Identifying Instrumental Variables for Social Movements*

Detecting Protest Events Using Pre-trained Language Model

Kazuhiro Terashita †

Abstract

This study attempts to explore instrumental variables in social movement research using a unique Korean dataset. With the development of causal identification methods, social scientists studying social movements have increasingly sought to develop new methodologies. While many researchers have used precipitation as an instrumental variable, its effect on social movement variables beyond mobilization remains unclear. Furthermore, there has been limited progress in examining random factors other than precipitation. Therefore, this paper estimates the effects of precipitation and temperature on protest occurrence, media coverage, and protest intensity in South Korea, utilizing data originally generated from Korean newspaper articles. The analysis reveals that these exogenous and random variables do not significantly affect event count, media coverage, or protest intensity—all important factors in social movements. These results provide methodological implications for examining social movement aspects beyond mobilization and highlight the need to consider the nature of civil society and specific protest characteristics.

*This is a working paper for JACP Annual Meeting 2025 and citation is declined. This study was supported by JSPS KAKENHI Grant Number 23KJ1150.

†Full-time Lecturer, Department of Area Studies, Graduate School of Arts and Sciences, The University of Tokyo. E-mail:kazuhiroterashita@outlook.jp

Introduction

What political and social outcomes do social movements bring about? Social scientists have identified many of the forces that social movements possess, both empirically and theoretically (Amenta, Caren, and Olasky 2005; Holly J. McCammon et al. 2001; Holly J. McCammon et al. 2007, 2008). In particular, political science has focused more on the dynamics and consequences of social movements than on how they arise. Since the 2010s, a period often characterized as the 'age of mass protest' due to the rise of global protests, accurately measuring the impact of social movements has become an imperative for social scientists.

However, measuring the consequences of social movements is just as difficult, if not more so, than identifying when and where they occur. In particular, because social movements are the result of the simultaneous emergence and interaction of various political phenomena (Tarrow 1994; Tilly and Tarrow 2007), identifying causal relationships requires consideration of a variety of confounding variables. Furthermore, analogous to other complex political phenomena, the inherent difficulty in randomly generating social movements and observing their outcomes precludes the straightforward application of randomized controlled trials for causal identification.

Consequently, many studies have undertaken protest event analyses to identify observable instances of collective action, subsequently employing these event datasets for various empirical tests (Ortiz et al. 2005). As a result, it has been suggested that social movements may influence legislator behavior, voting behavior, voter turnout, and policy. However, there are many confounding factors between these causes and consequences (Amenta et al. 2010). Initially, it is observed that social movements are more prone to emerge during periods of heightened political opportunities, and individual mobilization tends to increase when the perceived likelihood of policy realization is greater (Kitschelt 1986). Given that this political opportunity structure encompasses a multitude of social, cultural, and political factors, constructing a comprehensive model that adequately accounts for all these elements presents a considerable methodological challenge.

Consequently, social scientists have endeavored to identify the causal relationship between protest behavior and its outcomes by employing causal identification strategies, notably instrumental variables. Specifically, precipitation has frequently been utilized, with researchers leveraging random variations in weather conditions to examine the effect of protest size on political outcomes (Wasow 2020; Madestam et al. 2013). The underlying assumption is that adverse weather, such as rain, discourages participation, thereby resulting in smaller demonstrations. For precipitation to serve as a valid instrumental variable, it must satisfy the exclusion restriction, meaning it should have no direct effect on the outcomes of interest, such as voter turnout or policy implementation, independent of its effect on protest size. Social scientists have, to varying degrees, sought to empirically justify this assumption.

However, as Mellon (2024) critically observed, while numerous political scientists employ precipitation as an instrumental variable, a substantial number of studies fail to satisfy the crucial exclusion restriction: that the instrumental variable influences the dependent variable solely through its effect on the independent variable, and not via alternative causal pathways. In essence, this necessitates careful consideration of alternative causal pathways through which precipitation might influence the outcomes. Mellon (2024) further reviewed several empirical studies examining protest outcomes, highlighting their failure to adhere to the aforementioned exclusion restriction. For instance, studies positing that the size of Tea Party protests increases Republican turnout often neglect to account for the possibility that precipitation might influence Republican turnout through alternative pathways (Madestam et al. 2013). In sum, the effective application of precipitation as an instrumental variable necessitates a thorough consideration of causal mechanisms beyond the hypothesized pathway and a judicious selection of covariates to mitigate confounding.

Given that prior research has predominantly focused on mobilization, what alternative instrumental variables can effectively explain other protest outcomes? Addressing this critical question, the present study seeks to identify potential instrumental variables by conducting a decades-long protest event analysis for South Korea and constructing a unique dataset. South Korea is particularly compelling as a case study due to its unique combination of an exceptionally high volume of protests and a relatively low participation rate compared to the average of developed countries. This distinctive context implies that social movement organizations in South Korea face a strategic imperative in deciding whether to organize protests. Specifically, given the inherent nature of protests requiring substantial mobilization, organizations may strategically opt to scale back or cancel events when adverse weather conditions are anticipated. Leveraging this unique civil society context in South Korea, this study will investigate whether climate variables (e.g., precipitation, temperature) are indeed linked to previously under-examined protest outcomes, such as their occurrence or absence, the extent of media coverage, and the resulting public attention, rather than solely focusing on protest size.

The structure of this paper is as follows. First, this study will present the operational variables and social movement factors under examination. Subsequently, it will detail the analytical methods and operationalization employed herein. Next, the diachronic changes and results of social movements in Korea will be presented. Finally, the implications of these findings will be discussed.

Predict Instrumental Variables

Research on social movement outcomes has shown that a variety of factors condition political and social outcomes. Classically, the size and violence of a social movement, as well as the tactics it employs, are known to have both positive and negative effects on outcomes (McAdam and Su 2002; Fassiotto and Soule 2017).

Media attention, defined as the extent to which a social movement is reported, is also an important factor(Madestam et al. 2013). Policymakers and voters typically become aware of movements not through direct observation of demonstrations, but via television, newspapers, social networking sites, and other media. Therefore, the degree to which an event attracts attention is itself an indicator that influences the outcome of a social movement (Hunt and Gruszczynski 2019; Andrews and Caren 2010).

Even when estimating social movement outcomes, we cannot overlook the endogenous nature of their complex dynamics. Therefore, several studies have employed instrumental variables to estimate the impact of specific social movement factors on turnout and other outcomes. Precipitation is one variable that has been extensively utilized. Similar to voting, participation in social movements involves costs. Some studies, drawing on classical models of political participation, focus on this cost and assume that rainfall impedes participation in social movements.

If we focus on the cost of participation, weather variables are certainly strong candidates for instrumental variables. Particularly in areas where spontaneous (off-poll) participation is popular, there may be a number of unorganized mobilizers who will not participate if it rains. However, in areas where organized mobilization is the norm, light precipitation may not be enough to dissuade participation. Similarly, in regions where rainfall is infrequent or particularly frequent, rain may not represent a relative cost. If the focus is on climate, then common temperatures across various locations would also warrant attention. Given these arguments, this study assesses the appropriateness of precipitation and temperature as instrumental variables based on their cost implications.

Methods

Detecting Protest Events

Social scientists have used protest event analysis to determine when, where, and how protests take place. Prominent projects for understanding political events—not solely protest events—include The World Handbook of Political Indicators, the World-wide Integrated Crisis Warning System (ICEWS), Phoenix Data Project, Global Database of Events, Language and Tone (GDELT), and Event Data on Armed Conflict and Security (EDACS). Some of these databases leverage machine learning techniques to compile extensive datasets from vast numbers of global newspaper articles, enabling broad comparative analyses.

However, prior research highlights significant challenges. These issues primarily stem from problems inherent in media coverage, particularly when using newspaper articles and other media reports. First, no single medium can capture a comprehensive sample of all real-world protests and political phenomena (Jenkins and Maher 2016). For instance, even public records, such as police-generated data, may not capture rallies that aren't subject to policing or illegal gatherings.

Jenkins and Maher (2016) contend that the absence of a robust sampling methodology lies at the root of this issue. Consequently, it's difficult to ascertain the specific biases inherent in any given source. Observations indicate that data derived from a single source (e.g., a national newspaper like the New York Times) can differ significantly from data obtained from multiple or alternative sources. For instance, Hocke (1999) found that local German newspapers covered approximately 38% of protests recorded in police records, yet only a small fraction of those reported in news outlets were not captured in police records. Similarly, Barranco and Wisler (1999) indicated that approximately 50% of protests identified in police reports were covered by Swiss newspapers. Conversely, Fillieule (1998) reported that only 2-3% of police-reported protests were also covered in the French national press. However, these studies primarily focus on protests subject to policing, such as street rallies and demonstrations. The extent of bias regarding more moderate forms of protest—including symposia, seminars, press conferences, and lawsuits—as well as elite-directed political activities, remains unknown.

Several other biases emerge when newspaper articles serve as data sources, including those stemming from the characteristics of the news media itself. For instance, conservative-leaning newspapers may underreport progressive events, and vice versa. Furthermore, the volume of reported protest events can vary between regions with national newspaper bureaus and those without (Rafail, McCarthy, and Sullivan 2019). Additionally, the inherent characteristics of the events being reported can introduce bias. Commonly cited is the topical salience of the protest's issue and the protest itself. For example, events like anti-Vietnam War protests and civil rights movements in the U.S. may receive disproportionate, and at times excessive, coverage.

Even if these issues are partially mitigated and a broad range of newspaper articles are extracted, biases can still arise during the coding process, pointing to a significant problem of coding consistency. In many previous studies, coding reliability is typically assessed by employing two or more coders, measuring their inter-coder agreement, and explicitly reporting this agreement. However, agreement rates are often reported to be around 60%, which is considered relatively low (Ortiz et al. 2005). The low inter-coder agreement stems from the inherent complexity of coding protest events. For example, essential data points (e.g., date, time, location, protest format, grievances, targets, organizers, and co-sponsoring organizations) are not always explicitly available. The more detailed the information to be recorded, the greater the burden on the coder. Additionally, even with pre-established coding criteria, ambiguities can arise during the interpretation of newspaper articles regarding whether an event constitutes a protest, its specific format, or its claims. In such instances, coders may be compelled to make subjective judgments unless a protocol, such as flagging the event for re-discussion, is strictly followed. Given this substantial burden, inter-coder agreement for protest event data remains low, raising significant concerns about its reliability.

Beyond issues of media coverage, a critical disadvantage of protest event analysis is its substantial cost. Traditional manual analysis of protest events from newspaper articles incurs substantial human, financial, and temporal costs associated with subscribing, reading, and coding. Consequently, protest event analysis has traditionally been limited to universities and research projects possessing ample funding and human resources. Furthermore, manual analysis inherently fails to resolve the issue of coder consistency, a significant aforementioned bias. Historically, datasets were coded by assistants with specific expertise, often graduate students in political science or sociology. However, as previously stated, even the strictest coding manuals cannot fully guarantee consistency. Moreover, these disadvantages impede the evaluation and reproducibility of the data. Without the necessary resources to generate the data initially, researchers are often compelled to rely on existing datasets. Even if the costs of data creation could be borne, the challenges of coder consistency still hinder data evaluation and reproducibility. In essence, manual data input fundamentally cannot resolve the consistency problem.

To address these multifaceted challenges, many researchers have actively developed semi-automated methods to fully or partially automate the process from newspaper article collection to dataset construction (Bond et al. 1997; Chojnacki et al. 2012; Lorenzini et al. 2022; Nardulli, Althaus, and Hayes 2015; Zhang and Pan 2019; Wang et al. 2016). These approaches encompass machine learning, particularly supervised learning models trained on analyst-prepared ground truth data, and the development and operation of programs that automatically code document sequences based on mechanically defined coding rules.

King and Lowe (2003) observed that the accuracy of machine coding using the Virtual Research Associates Knowledge Manager was comparable to human coding, though they also found that accuracy varied significantly depending on the type of protest event. For instance, coding accuracy for political graffiti ranged from 25-50%, while for protests, it reached 55-70% depending on the specific protest format. A further challenge when synthesizing information from multiple news sources is the necessity of processing duplicate events (Wang et al. 2016). Wang et al. (2016) found that among automated protest event analysis projects, ICEWS demonstrated strong performance in identifying protest events, though they noted that even with ICEWS, approximately 20% of events were duplicates.

However, concerns persist that human modification during classification may inadvertently introduce bias. Furthermore, the manual classification process, which was intended to reduce costs and enhance reproducibility (i.e., coder consistency) through automation, could paradoxically increase costs and undermine reproducibility. Indeed, Lorenzini et al. (2022) reported that 26.6% of machine-selected articles contained no events, and approximately one-third (32.6%) contained duplicate events previously reported in other articles, which were subsequently removed manually. Conversely, the semi-automated method proves effective in its iterative nature, requiring multiple checks and refinements through trial and error rather than a single, one-time analysis. This implies that improving accuracy hinges on iterative manual refinement, rather than relying solely on one-time machine learning applications.

As detailed above, the integration of methods like machine learning and natural language processing can resolve most of the aforementioned challenges associated with manual analysis. Specifically, automated or semi-automated methods offer the distinct advantage that once classification criteria are defined, data generation can be automated by a computer. Furthermore, the existence of well-defined classification criteria inherently ensures coding consistency. Another benefit is the enhanced reproducibility of the analysis, achieved by applying the identical method to the same newspaper article data. Provided that classification criteria and methods are publicly available, the analytical approach can be refined for greater accuracy and precision through the evaluation of trends and errors in the generated data. Indeed, some researchers, like Hanna (2017), opt to publicly release the analysis program rather than the dataset itself, thereby enhancing data availability and fostering further development.

However, the inherent limitations of relying on newspaper reports present challenges that are difficult to resolve purely through technical means. Given the finite capacity of newspaper pages and reporters, it's simply not feasible to observe and comprehensively document every protest event. Therefore, judicious selection of data sources and analytical methods is crucial, dependent on the specific aspects of protest under investigation.

Building upon these considerations, this paper will address the aforementioned challenges by leveraging a unique database comprising newspaper articles and police data. To further enhance coding accuracy, multiple machine learning models will be applied to classify the article data, and their respective results will be comparatively analyzed. This approach will not only demonstrate which learning models are most effective for conducting protest event analysis but also explore the potential variability of usable analytical variables contingent on the specific data and methodology employed.

The specific methods are as follows: First, I collected articles related to social movements from BIGKinds, a large-scale news archive in Korea. This archive, like most major newspaper databases, allows users to search for included newspaper articles by date, media, keywords, and publication page. Additionally, it offers the advantage of free cross-search of published newspapers and basic keyword analysis. However, BIGKinds only provides the first 200 words of an article for free, so the full text was not used in this analysis. Nevertheless, I believe there is some validity in analyzing and classifying protest event texts, as a certain text length is necessary, and much information about an event is often presented at the beginning of an article.

To accommodate various media coverage, previous studies have recommended the use of multiple media sources. In this paper, all national and regional newspapers included in BIGKinds were analyzed. The period covered was from January 1, 2000, to December 31, 2022. Duplicate articles, editorials, and international news were excluded beforehand. As a result, 205,931 newspaper articles were collected.

The text of the collected articles was then extracted and segmented into separate sentences. The 'gibasa' package and XLMRoberta's Tokenizer were used for this segmentation. After removing unnecessary words (stopwords) such as particles, alphabetic characters, and numbers, a document-feature matrix was created. For stopwords, a list of Korean words from 'marimo', an R package developed for multilingual text analysis, was used.

In this paper, the processed texts were classified to determine whether they are 'related to social movement events' or not. Three different supervised learning models were employed for this determination, with each model attempting to address the imprecision inherent in machine learning for identifying protest events. Specifically, Naive Bayes, Random Forest, and XLMRoberta (a pre-trained language model) were utilized (Conneau et al. 2020). To use these models, one native Korean speaker was asked to classify 1,000 pre-selected, randomly chosen datasets to generate training data.

Using the created training data, Naive Bayes and Random Forests were simply trained, while XLMRoberta was fine-tuned to obtain appropriate learning results. The accuracy of each is shown in Table 1. These accuracies are generally similar to those reported by previous studies. It is worth noting that there is actually no significant difference in simple classification accuracy between a simpler model like Naive Bayes and a complex and costly model such as XLMRoberta. This may be due to the limited amount of data used for fine-tuning, but it could also indicate the inherent difficulty of identifying protest behavior.

Table 1: Classification Performance of Protest Detection Models

Model	Accuracy	Recall	F1 Score
Naive Bayes	0.72	0.69	0.65
Random Forest	0.70	0.71	0.76
XLM-RoBERTa	0.72	0.70	0.70

Although both models exhibit low accuracy in binary classification tasks when supervised learning is applied inherently, this paper expects to obtain consistent results regardless of the model used, while acknowledging that a certain degree of bias exists in each. For this reason, in future analyses, we will prioritize the use of multiple results over individual model accuracy.

I applied these models and utilized newsmap, a quasi-supervised learning model, to obtain geographic information for the classified data (Watanabe 2018). This model employs an arbitrary geographic dictionary to estimate the geographical location of the articles. The model can also estimate geography in finer units by allowing the analyst to create a new dictionary instead of a pre-prepared one. I created a dictionary for the protest event analysis by setting up seed words related to basic local governments in Korea. Since this paper omits international news beforehand, it is generally possible to classify only domestic news reports.

Finally, to identify the repertoire and themes of the protests, we utilized SeededLDA, a quasi-supervised learning model, to estimate the topics of the articles (Watanabe and Baturo 2023). This involved two distinct thematic categories: one related to repertoire and the other to the themes of protest actions. Original seed terms were set for each category, allowing for differentiated classifications. In the analysis, geographical words were removed beforehand, as they pose obstacles to classification.

Identifying Instrumental Variables

This paper examines whether certain instrumental variables are indeed appropriate for estimating the impact of various factors on multiple social movements. For each test, we'll combine variables and cases to determine if these instrumental variables function effectively under the most favorable data conditions.

This paper focuses on aspects of social movements that have received less attention in previous instrumental variable studies: their occurrence, the amount of media coverage, and their intensity. The generation of social movements, or event count, is obtained by estimating each machine learning model using the aforementioned methods. The amount of media coverage is a variable obtained by aggregating events that would normally be considered the same (i.e., sharing the same date, time, location, issue, and repertoire), given the use of multiple newspaper articles. Additionally, the intensity of social movements is measured by the sum of conflict-ridden social movements estimated by the topic model. This includes, for example, clashes with police or other social movements, and illegal occupations.

The next consideration is the choice of instrumental variable. Consistent with previous research, we will utilize a random exogenous variable: weather. Beyond precipitation, this study will specifically focus on temperature, a variable not extensively examined in prior instrumental variable research. Precipitation and temperature data were obtained from the Korea Meteorological Administration (KMA). As these data are provided in station units, they were processed using the Inverse Distance Weighting (IDW) method, which weights and estimates values from stations closer to the center of the municipality. A simple, inverse relationship between precipitation and participation was assumed: increased rainfall leads to decreased participation. However, a linear relationship is difficult to assume for temperature, as both excessively hot and cold conditions are presumed to deter participation. Therefore, the primary temperature variable employed in this paper is the absolute deviation from 20 degrees Celsius. Additionally, a nonlinear model incorporating a squared temperature term will be used as a preliminary measure. More details on the data are available in the Supporting Information.

Results

Protest Events

First, we will assess how the protest events estimated in this paper align with actual data. Although no comprehensive documentation of all protest events exists, police records serve as a useful official source. Figure 1 compares the police data with the estimation results of each model, demonstrating that the trends are generally consistent. Police data not only includes very small, unreported social movements but may also contain duplicate counts where multiple applications are accepted for what is effectively a single rally. In any case, the results are similar across models, reflecting, for example, the effects of legal restrictions during the

COVID-19 pandemic (2020) and the increase in impeachment politics (2016).

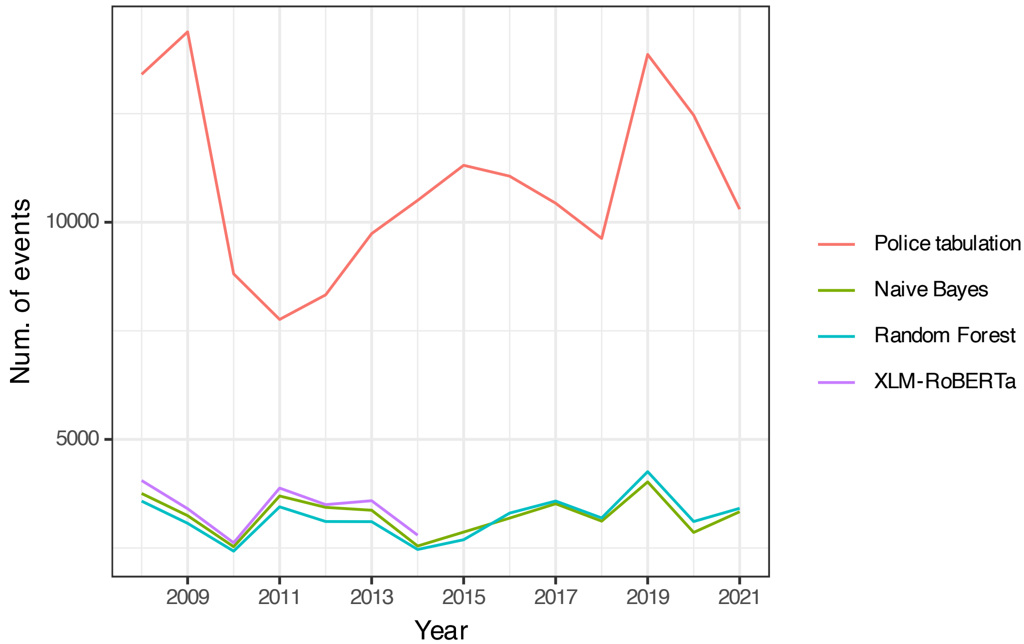

Figure 1: Comparison Between Protest Data and Police Records (ALL)

Figure 2 supplements the estimation characteristics in more detail using Seoul-specific data. Unlike the national data discussed earlier, discrepancies exist between police data and newspaper article data in certain areas. Notably, while police data show fewer values in 2011, other estimates are higher. The implications of this trend difference will require future examination.

We should also examine the geographic distribution. Figure 3 displays a map of the aggregated event results for each regional municipality, where darker colors indicate a higher incidence of events. This visualization demonstrates that events primarily occur in metropolitan and urban areas. Given that police data also confirm several thousand events in Seoul alone, we believe this estimation is appropriate.

Instrumental Variables

Using the estimated protest event data, we will examine the effect of random temperature variations on the occurrence, number of reports, and militancy of social movements. Table 2

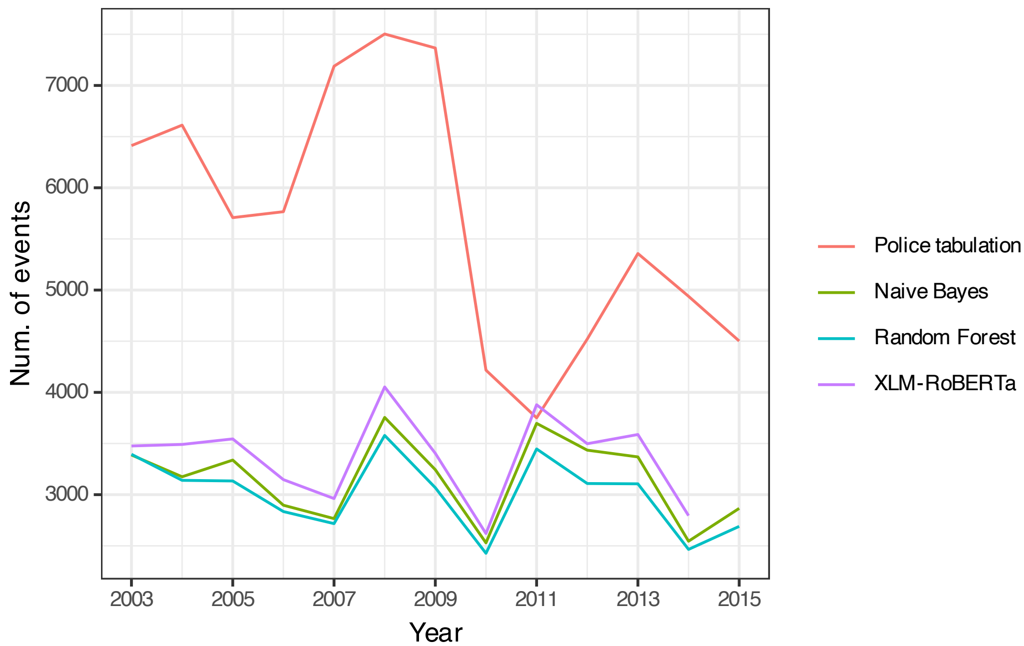

Figure 2: Comparison Between Protest Data and Police Records (Seoul)

presents the results of regression analyses examining the impact of precipitation and temperature differences (calculated as deviation from 20 degrees Celsius, where a larger absolute value indicates more unpleasant conditions) on the data obtained from the three estimation models and the three outcomes. As is clear from the table, none of the models or variables show statistical significance. These results indicate that climate variables, traditionally used as instrumental variables for mobilization, cannot be employed for the causal identification of at least event occurrence, media coverage, and event militancy.

Other analyses yield similar findings. Table 3 presents similar estimates derived from aggregate data collected one week prior to the 2018 election. Unlike earlier findings, some coefficients are positive or negative, but they lack statistical significance or indicate substantial effects. Furthermore, precipitation, which typically has a negative effect on mobilization numbers, exhibits positive values for all variables. This suggests that instrumental variables should be reconsidered when focusing on social movement factors beyond mere mobilization numbers.

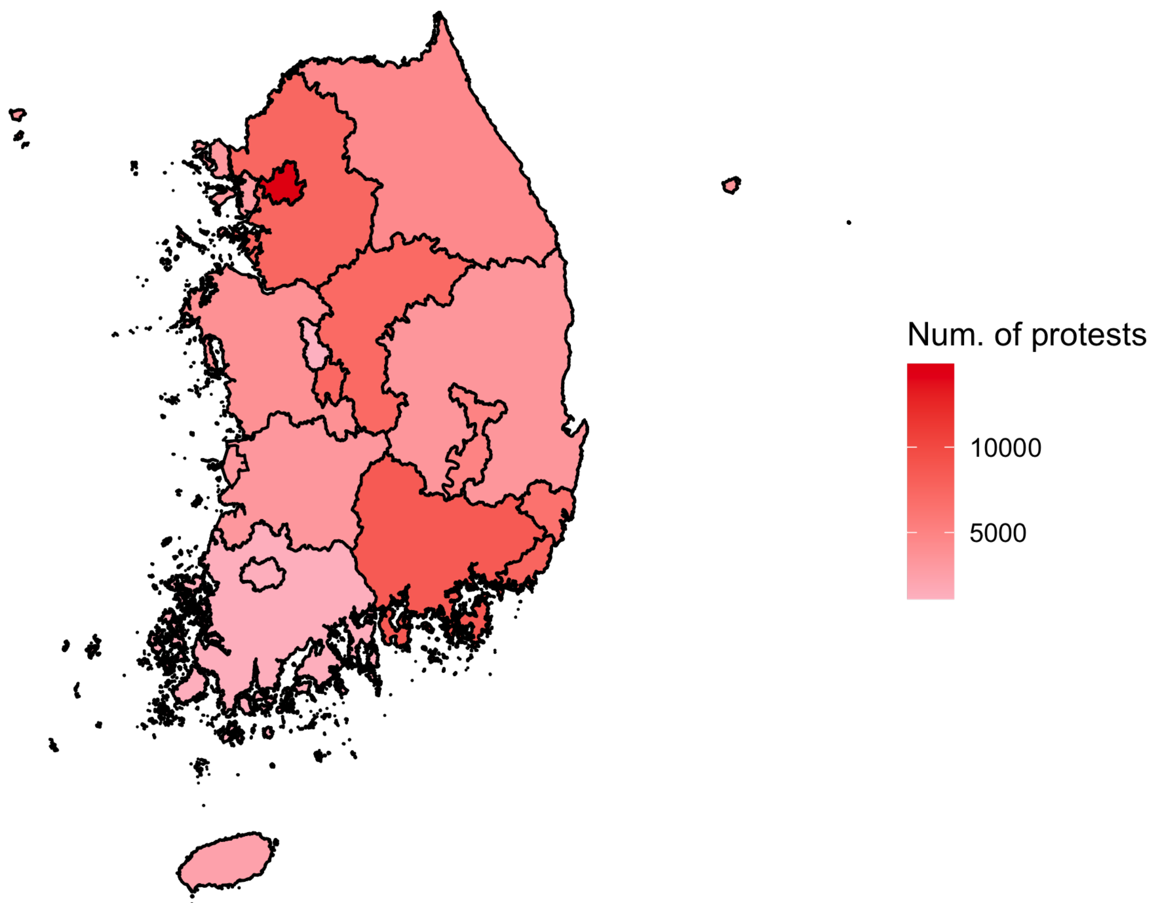

Figure 3: Geographic Distribution of Protest Events

Table 2: Estimation Results of Protest Activity by Day

	NB-Protest	RF-Protest	XLMR-Protest	NB-Report	RF-Report	XLMR-Report	NB-Clash	RF-Clash	XLMR-Clash
Rainfall	0.000	0.000	0.000	0.000	0.000	0.000	0.000+	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Temp_dev	0.000	0.000	0.000	0.000	0.000	0.000+	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Num.Obs.	294 417	294 417	294 417	294 417	294 417	294 417	294 417	294 417	294 417
R2	0.223	0.270	0.090	0.176	0.226	0.052	0.034	0.031	0.013

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors clustered at the metropolitan (province) level.

NB = Naive Bayes, RF = Random Forest, XLMR = XLM-RoBERTa. Outcomes are protest counts, news coverage, and protest with clashes.

Table 3: Effects of Predicted Protest Activity on Pre-Election Averages

	NB-Protest	RF-Protest	NB-Report	RF-Report	NB-Attack	RF-Attack
(Intercept)	0.746	−0.672	0.834	0.834	0.100	0.100
	(1.237)	(1.653)	(1.285)	(1.285)	(0.102)	(0.102)
Rainfall	0.056	0.029	0.055	0.055	0.007	0.007
	(0.056)	(0.054)	(0.065)	(0.065)	(0.007)	(0.007)
Temp_dev	−0.037	0.032	−0.042	−0.042	−0.005	−0.005
	(0.059)	(0.080)	(0.062)	(0.062)	(0.005)	(0.005)
Temp	−0.024	0.042	−0.027	−0.027	−0.005	−0.005
	(0.057)	(0.078)	(0.060)	(0.060)	(0.005)	(0.005)
Num.Obs.	160	160	160	160	160	160
R2	0.008	0.007	0.006	0.006	0.016	0.016

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

NB = Naive Bayes, RF = Random Forest.

Outcomes are protest counts, news coverage, and protest with clashes.

Finally, to further confirm these null findings, I tested a model incorporating a day-of-the-week dummy variable (reference category: Monday) and multiple climate variable inputs. The results, presented in Table 4, indicate that, with the exception of a few variables, there remains no statistically significant effect. The coefficients of the few significant variables are small, suggesting virtually no substantive effect.

Conclusion

Methodological rigor has long been a challenge in the social and political sciences, and empirical research on social movements still struggles to find effective methods for causal identification. Instrumental variables offer a potential solution, but it's been unclear whether random variables, beyond just mobilization numbers and climate, can be effectively utilized.

Table 4: Testing the Null Association Between Weather Conditions and Protest Activity

	NB-Protest	RF-Protest	XLMR-Protest	NB-Report	RF-Report	XLMR-Report	NB-Attack	RF-Attack	XLMR-Attack
Rainfall	0.000	0.000	0.000	0.000	0.000	0.000	0.000*	0.000*	0.000+
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Temp_dev	0.000	0.000	−0.001**	0.000	0.000	−0.001**	0.000*	0.000+	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Temp	0.000	0.000	−0.001***	0.000	0.000	−0.001***	0.000*	0.000+	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
I(Temp^2)	0.000	0.000	0.000*	0.000	0.000	0.000*	0.000	0.000	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
WeekdayTuesday	0.001	0.001	0.000	0.004	0.003	0.000	0.000	0.000	0.000
	(0.002)	(0.002)	(0.001)	(0.004)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)
WeekdayWednesday	0.000	0.000	−0.001	0.003	0.002	0.000	0.000	0.000	0.000
	(0.002)	(0.002)	(0.001)	(0.004)	(0.003)	(0.001)	(0.001)	(0.000)	(0.000)
WeekdayThursday	−0.019***	−0.016**	−0.004**	−0.021***	−0.016**	−0.004**	−0.001*	−0.001*	0.000
	(0.006)	(0.005)	(0.001)	(0.006)	(0.005)	(0.001)	(0.001)	(0.000)	(0.000)
WeekdayFriday	−0.056***	−0.052***	−0.013**	−0.062***	−0.055***	−0.017**	−0.003***	−0.003***	0.000*
	(0.015)	(0.014)	(0.004)	(0.016)	(0.014)	(0.005)	(0.001)	(0.001)	(0.000)
WeekdaySaturday	−0.048***	−0.043***	−0.013**	−0.057***	−0.050***	−0.016**	−0.003***	−0.002***	0.000**
	(0.013)	(0.012)	(0.004)	(0.015)	(0.014)	(0.005)	(0.001)	(0.001)	(0.000)
WeekdaySunday	−0.006*	−0.003	−0.003*	−0.005+	0.000	−0.003+	0.000	0.000	0.000
	(0.002)	(0.002)	(0.001)	(0.003)	(0.004)	(0.002)	(0.000)	(0.000)	(0.000)
Num.Obs.	294 417	294 417	294 417	294 417	294 417	294 417	294 417	294 417	294 417
R2	0.216	0.262	0.060	0.168	0.218	0.053	0.026	0.024	0.005

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors clustered at the metropolitan (province) level.

NB = Naive Bayes, RF = Random Forest, XLMR = XLM-RoBERTa. Outcomes are protest counts, news coverage, and protest with clashes.

The analysis in this paper demonstrates that random variables such as precipitation and temperature are unlikely to significantly affect event occurrence, media coverage, or the militancy of protest formats, at least in South Korea. Several additional analyses corroborate this finding. We also employed various machine learning methods for analyzing protest events, consistently showing that no estimation method can detect a climate influence. This paper, as further discussed, offers specific contributions to understanding protest events.

What does it signify that climate variables cannot serve as effective instrumental variables? First, we must consider the stagnant nature of civic participation in a "normal civil society" with a mature democracy. South Korea, for instance, recently experienced a state of emergency followed by presidential impeachment and removal. However, prior to this, professional activists, often with experience in the democracy movement, typically led advocacy efforts without widespread citizen street mobilization. This structure of civil society is underscored by the low level of non-conventional political participation among Korean citizens, which ranks low among OECD countries (Inglehart et al. 2014).

Given these confirmed facts, the occurrence, media coverage, and repertoire of events are likely driven by the strategies of organized activists and leaders. Such seasoned professionals are unlikely to cancel events or alter strategies due to minor climate shifts. While previous studies have identified the potential for instrumental variables in the United States—a social movement environment that is, in some sense, unique—it becomes crucial to explore what causal identification strategies are viable in developed countries, which are largely grappling with the universal phenomenon of declining political participation in mature societies.

Additionally, this study contributes to the development of protest event analysis. While many research projects aim to automate or semi-automate the analysis, often at great manpower and expense, the core problems with these analyses are their cost and inflexibility. This study demonstrates a certain degree of accuracy by utilizing a combination of multiple machine learning models and text classification methods, thereby offering a solution to both of these problems. Future work to confirm and improve the accuracy of the results obtained with these methods will foster the development of analysis methods applicable even to researchers with limited resources. Furthermore, this analysis method is reproducible and flexible, provided that the data and dictionaries used for training are made publicly available. For example, it's possible to focus the analysis on specific issue repertoires, such as environment or gender, and to apply it in multiple languages. While many large-scale projects have targeted large-scale protest events—a scope not always aligned with the interests of many researchers who require more flexible coding—it should be noted that researchers' interests are varied. This customizability will also need to be improved while scrutinizing accuracy.

References

Amenta, Edwin, Neal Caren, Elizabeth Chiarello, and Yang Su. 2010. "The political consequences of social movements." *Annual Review of Sociology* 36: 287–307. https://doi.org/10.1146/annurev-soc-070308-120029.

Amenta, Edwin, Neal Caren, and Sheera Joy Olasky. 2005. "Age for leisure? Political mediation and the impact of the pension movement on U.S. old-age policy." *American Sociological Review* 70 (3): 516–38. https://doi.org/10.1177/000312240507000308.

Andrews, Kenneth T., and Neal Caren. 2010. "Making the news: Movement organizations, media attention, and the public agenda." *American Sociological Review* 75 (6): 841–66. https://doi.org/10.1177/0003122410386689.

Barranco, José, and Dominique Wisler. 1999. "Validity and systematicity of newspaper data in event analysis." *European Sociological Review* 15 (3): 301–22. https://doi.org/10.1093/oxfordjournals.esr.a018265.

Bond, Doug, J. Craig Jenkins, Charles L. Taylor, and Kurt Schock. 1997. "Mapping mass political conflict and civil society: Issues and prospects for the automated development of event data." *Journal of Conflict Resolution* 41 (4): 553–79. https://doi.org/10.1177/0022002797041004004.

Chojnacki, Sven, Christian Ickler, Michael Spies, and John Wiesel. 2012. "Event Data on Armed Conflict and Security: New Perspectives, Old Challenges, and Some Solutions." *International Interactions* 38 (4): 382–401. https://doi.org/10.1080/03050629.

2012.696981.

Conneau, Alexis, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. "Unsupervised Cross-Lingual Representation Learning at Scale." https://arxiv.org/abs/1911.02116.

Fassiotto, Magali, and Sarah A. Soule. 2017. "Loud and clear: The effect of protest signals on congressional attention." *Mobilization* 22 (1): 17–38. https://doi.org/10.17813/1086-671X-22-1-17.

Fillieule, Olivier. 1998. "'Plus ca change, moins ca change': Demonstrations in France during the Nineteen-Eighties." In *Acts of Dissent : New Developments in the Study of Protest*, edited by Dieter Rucht, Ruud Koopmans, and Friedhelm Neidhardt, 199–226. Sigma. https://cir.nii.ac.jp/crid/1130282272232447232.

Hanna, Alex. 2017. "MPEDS: Automating the Generation of Protest Event Data." Preprint. SocArXiv. https://doi.org/10.31235/osf.io/xuqmv.

Hocke, Peter. 1999. "Determining the Selection Bias in Local and National Newspaper Reports on Protest Events." *Acts of Dissent: New Developments in the Study of Protest*, no. October: 131–63.

Hunt, Kate, and Mike Gruszczynski. 2019. "The influence of new and traditional media coverage on public attention to social movements: the case of the Dakota Access Pipeline protests." *Information, Communication & Society*, September, 1–17. https://doi.org/10.1080/1369118X.2019.1670228.

Inglehart, Ronald, Christian Haerpfer, Alejandro Moreno, Christian Welzel, Kseniya Kizilova, Juan Diez-Medrano, Marta Lagos, Pippa Norris, Eduard Ponarin, and Bjorn Puranen, eds. 2014. "World Values Survey: Round Six - Country-Pooled Datafile Version." https://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp.

Jenkins, J. Craig, and Thomas V. Maher. 2016. "What Should We Do about Source Selection in Event Data? Challenges, Progress, and Possible Solutions." *International Journal of Sociology* 46 (1): 42–57. https://doi.org/10.1080/00207659.2016.1130419.

King, Gary, and Will Lowe. 2003. "An Automated Information Extraction Tool for International Conflict Data with Performance as Good as Human Coders: A Rare Events Evaluation Design." *International Organization* 57 (3): 617–42. https://doi.org/10.1017/s0020818303573064.

Kitschelt, Herbert P. 1986. "Political Opportunity Structures and Political Protest: Anti-Nuclear Movements in Four Democracies." *British Journal of Political Science* 16 (1): 57–85. https://doi.org/10.1017/S000712340000380X.

Lorenzini, Jasmine, Hanspeter Kriesi, Peter Makarov, and Bruno Wüest. 2022. "Protest Event Analysis: Developing a Semiautomated NLP Approach." *American Behavioral Scientist* 66 (5): 555–77. https://doi.org/10.1177/00027642211021650.

Madestam, Andreas, Daniel Shoag, Stan Veuger, and David Yanagizawa-Drott. 2013. "Do political protests matter? Evidence from the tea party movement." *Quarterly Journal of Economics* 128 (4): 1633–85. https://doi.org/10.1093/qje/qjt021.

McAdam, Doug, and Yang Su. 2002. "The war at home: Antiwar protests and congressional voting, 1965 to 1973." *American Sociological Review* 67 (5): 696–721. https://doi.org/10.2307/3088914.

McCammon, Holly J, Karen E Campbell, Ellen M Granberg, and Christine Mowery. 2001. "How Movements Win: Gendered Opportunity Structures and U.S. Women's Suffrage Movements, 1866 to 1919." *American Sociological Review* 66 (1): 49–70. https://doi.org/10.2307/2657393.

McCammon, Holly J., Soma Chaudhuri, Lyndi Hewitt, Courtney Sanders Muse, Harmony D. Newman, Carrie Lee Smith, and Teresa M. Terrell. 2008. "Becoming full citizens: The U.S. women's jury rights campaigns, the pace of reform, and strategic adaptation." *American Journal of Sociology* 113 (4): 1104–47. https://doi.org/10.1086/522805.

McCammon, Holly J., Courtney Sanders Muse, Harmony D. Newman, and Teresa M. Terrell. 2007. "Movement framing and discursive opportunity structures: The political successes of the U.S. women's jury movements." *American Sociological Review* 72 (5): 725–49. https://doi.org/10.1177/000312240707200504.

Mellon, Jonathan. 2024. "Rain, rain, go away: 194 potential exclusion-restriction violations for studies using weather as an instrumental variable." *American Journal of Political*

Science, no. July 2022: 1–18. https://doi.org/10.1111/ajps.12894.

Nardulli, Peter F., Scott L. Althaus, and Matthew Hayes. 2015. "A Progressive Supervised-learning Approach to Generating Rich Civil Strife Data." *Sociological Methodology* 45 (1): 148–83. https://doi.org/10.1177/0081175015581378.

Ortiz, David, Daniel Myers, Eugene Walls, and Maria-Elena Diaz. 2005. "Where Do We Stand with Newspaper Data?" *Mobilization: An International Quarterly* 10 (3): 397–419. https://doi.org/10.17813/maiq.10.3.8360r760k3277t42.

Rafail, Patrick, John D. McCarthy, and Samuel Sullivan. 2019. "Local receptivity climates and the dynamics of media attention to protest." *Mobilization* 24 (1): 1–18. https://doi.org/10.17813/1086-671X-24-1-1.

Tarrow, Sidney. 1994. *Power in Movement: Social Movement, Collective Action and Politics*. Cambridge University Press. http://ci.nii.ac.jp/ncid/BA23651908.

Tilly, Charles, and Sidney G Tarrow. 2007. *Contentious Politics*. 2nd ed. fu. Oxford University Press. http://ci.nii.ac.jp/ncid/BB19452290.

Wang, Wei, Ryan Kennedy, David Lazer, and Naren Ramakrishnan. 2016. "Growing pains for global monitoring of societal events: Automated event coding raises promise and concerns." *Science* 353 (6307): 1502–4.

Wasow, Omar. 2020. "Agenda Seeding: How 1960s Black Protests Moved Elites, Public Opinion and Voting." *American Political Science Review* 114 (3): 638–59. https://doi.org/10.1017/S000305542000009X.

Watanabe, Kohei. 2018. "Newsmap: A semi-supervised approach to geographical news classification." *Digital Journalism* 6 (3): 294–309. https://doi.org/10.1080/21670811.2017.1293487.

Watanabe, Kohei, and Alexander Baturo. 2023. "Seeded Sequential LDA: A Semi-Supervised Algorithm for Topic-Specific Analysis of Sentences." *Social Science Computer Review* 0 (0): 1–25. https://doi.org/10.1177/08944393231178605.

Zhang, Han, and Jennifer Pan. 2019. *CASM: A Deep-Learning Approach for Identifying Collective Action Events with Text and Image Data from Social Media*. Vol. 49. 1. https://doi.org/10.1177/0081175019860244.

Support Information

https://kazuhiroterashita.com/supplementary/