. C. Baum (Ed.), *The Blackwell companion to organizations*, (pp. 541-556). Wiley-Blackwell.
- Rosenstiel, T. (2005). Political polling and the new media culture: A case of more being less. *Public Opinion Quarterly*, 69(5), 698-715.
- Schudson, M. (1989). The sociology of news production. *Media, Culture & Society*, 11(3), 263-282.
- Schumann, D., Shamon, H., & Hake, J.-F. (2020). The importance of house effects for repeated public opinion surveys. *International Journal of Public Opinion Research*, 32(4), 769-779.

- Searles, K., Ginn, M. H., & Nickens, J. (2016). For whom the poll airs: Comparing poll results to television poll coverage. *Public Opinion Quarterly*, 80(4), 943-963.
- Searles, K., Smith, G., & Sui, M. (2018). Partisan media, electoral predictions, and wishful thinking. *Public Opinion Quarterly*, 82(S1), 888-910.
- Shirani-Mehr, H., Rothschild, D., Goel, S., & Gelman, A. (2018). Disentangling bias and variance in election polls. *Journal of the American Statistical Association*, 113(522), 607-614.
- Shoemaker, P. J., & Vos, T. P. (2009). *Gatekeeping theory*. Routledge.
- Singer, J. B. (2005). The political j-blogger: 'Normalizing' a new media form to fit old norms and practices. *Journalism*, 6(2), 173-198.
- Stan Development Team (2024) Stan modeling language users guide and reference manual. Version 2.35. Retrieved from <https://mc-stan.org>
- Strömbäck, J., & Kaid, L. L. (Eds.). (2009). *The handbook of election news coverage around the world*. Routledge.
- Tandoc, E. C., Jr. (2014). Journalism is twerking? How web analytics is changing the process of gatekeeping. *New Media & Society*, 16(4), 559-575.
- The Washington Post. (2025, March 25). Policies and standards. Retrieved from <https://www.washingtonpost.com/policies-and-standards/>
- Thomson, T. J., Phan, A., Holland-Batt, S., Seevinck, J., Miller, E., & Regi, S. (2025). Who or what gets referenced by whom, how often, and in which ways? Exploring journalists' sourcing practices in the context of aged care coverage. *Journalism Practice*, 19(4), 822-842.
- Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometric: Journal of the Econometric Society*, 24-36.
- Toff, B. (2019). The 'Nate Silver effect' on political journalism: Gatecrashers, gatekeepers, and changing newsroom practices around coverage of public opinion polls. *Journalism*, 20(7), 873-889.
- Tourangeau, R., Presser, S., & Sun, H. (2014). The impact of partisan sponsorship on political surveys. *Public Opinion Quarterly*, 78(2), 510-522.
- Wang, R., & Shin, J. (2023). Determinants of alliance formation and dissolution among international health organizations: The influence of homophily and institutional power in affinity communication networks. *Management Communication Quarterly*, 37(4), 913-940.
- WAPOR. (2023). *Freedom to conduct and publish opinion polls: Global report 2023*. Retrieved from <https://wapor.org/publications/freedom-to-publish-opinion-polls/freedom-report-2023/>
- Weber, M. S. (2012). Newspapers and the long-term implications of hyperlinking. *Journal of Computer-Mediated Communication*, 17(2), 187-201.
- Zeng, J., & Schäfer, M. S. (2021). Conceptualizing "dark platforms". Covid-19-related conspiracy theories on 8kun and Gab. *Digital Journalism*, 9(9), 1321-1343.
- Zerback, T., & Schneiders, P. (2024). Noticed and appreciated? The role of argument diversity in enhancing news credibility and reader satisfaction. *Journalism*, 25(12), 2523-2542.