The geography of innovation and technology news – An empirical study of the German news media

Burcu Ozgun & Tom Broekel

Papers in Evolutionary Economic Geography

21.10

Utrecht University   Human Geography and Planning

The geography of innovation and technology news - An empirical study of the German news media

Burcu Ozgun[1] and Tom Broekel[2]

[1]Department of Human Geography and Spatial Planning, Utrecht University, Utrecht, Netherlands
[2]Business School, Stavanger University, Stavanger, Norway

Abstract

Variations in the frequency and tone of news media are the focus of a growing literature. However, to date, empirical investigations have primarily confirmed the existence of such differences at the country level. This paper extends those insights to the subnational level. We provide theoretical arguments and empirical support for systematic regional variations in the frequency and sentiments of news related to innovation and new technologies. These variations reflect regional socio-economic structures. We find that the average newspaper circulating in urban areas features more news on innovation and new technologies than media in more rural areas. Similar findings hold for locations in East Germany and to a certain degree for regions with low unemployment. The sentiments of innovation and new technology news are negatively associated to the unemployment rate, and they tend to be lower in regional newspapers than in national ones. Overall, our results suggest a strong link between the regional socioeconomic conditions and how newspapers circulating in these places report on innovation and new technologies.

Keywords— innovation, technology, news media, sentiment analysis, topic modeling.
JEL classification— O33, R12, L82.

1 Introduction

Innovation is undoubtedly a crucial ingredient of technological and economic growth. However, innovation and technological progress may also induce negative social and economic effects (González-Romá and Hernández, 2016). The discussion of artificial intelligence (AI) nicely illustrates this two-sided nature of innovation (Agrawal et al., 2019; Korinek and Stiglitz, 2017). The sheer unlimited potential of ever-growing mountains of data coupled with breathtaking advances in analytical systems, bear huge potentials for future economic growth and, thus attract the fascination of researchers and businesspeople alike (Aghion et al., 2017; Goldfarb and Trefler, 2018). On the other hand, AI is closely linked to the automation of human tasks and is widely expected to transform and replace the "routine, non-cognitive tasks that have been primarily performed by middle-skilled

workers" (Buarque et al., 2019). Consequently, many occupations are in danger of being automated or replaced by AI-based systems (Frey and Osborne, 2017; Acemoglu and Restrepo, 2019, 2020), spurring growing public concerns (Furman and Seamans, 2019; Inhoffen, 2018; Fast and Horvitz, 2017).

How the public thinks about new technologies matters; since it triggers and shapes political debates that eventually translate into concrete policies. In turn, these have the potential to create institutional structures that significantly shape the development, diffusion, application, and ultimately, the socioeconomic benefits that new technologies may unfold (Stone et al., 2016). The scientific literature has long recognized that in this context, the news media play a crucial role, with their substantial influence on public opinion and expectations regarding innovations and new technologies (e.g., Marks et al., 2007; Priest, 1994). This alignment of public opinion and news coverage is not one-directional. Studies show that expectations and citizens' general attitude toward such issues affect their presentation in news media (Watt Jr and Van Den Berg, 1978; Gentzkow and Shapiro, 2010). Thus, individuals' perceptions of and attitudes toward innovation are also likely to influence coverage of these issues in the news. Notably, existing studies identify substantial variations between countries in the intensity and sentiments of media coverage of these topics, raising the question of whether such variations occur only at the country level.

In this paper, we argue that while the national level is of unquestioned importance, the subnational (regional) level has been vastly overlooked in this context. Innovation processes are highly localized, with regions' institutional, political, and economic contexts frequently determining the emergence and diffusion of technologies (Cooke et al., 1997; Bednarz and Broekel, 2020). At the same time, news coverage is not uniform across localities, implying that people in different regions are exposed to heterogeneous sets of information and varying evaluations regarding innovations. Consequently, news media may be an important, albeit widely overlooked social institution, that both explains and is explained by heterogeneity in regional attitudes toward innovation and technologies. This paper marks a first step in this direction. We discuss and empirically explore the extent to which regionally circulating news media cover innovation and new technologies differently, in terms of frequency and sentiments. In addition, we explore whether these differences are related to regional socioeconomic conditions.

The empirical study utilizes a newly established source of news information, the *RegNeS* database, which covers the most important national and regional newspapers in Germany. We used a range of text-mining methods to identify and evaluate news on innovation and new technologies in terms of sentiments, and we model their geographic distribution with the help of newspaper circulation data.

Our empirical findings at the level of German spatial-planning regions suggest that the average newspaper in urban areas is more likely to report on innovation and new technologies than those in other places. Our results also suggest the existence of an East-West divide: News on new technologies and innovation more frequently appears in newspapers circulating in regions in the former East Germany. A strong negative association exists between sentiments on innovation and new technology news and the unemployment rate in the newspapers' circulation area. Crucially, this result holds, even when controlling for the general sentiments of articles in the newspaper. In sum, our results

2

point toward significant regional variations in the frequency and sentiments to which newspapers expose readers, regarding news on innovation and new technologies. Consequently, looking at such news at a subnational level seems fruitful for gaining a better understanding of the spatial diffusion of new technologies and may represent a distinguishing factor of regional technological systems that thus far has received little attention.

The rest of the paper is organized as follows. Section 2 presents the theoretical discussion of the relationship between geography and news on innovation and technologies, particularly focusing on the regional level. Section 3 describes the data and our empirical approach. Section 4 presents the estimation results, and Section 5 concludes by discussing the implications of the findings.

2   Motivation and theoretical background

2.1   News and public expectations

A central element underlying the relation of news and innovation is expectations. More precisely, expectations are fundamental factors shaping the development, diffusion, and use of new technologies (Konrad, 2006; Borup et al., 2006; Geels and Verhees, 2011; Budde et al., 2012). Expectations usually consist of socio-technological visions describing a future world based on assumptions and empirical observations that translate into a set of scenarios of hopes and fears (Konrad, 2006). Individuals have expectations about the future development and use of specific technologies, about their structures, rules, and regulatory regimes. Expectations regarding innovation and new technologies form and materialize at different levels: micro, meso, and macro. They crucially shape search activities, the selection of technologies and their legitimization (Van Lente, 1995).

The role of expectations in technological development is directly linked to the concepts of risk and uncertainty (Berkhout, 2006). First, in an uncertain environment, expectations create coordination mechanisms for economic actors and activities, which can achieve alignment of interests (Alkemade and Suurs, 2012; Eames et al., 2006). Second, many instances of newly emerging technologies do not immediately meet existing markets and commonly lack structural components, such as regulations, infrastructure, user practices, and maintenance networks (Geels, 2002). Due to high degrees of uncertainty in these processes, technologies must attain legitimacy before the creation of such structural components (Bergek et al., 2008; Geels and Verhees, 2011)[1]. New technologies need cognitive legitimacy to attain an institutionalized diffusion of knowledge (Aldrich and Fiol, 1994). Societal embedding is a way of gaining this legitimacy, depending in turn, on societal norms and beliefs (Deuten et al., 1997). Once positive societal expectations have created legitimacy, required components emerge and diffuse (Alkemade and Suurs, 2012; McCormick, 2010), further stimulating the provision of resources, support, and investment in new technologies (Eames et al., 2006; Geels and Verhees, 2011; Borup et al., 2006). Moreover, collective expectations give economic agents a sense of *the way things are going* and, consequently, provide a guide for future research activities (Eames et al., 2006). In sum, collective expectations are part of the social repertoire, integral to

---

[1]Bergek et al. (2008) defines legitimization as *the politics of shaping expectations and of defining desirability.*

the socio-technological landscape (Konrad, 2006). Consequently, empirical research confirms that positive expectations are a crucial precondition for successful diffusion of innovation (Geels and Verhees, 2011; Budde et al., 2012).

Although positive expectations held by the wider society may create momentum for policy support and private investments into R&D efforts, sometimes they may harm the very process that they are promoting. A well-known example is the so-called "hype cycles". A hype cycle is a sudden increase in the attention and visibility that a technology gets (Van Lente et al., 2013). While a technological hype helps to generate initial interest and promote a technology, it sets high unconscious expectations, hard for most technologies to meet in the short run. If the promises are not fulfilled in the short run, public opinion and attention may quickly turn away, without giving the technology a proper chance for its benefits to materialize (Caulfield, 2004). Hence, hype cycles can lead to overly quick disappointments and consequential withdrawal of support (Bakker, 2010). However, in some cases, the institutionalization processes that such hypes trigger may continue after the hype has ended and keep promoting further diffusion and development of the technology (Ruef and Markard, 2006). Accordingly, the relationship between innovation activities and public expectations is neither straightforward, linear, nor one-sided. On the contrary, it is rather complex and not yet fully understood.

Given the increasing deterioration of the boundaries between science, technology, and society (Gibbons, 1999), what societies expect from emerging technologies is growing in importance, and the expectation becomes increasingly crucial for their future development and application. This raises the question of how such expectations form and what their determinants are.

A range of factors shape public opinion and expectations, with media being among the most crucial. In particular, framing the societal discourse (Konrad, 2006) is essential. As a crucial source of information, visuals, and interpretations of events external to peoples' direct observation (Lippmann, 1922), media have the power to influence the salience of attitudes, and, to certain extents, agenda-setting (McCombs and Shaw, 1972; Hester and Gibson, 2003). In many instances, how the media frame an issue shapes how people understand and remember it. The media also contribute to evaluation of and reaction to those issues (Entman, 1993). By helping people construct meanings and giving them orientation, the influence of media on public opinion is particularly substantial for such complex topics as science and technological development. Generally, people turn to the media to make sense of complexities with which they have little direct experience, for information on which they can rely (Boykoff, 2009; Mast et al., 2005; Zucker, 1978). Consequently, media are an essential element in disseminating science, technology and innovation-related information, opinions, and expectations to the wider public.

Ample evidence supports media coverage shaping public opinion and expectations regarding various technologies. For example, Gamson and Modigliani (1989) show that media discourse has been an essential context for the formation of public opinion on nuclear power since 1945. Another example is the biotechnology debate. In the late 1990s, public opinion on this technology drastically changed, as the public became highly concerned after the extensive media coverage of the cloning of the sheep, Dolly (Petersen, 2001).

4

In light of this evidence, variations in media presence, coverage, focus, and tone are likely reasons for the spatial variation. Gaskell et al. (1999) and Mazur (2006) highlight that the quantity of news media coverage explains the risk attitude towards and consumption of biotechnology and genetically modified food in different countries, as increased news coverage conveys a sense of hazard and uncertainty. It may also lead to greater awareness of the alleged risks in society. Thus, the influence of media and its spatial variation is important. For instance, Skjølsvold (2012) investigates how the news media of Sweden and Norway have covered and communicated about bioenergy. While news media in Norway focus on technological and economic ambivalence, news media in Sweden promote optimism and highlight green consumption features. This contributed to the development of different systems of innovation and diffusion patterns in the two countries, with respect to these two technologies. Similarly, Negro et al. (2012) study the presentation of wind-power technology in the media. They conclude that there was a lack of legitimacy of this technology in Sweden, which, apparently strongly shaped by the negative presentation of the wind power technology in the media.

The relationship between news coverage and spatial variation in perceptions regarding new technologies is not one-directional. As the insights in the literature on mass-media effects suggest, just as media may facilitate a specific agenda, they do not do so independent of their audience. In general, news media must provide for their audiences by reflecting their preferences and corresponding to existing views and interests. In this sense, audiences strongly shape the media's agenda (i.e., reverse agenda-setting). Research well establishing that contention, showing that media outlets tend to communicate information in ways that confirm their news consumer's prior beliefs and adapt their slant to the political stance of their readers (Gentzkow and Shapiro, 2006, 2010). Media's effects more strongly reinforce existing opinions than create or alter them (Klapper, 1960). Consequently, the existing salience of a topic to an audience often predicts its frequency and tone of coverage in the media (Watt Jr and Van Den Berg, 1978). Put differently, how the press reports an event or issue depends on the target audience. Studies show that this is also true for news coverage of technology-related issues. For example, Marks et al. (2007) find that the existence of a local focus (such as a local incident related to the technology) significantly impacts the news coverage of biotechnology issues, causing countries to differ in their reporting of them. New technologies were also observed to be covered more frequently where there is a greater local significance (Marks et al., 2007). Consequently, we can expect news coverage of innovation and new technologies to vary geographically, because they differ in their relevance to regional societies and, therefore, will be newsworthy to different degrees.

This expected alignment of news media and their audiences is at the center of the present paper, which seeks to assess the degree to which news-media coverage of innovation and new technologies varies systematically in geographical space.

2.2   A regional perspective on innovation news

Above, we indicate that many studies analyze national media discourse on innovation and new technologies (Dudo et al., 2011; Mejía and Kajikawa, 2019). Many of these studies hint at the importance of geography in this context. However, so far, previous research has not adequately addressed the potential variations in news media's content and sentiment with respect to technolo-

gies, on a smaller scale, i.e., subnational and regional. An exception is Stephens et al. (2009), who describe substantial variations in the news concerning wind energy, among the U.S. states of Texas, Minnesota, and Massachusetts.

From a theoretical perspective, this observation does not come as a surprise. Foremost, media themselves are regionalized. Besides national television networks and newspapers, there are large numbers of regional and local news broadcasters and outlets. While such news outlets may have some coverage overlap with national and international news, in light of the previous discussion, they generally must adapt to their regional audience, i.e., they select, frame, and present national as well as regional news, in a way that meets the (perceived) regional demand of their customer base. Although individual traits, such as political views and socio-economic status play important roles in one's exposure to news (Price and Zaller, 1993), two individuals who have similar individual traits may be exposed to different sets of news information just because they reside in different locations. Indeed, Althaus et al. (2009) report that news exposure is strongly related to characteristics of the local news market, even after controlling for individual-level variables. The study's results imply that different regions vary in their preferences for certain types of news. Multiple studies establish geography's role as an important influence on news exposure and consumption (Bogart, 1989; Webster and Lichty, 1991). In the context of the present study, the existing empirical evidence implies that regional characteristics relate to and shape local news consumption, implying that different patterns of news consumption can be expected to exist in structurally heterogeneous regions such as, e.g., cities and rural locations.

In addition to the consumption of innovation and technology news, the supply will likely be strongly regionalized. Notably, the region- and technology-specific institutional setup of innovation activities (frequently labeled "regional innovation systems") greatly impacts the likelihood of novelty creation, adoption, and application (Cooke et al., 1997). Part of this institutional set-up is the (local) news media, which may act as a facilitator by diffusing information, providing coordination and mobilizing (public) support for novelty implementation and experimentation. By disseminating information on what is happening in a region, news media disseminate opportunities and strengthen collective expectations (Nordfors, 2004), thereby building interrelations involved actors and contributing to connecting relevant stakeholders (Blasini et al., 2013). This also relates to the *local buzz* argument. Accordingly, organizations get exclusive access to localized information flows by being present in specific locations, *by being there* (Bathelt et al., 2004). While localized information flows usually refer to labor mobility, collaboration, and spontaneous interaction, in practice, this also includes local news whose limited range involuntarily excludes actors outside of their respective distribution areas. Notably, this does not imply that these information flows are inaccessible to outsiders, per se. Rather it is about actors being made spontaneously and in an unplanned fashion aware of topics, information, and potential contacts (Broekel and Binder, 2007). Arguments supporting a close link between innovation activities and regional news also appear in the literature on technological transition. Here, demand side, representing a crucial factor in the successful emergence and expansion of new technologies and products, receives particular attention (Geels, 2004). More precisely, the argument is that local demand, in combination with local institutions, is essential in the creation of local market niches that allow new technologies (or products) to grow

and evolve until they have reached a developmental stage that gives them a fair chance at non-local markets (Schot and Geels, 2008). This requires a mobilization of early local demand, for which local media besides word-of-mouth communication is essential. Recently, Bednarz and Broekel (2020) empirically confirmed the importance of local demand for the emergence and growth of the German wind-energy industry, even though they do not find that the producers have mobilized this demand.

In sum, we expect regional media to play a crucial role in the information set regarding innovations and new technologies to which individuals are exposed. Moreover, the frequency and content of these information flows are important in the emergence, acceptance, and spatial diffusion of new technologies, at least in the long run. Consequently, better understanding variations in the presentation, discussion, and evaluation of new technologies in the news, at the regional level, is important. This paper is a first step in that direction.

2.3 Regional newsworthiness of innovation

The major objective of this paper is to explore regional differences in innovation news coverage. To this end, we aim to understand how an event or issue becomes news in a region, with what frequency and tone, what determines which event is reported, when and where, in other words, its newsworthiness. Newsworthiness is the likelihood of a news item's selection for publication (Kepplinger and Ehmig, 2006). Media scholars have found that newsworthiness depends on both the nature of the event and the journalistic assessment of the event's relevance to an audience (Staab, 1990; Allern, 2002; Caple and Bednarek, 2016; Eilders, 2006). Kepplinger and Ehmig (2006) disentangle and name these two components of newsworthiness as "news factors" and "news values". News factors are the inherent characteristics of an event, and news values are the judgments about the relevance of the event to the respective audience. News values arise from the fact that news is an economic commodity, and as with all other commodities, its content partly depends on the tastes and preferences of individuals who demand it (Hamilton, 2004). In other words, what people want to hear or read about, and how they feel about certain issues, impact what becomes news in a particular region. Next, we elaborate on how some news factors and values may generate regional variation in news coverage of innovation and news technology-related events.

One determinant of newsworthiness is *unexpectedness* (Galtung and Ruge, 1965), the surprise element of an event or a discussion, arising from novelty, deviance, or unusualness. Although some events are inherently more unexpected than others, since expectations may vary, unexpectedness also partly depends on the target audience (Bednarek and Caple, 2017). From a regional perspective on innovation news, this suggests that the more unexpected an innovation activity or innovation-related event is in a region, the more likely it will be covered. For instance, von Bloh et al. (2019), argue and support empirically that in highly entrepreneurial regions, founding a new enterprise is a relatively less surprising (and, consequently, less newsworthy) event than it is in regions that hardly experience any positive economic dynamics. Accordingly, this suggests that regions more frequently exposed to innovation and new technology-related events and discussions observe that reporting about them is less likely.

Another determinant is *magnitude*. The intensity or potential impact of events increases the likelihood of news coverage (Harcup and Oneill, 2017; Galtung and Ruge, 1965). This suggests

that innovations with a greater (societal) impact are more likely to be presented in the news. Such innovation news is also more likely to emerge in regions with intensive innovation activities, implying generally more innovation events from which to select. Consequently, this effect counters the one previously discussed, since it suggests that in innovation-intensive regions, more high-impact innovation-events are likely to enter the news system, and the chances of innovation-related news being published may be greater.

Relevance is another determinant; the more an issue is perceived as relevant, the more likely is its coverage in the news (Harcup and Oneill, 2017). Consequently, in regions where individuals are more interested in innovation and new technologies, these events are more likely to pass through the journalistic process. The dimensions of *proximity* (geographical and cultural) are a subset of *relevance*, as events occurring close-by tend to be more *relevant* to individuals than those happening far away. Put differently, more geographic or cultural distance between an innovation/technology event and the news outlet decreases the event's newsworthiness (Shoemaker et al., 2007). Proximity is particularly relevant in a regional news context, because the judgment of newsworthiness varies between news outlets (Allern, 2002), and parts of this variation are rooted in geography. Boukes and Vliegenthart (2020) find that domestic stories are over-represented in regional newspapers, compared to some other news media types.

In sum, it is the interplay between the supply of innovation-related events in the news system and the demand of the audience for this kind of story, as well as the inherent characteristics of innovation- and new technology- related events and their relevance to the respective audience, shape the regional coverage of innovation.

How an outlet covers an issue is as important as the frequency of coverage. Thus, the effects above are also likely to shape the sentiments toward innovation and new technologies in the news. We expect that by and large, journalists will seek to comply with existing sentiments among their regional readership, toward innovation, in general, and with respect to specific technologies, in particular. Accordingly, the visibility of innovation and the ways of presenting new technologies in the news media will be a more or less accurate proxy for existing regional public opinion and collective expectations (Fenn and Raskino, 2008; Melton et al., 2016). Nevertheless, we suspect some interference from the journalistic process (e.g., political-sympathies of journalists, events on the national stage), in contrast to the *relevance* effect, might decrease the observable systematic variations at the regional level.

In light of these factors, we aim to investigate if systematic regional variations in the frequency and sentiments of innovation and technology news exist, and to identify their primary determinants (regional characteristics). More precisely, we hypothesize that variations in frequency and tone of news reporting about new technologies are not random, but rather reflect systematic structural differences among the regions. Thus, the study will further deepen our knowledge of the link between news media and conditions for innovation and technology development at the regional level.

We empirically explored this hypothesis by focusing on newspapers. We believe that the afore-mentioned geographical aspect of news exposure can be well understood by looking at regional newspaper readership. While regional newspapers are not the only form of news media, they are the most prevalent form at the regional level (Hutchins, 2004). Regional newspapers are the primary

source of news (Hutchins, 2004) and they contribute to defining the norms of their communities, by producing a powerful discourse (Ewart, 2000). Consequently, they are good proxies for regional news media in general.

3 Data and empirical approach

3.1 Readership shares of newspapers

For the empirical investigation, we needed information on news at the regional level. More precisely, we needed to know what news circulates and what news do inhabitants consume. Unfortunately, most commonly used news databases primarily cover national or international newspapers and provide little information on news at the regional level. Therefore, we used the recently established *Regional News Syndication (RegNeS)* database. This data base provides a daily collection of German-language newspaper headlines and snippets since July 2019. The database covers more than 300 print and online newspapers. As of December 2020, the database consists of more than 6 million unique news entries. The set of newspapers includes regional and national outlets from the German-speaking world (i.e., Germany, Austria, Lichtenstein, Luxembourg, Switzerland).

To simplify data collection and avoid national differences, we exclusively considered newspapers from Germany. Regional newspapers are an essential part of the media landscape and have an extensive readership in Germany (Mangold et al., 2017; Humprecht and Esser, 2018). Newman et al. (2019) reports that weekly usage of a regional or local news media in Germany is 34%, among the highest percentages in Europe. The same report also shows that concerns here about misinformation and disinformation regarding Internet news media are among the lowest, compared to other countries. Consequently, regional news matter in Germany and its outlets are likely to be perceived as trustworthy. Also, innovation processes are much less spatially concentrated in this country than in others and show a strong regional dimension (Brenner and Broekel, 2011).

From the *RegNeS* database, we obtained location information on the regional section in which a news article was published. These sections are very heterogeneous. In larger newspapers, many of them almost qualify as independent daughter newspapers, that share specific subsections of the mother newspaper. These joint sections usually cover topics like national and international politics, as well as economic overviews. The remaining parts are readership- and location-specific. In case of smaller newspapers, these sections refer to sets of news that the paper's editorial offices deem relevant to a specific location. Typical examples are reports on local sports results, information on local cultural events, and the like. For smaller, local and regional newspapers, this information is relatively accurate and reflects the locations in which this newspaper holds a significant readership share. In contrast, for larger, multiregional and national newspapers, this information is insufficient to accurately model their spatial distributions of readership, as it primarily reflects regions consisting of one or more federal states in Germany.

Therefore, we enrich this information with actual regional readership data obtained from the German Audit Bureau of Circulation (IVW). This organization records and audits the distribution of advertising media in Germany and covers most (but not all) newspapers in Germany. From this

organization, we obtained the number of print and digital subscriptions for each newspaper in its database in each district (NUTS3), per day. This data required processing in multiple steps to become useful in the context of the present study. First, weekly readership was calculated by summing the daily numbers, which allows fairly considering weekday and weekend newspapers, as well as those that are published during both parts of the week. Second, the district-level numbers were aggregated to the level of spatial-planning regions (see the discussion below)[2]. Third, the numbers were assigned to the corresponding newspapers in the *RegNeS*-database. For 127 newspapers, a one-to-one matching (based on the newspaper's name) between the two databases is possible. We denote the number of subscriptions newspaper n has in region r as $\text{NR}_{n,r}$.

However, this leaves some of the newspapers featured in the *RegNeS*-database unassigned, many of which are smaller local and regional newspapers. To not lose this information, we modeled their shares based on three assumptions.

A) The number of readers of newspapers not included in the IVW-database (IVW) does not systematically vary between planning regions. Hence, the fourth step was, conditional on this, obtaining an estimate of the total number of newspaper readers in region r by summing the corresponding readership information over all newspapers (N) with readers in region r for all $n \in \text{IVW}$.

$$NR_r = \sum_{n \in IVW} NR_{n,r}. \tag{1}$$

Straightforwardly, we calculated the readership shares (RS) of newspapers with a match in *RegNeS* and *IVW* by using

$$RS_{n,r} = \frac{NR_{n,r}}{NR_r} \quad \text{for } n \in IVW \cap RegNeS. \tag{2}$$

Fifth, this assumption also allows us to calculate the total share of readership that is not accounted for by newspapers with a match in *RegNeS* and IVW (RS_r^*):

$$RS_r^* = \frac{1 - \sum_n RS_{n,r}}{NR_r} \quad \text{for } n \in IVW \cap RegNeS. \tag{3}$$

For the sixth step, we assume that B) all readers not reading IVW newspapers, with a match in RegNeS, buy newspapers that are listed in the *RegNeS* database that lack a match in the IVW database. Conditional on this assumption, we can distribute this shares of readers RS_r^* to the set of newspapers included in the *RegNeS* database, $RS_{n,r}^*$ for $n \in \text{RegNeS}\backslash\text{IVW}$. This assumption implies that all newspapers that are part of IVW are somehow included in RegNeS, but for some reason, no match with IVW could be established.

In the third and final assumption, C) all newspapers in the *RegNeS* database with no match in IVW are assumed to hold an equal share of regional readership, i.e.,

$$RS_{n,r}^* = RS_{m,r}^* \quad \text{for all} \quad m, n \in RegNeS \setminus IVW$$
$$\text{and} \quad n \neq m. \tag{4}$$

---

[2]Spatial planning regions: Raumordnungsregionen (ROR).

On the basis of these assumptions, we calculated the readership shares of any newspaper n in the *RegNeS*-database with at least one regional section associated with a location in r, but lacking a match in IVW. It is the share of readership in IVW, not covered by newspapers with a match in *RegNeS* (RS_r^*), divided by the number of such newspapers. In our data, 104 newspapers fall into this category.

$$RS_{n,r}^* = \frac{RS_r^*}{N} \tag{5}$$

where $n \in$RegNeS\IVW and N is the total number of newspapers such that $n \in$RegNeS\IVW.

Unfortunately, we lack empirical support for these assumptions. Therefore, we made use of the calculated shares in a very conservative manner, by using them in only two instances. First, all news articles in a newspaper, were assigned to the region in which the newspaper has some readership, i.e., $RS_{n,r} > 0$ and $RS_{r,n}^* > 0$. That is, the actual value of the readership share is used in a binary manner, with positive values just seen as an indication of the paper circulating in the region. Second, the shares were used to test the robustness of this allocation procedure, i.e., our estimations were be repeated on subsamples that are identified on the basis of these values.

In total, 231 newspapers were included in this study, with estimated regional readership shares. These newspapers published about 4.3 million news articles within the one year from 01 July 2019 to 30 June 2020. In some instances, the same news article was published by multiple outlets. This might be the case when newspapers share specific sections, or draw from the same pool of articles that national news agencies distribute. Therefore, we assigned a unique identification code to all articles with the same heading and snippet published on the same day. Consequently, the same article was assigned to multiple regions if published by multiple outlets[3].

Our study crucially depends on the choice of an appropriate spatial unit. Yet, this choice involves a trade-off. On the one hand, when using a very fine-grained spatial delineation, the units of observation are likely too small to correspond to the main circulation areas of newspapers. In this case, these units' socioeconomic characteristics were unlikely to be decisive for newspapers' choices about what to report and how. On the other hand, when the spatial units are too large and greatly exceed newspaper circulation areas, we also might not detect a relation because large portions of these regions are not relevant for the newspaper. Moreover, larger territories imply less spatial information in our empirical models, which reduces the chances of identifying any relation. The last aspect, the lack of spatial variance, rules out using the sixteen federal states in Germany. We are left between three spatial levels for which socioeconomic information is available: the city (10,232), district/NUTS3 (429), and planning region (96) levels. To choose between these three, we calculated the distribution of each newspapers' readers for each of the three levels. Subsequently, for each newspaper, we identified the region with the largest share of readers. Figure 1 shows the distribution of these regions' shares, i.e., how many of a newspapers' readers concentrate in a single region for each of the three spatial levels.

We observe that, for more than 50 % of newspapers (excluding national ones), the vast majority of their readers (almost 90 %) concentrate in just one planning region. Accordingly, for most news-

---

[3]On average, an individual news article is associated with 1.2 newspapers and 12.8 spatial-planning regions.

papers, these regions appear to cover almost all of their circulation areas. This share is substantially lower for city-regions (65 %) and NUTS3-regions (75 %). Planning regions also offer sufficient spatial variance with 96 distinct units and therefore, are employed.
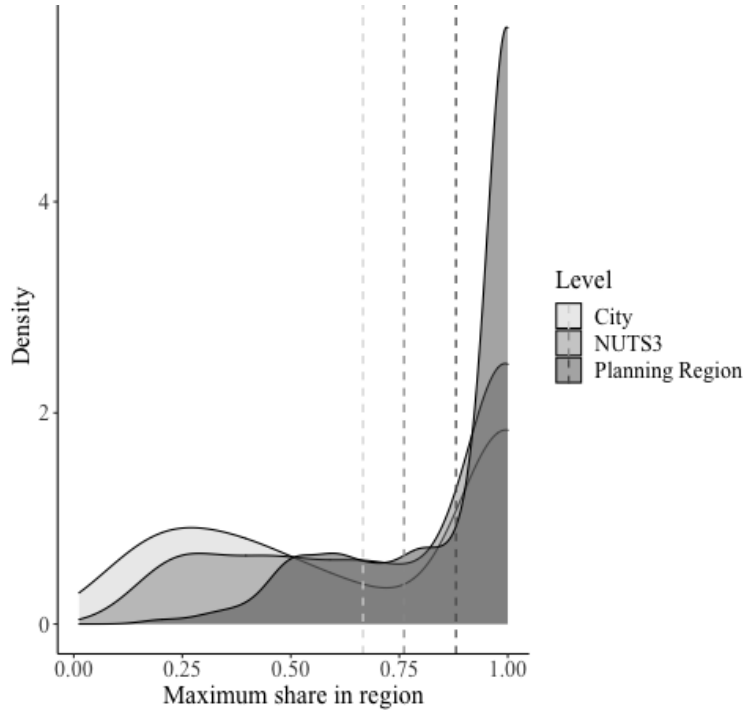

Figure 1: Distribution of the largest share of readers in a single region across newspapers

3.2   News coverage on innovation and new technologies

To identify news articles covering issues of innovation and new technologies, knowledge of topics that the articles cover is crucial. To select relevant articles, we applied a two step-procedure[4]. In the first step, using string-matching algorithms, we identified articles containing at least one of the three keywords on general innovation and new technologies, namely, *technology*, *innovation*, and *science*, very general words that we expected to see in newspaper headlines. Added to these were keywords (AI, automation, and robotics) representing specific technologies that were currently diffusing or emerging in Germany. We based their selection on recent debates in the scientific literature regarding technologies with potentially major impacts on society (see, e.g., Acemoglu and Restrepo, 2019, 2018; Furman and Seamans, 2019; Makridakis, 2017). Related keywords for these technologies were *artificial intelligence*, *automation*, and *robot*[5]. Despite our confidence that these search terms gave us a representative picture of news associated with new technologies and innovation, the choice was

---

[4]Before any text-based analysis, we cleaned the corpus of numbers, punctuation, and stop-words.

[5]The original (stemmed) search strings in German are: *technolog, innovat, wissenschaft, künstlich intelligenz, artifiziell intelligenz, AI, KI, automatisier*, and *robot*.

somewhat arbitrary. Consequently, not all articles were equally relevant, as some containing the word *technology* might not necessarily discuss a new technology or an innovation-related issue. To solve this potential issue, we implemented a second step, using a topic-modeling approach.

We applied the topic-modeling procedure to the set of articles containing any of the above keywords. More precisely, we applied a Latent Dirichlet Allocation (LDA) model. LDA is a probabilistic topic-modeling technique that helps in the automatic discovery of themes/topics/categories in a collection of text documents (Blei et al., 2003). It is one of the most prominent methods for topic modeling and is applied in various fields, e.g., medical sciences, software engineering, geography, and political sciences (Jelodar et al., 2019).

In the context of the present paper, having supervised the probabilities with which every word appears in a certain topic and the probability with which a news article belongs to a specific category, we eliminated news belonging to irrelevant topics and confirmed our associated topics (innovation/new technologies/specific technologies). That is, by applying topic modeling to the preselected sample of articles and classifying these into subtopics, we assessed the extent to which topics the LDA extracted actually reflected topics in which we were originally interested. For instance, an article may include the keyword "innovation" and hence be selected in the first step. In the second step, the LDA might classify the article into a subtopic clearly focusing on innovation and new technologies, while in another instance, such an article might be classified into a non-innovation-related topic such as "education". Consequently, this procedure minimizes the false-positive error -that is, even while an article contains the word "innovation", among all such articles it might be classified as primarily about education. In this case, it is most likely that the article will be about education with "innovation" being not much more than a side issue. In the following, we exclusively considered articles to be related to innovation and new technologies if they contained any of the above keywords, and LDA classified them as so related. Of course, the minimization of the false-positive comes at the expense of the false-negative error rate. That is, we may not consider all articles related to innovation and new technologies. However, we believe that applying a more conservative approach was more appropriate, given the large number of articles in our database. More details regarding LDA and the topic-modeling procedure appear in A.

At the end of the first step, $23,849$ unique news articles were found to include one of our search keywords. At the end of the second step, i.e., topic modeling, we found $20,302$ were indeed about innovation and new technology-related events or discussions[6] .

3.3 Sentiment analysis

In addition to the variation in the frequency of innovation news reporting, we were interested in potential differences in the sentiments with which they are presented. To obtain a measure of articles' sentiments, we used an automated sentiment analysis tool introduced by Rauh (2018). The tool is built on two widely used German sentiment lexicons, namely, *Sentiment Wortschatz*, and *German Polarity Clues*, developed by Remus et al. (2010) and Waltinger (2010), respectively. Crucially, this tool considers negation. With this approach, we identified the number of positive and

---

[6]To improve the readability of the paper, we will exclusively refer to "innovation news" and "news on innovation" from now on, which, nonetheless, includes news on new technologies.

negative terms and calculated the sentiment polarity score of each article, by using the following formula:

$$SENT = \frac{\#POS - \#NEG}{\#POS + \#NEG} \tag{6}$$

where $\#POS$ and $\#NEG$ denote the total number of positive and negative terms, respectively. The denominator equals the total number of sentiment bearing words.
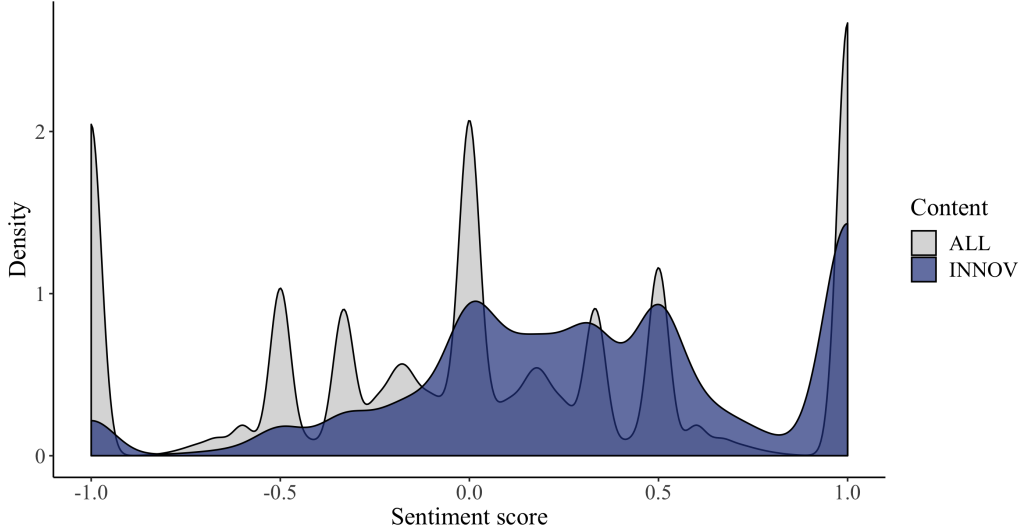

Figure 2: Sentiment distribution of innovation news

The distribution of sentiment of the articles appears in Figure 2, where the lighter color shows the distribution for all news articles and the darker color for the innovation news (INNOV). Comparing the sentiments towards innovation to the general news sentiment, we see that they mostly receive positive coverage in German news media. Their mean sentiment polarity score is 0.33, whereas the average score of all news is 0.05. Given these results, we infer that news generally bears a neutral sentiment, while innovation and new technologies are indeed (absolutely and relatively) good news in Germany!

3.4 Regionalizing the news data

Our initial observations are individual news articles. However, newspapers captured in *RegNeS* vary greatly in the number of articles they publish. Moreover, decisions about articles (e.g., content, sentiment, in what regional section in which to publish) are made at the newspaper level. Inversely, readers do not pick what newspaper to read on the basis of an individual article but, rather, by assessing the entire package of articles that newspapers present over a certain time. Consequently, the link between regional characteristics and what is read in a region must be modeled at the newspaper level. Therefore, we aggregated the article-level information at that level. We also must consider that most of the newspapers serve readers in multiple regions, implying that our unit

of observation ,called *newspaper-regions* in the following, is the combination of newspapers and regions in which they have a positive readership share ($RS_{n,r} > 0$ and $RS_{n,r}^* > 0$; see Section 3.1). Accordingly, in our final data, each newspaper appears as many times as there are regions in which it has a positive readership.

To capture the variations with which articles in newspaper-regions contain news about innovation, we constructed our first dependent variable, $INNOV_{n,r}$. The value of $INNOV_{n,r}$ equals the number of articles that deal with innovation in newspaper n circulating in region r. INNOV is expected to increase with the total number of articles published by the respective newspaper. Therefore, including the total number of news articles published in the respective newspaper-region, $NNEWS_{n,r}$, in any kind of evaluation is essential.

In Figure 3, we illustrate the share of innovation news in total news (INNOV/NNEWS), aggregated at the spatial-planning region level. It can be interpreted as representing the likelihood of reading about innovation when randomly picking up a newspaper circulating in the region. The map provides a first insight into the regional variation in the frequency with which newspapers cover innovation. It reveals a strong imbalance in the share of innovation news coverage, where generally lower shares seem to characterize larger metropolitan regions (Berlin, Hamburg, Munich, Frankfurt).
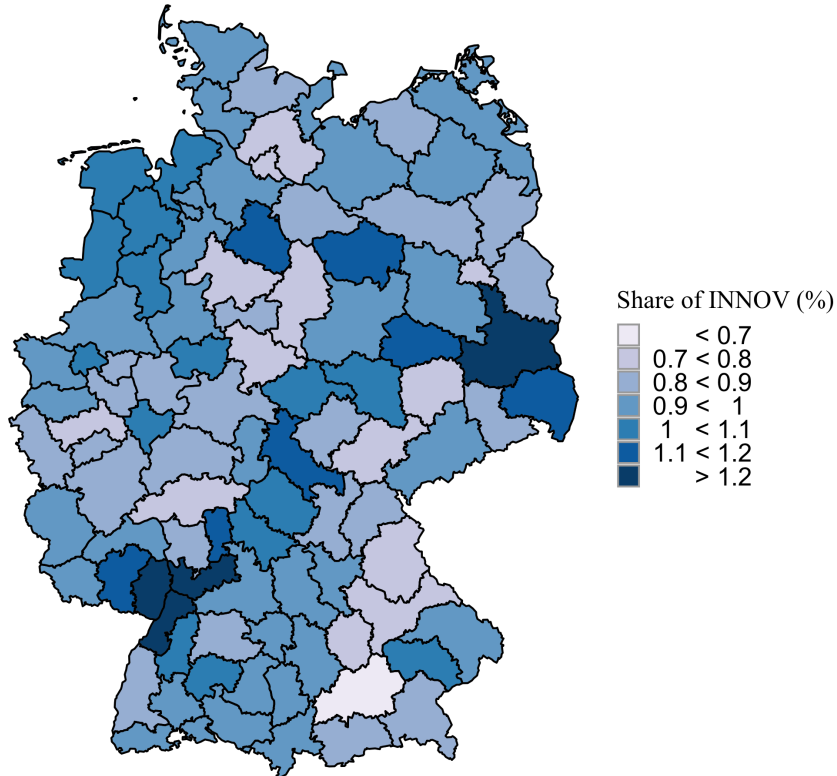

Figure 3: Share of innovation news in total news

Figure 4 represents the sentiments of innovation news ($\text{SENT}_{n,r}$), corresponding to the average sentiment of articles covering innovation and new technology-related issues in region r. Darker colors correspond to a relatively more positive sentiment towards innovation. As in the case of the frequency of innovation news, we observe a substantial spatial imbalance, explained later. However, in contrast to the previous map, in this case, a clear North-South pattern is visible. Regions in Bavaria and Baden-Wuerttemberg, as well as Thuringia, seem to have newspapers that report these issues more positively.
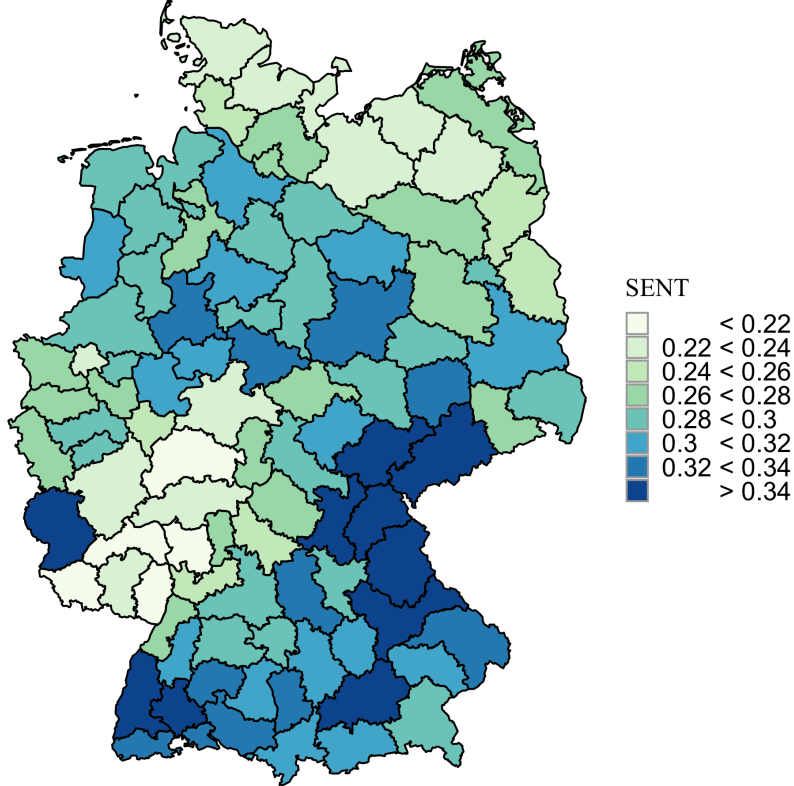
Figure 4: Average sentiment of innovation news per planning region

3.5 Regional variables

To investigate if variations in the frequency and tone of innovation news vary systematically with regions' socioeconomic characteristics, we considered a range of variables. We differentiated between urban or rural regions by means of their population density (POPDENS). Urban regions are expected to generate a larger number of "activities" and "events" to report, including more frequent innovation-related events. Accordingly, there is a greater probability that some of them to show up in the news. However, as discussed earlier, what is *newsworthy* is a relative concept, and differences

between urban and rural regions might emerge because of innovation news differing in degrees of *unexpectedness*.[7]

The economic development of regions is straightforwardly approximated by the gross domestic product per capita[8] (GDPC). We expect more economically more developed regions to generate and their news outlets to report more innovation news. However, that might also trigger a reduction in *unexpectedness* and lower the frequency of innovation news and its sentiments.

In light of the discussion above on some new technologies potentially threatening the demand for human labor, we include the regional unemployment rate (UNEMP) as a potential explanatory factor. Due to the *relevance* effect, we expect more frequent innovation news coverage with a more negative tone in regions with higher unemployment rates.

Given the peculiar history of the two parts of Germany, there (still) seem to be systematic differences in journalistic activities (Haller, 2012) and this may potentially impact the judgment of newsworthiness and sentiment of coverage. For this reason, we control for regions belonging to the former East Germany with the dummy variable EAST. The variable takes a value of 1 for the spatial-planning regions that were in the former German Democratic Republic (GDR) and 0 for those remaining.

The regional socioeconomic indicators were obtained from the statistical offices of the German federal states and has been sourced from the INKAR database[9]. We used the most recent data available for each variable (2018). Given the cross-sectional nature of our research and the limited temporal variance of these variables, we are confident that they are a sufficient match of our 2019-2020 news data.

Finally, regions vary in terms of the number of newspapers available, which might impact the likelihood of reading about innovation. Larger numbers of newspapers may bring a certain degree of diversity to regions, with newspapers specializing in different topics, or fierce competition may push newspapers to publish what is found the most newsworthy in the region, making them similar in content. Thus, in order to capture how this diversity impacts innovation news coverage, we include NNP_r, the number of newspapers in region r for all n such that $RS_{n,r} > 0$.

3.6 Control variables

Although our dependent variable is at the newspaper-region level, and we are rather interested in its variation; depending on the regional variables, some newspaper-level variables may potentially impact our results. Accordingly, we control for these.

First is the average length of articles that newspapers published. The length might impact the likelihood of detecting search keywords and, thus, might have an impact on the number of innovation news items found. Since the sentiment index depends on the total number of sentiment-

---

[7]We also considered the number of patents granted (PATC) per capita, as an indication of the research and development activities. More R&D increases the chances of having more "spectacular" innovations that may find their way into newspapers. However, the variable turned out to be correlated with POPDENS and did not add to the models. Therefore, we dropped it and refrained from reporting the corresponding results at this point. They can be obtained from the authors upon request.

[8]The variable is measured in thousands.

[9]www.inkar.de.

bearing words, the total number of words in an article is also likely to have an impact on the assigned sentiment score. The $NWORD_n$ variable is constructed to this end and it simply shows the average number of words an article contains for each newspaper.

Second, since newspapers significantly vary in their reporting styles, the sentiment conveyed in reporting innovation news also depends on the general tendency of newspapers to frame issues more positively or negatively. To control for newspapers' general sentiment, we construct the control variable $NPSENT_n$, which represents the average news sentiment (regardless of the topics covered) for each newspaper.

Third, while some newspapers in our data set are regional and local newspapers, some are national, available and with readership in almost all regions. Since journalistic practices might systematically differ between national and regional newspapers, we control for this with the dummy variable, $NATIONAL_n$. The variable takes the value of 1 for the newspapers with readership share in more than 50 of 96 spatial-planning regions, and 0 otherwise.

Table 1 and Table 2 provide the relevant summary statistics for all variables used in the paper. Correlation of the variables appears in Table B1 in the B.

Table 1: Summary statistics

	# Obs	# NA	Min	Max	Median	Mean	Std.dev
INNOV	2002	0	0.00	1,684.00	114.00	253.34	351.49
NNEWS	2002	0	455.00	116,666.00	23,312.00	29,695.66	24,825.08
SENT	1985	26	-0.50	1.00	0.26	0.29	0.17
POPDEN	96	0	42.00	4,055.00	180.00	335.17	521.37
GDPC	96	0	24.20	65.90	34.75	36.51	8.51
UNEMP	96	0	2.20	10.90	5.25	5.55	2.17
NNP	96	0	11.00	40.00	18.00	20.85	6.39
NWORD	231	0	5.50	435.70	31.90	42.80	59.70
NPSENT	231	0	-0.74	0.52	-0.01	-0.01	0.10

Table 2: Summary statistics for categorical variables

	Levels	# Obs	Perc (%)
EAST	1	22	22.92
	0	74	77.08
	All	96	100.00
NATIONAL	1	8	3.46
	0	223	96.54
	All	231	100.00

3.7 Empirical approach

As pointed out above, the unit of observation is newspaper-regions, i.e., we observe the frequency of newspaper n featuring articles about new technologies and innovation in region r, as well as the sentiments they convey. As newspapers transcend regions, this creates a relatively complex dependency structure among the observations, which renders using standard regression analyses invalid. This becomes obvious in the construction of our dependent variables, which are partly based on newspaper-level information that is the same in all regions in which the newspaper circulates. Put differently, the values associated with newspaper n in region r_i are more likely to be similar to those in r_j than to those held by chance. Such types of problems are common in spatial research, which lends us the methodologies addressing this. More precisely, we use cross-sectional spatial regression approaches and several Lagrange Multiplier (LM) tests (introduced by Anselin et al. (1996) and Anselin (2013), and implemented by Bivand and Piras (2015)) to identify the appropriate models. In all cases, these tests recommend the use of the so-called spatial error model (SEM) (LeSage and Pacey, 2009), which takes the form:

$$y = \beta X + u \ \text{ where } \ u = \lambda W u + e. \tag{7}$$

In Equation 7, y denotes the dependent variable, X denotes the matrix of explanatory variables, and β is the corresponding vector of coefficients. The disturbance term u is spatially auto-correlated, where λ is the spatial auto-regressive parameter and e is the usual independent and identically distributed disturbance. W denotes the spatial weight matrix capturing the spatial dependencies among the observations. Usually, the spatial weight matrix is defined on the basis of geographical neighborhoods or distances, which represent the underlying spatial structures. However, in this study, dependencies arise from the fact that the same newspaper is circulating in multiple regions. Consequently, we construct the weight matrix on this basis. That is, we consider observations (newspaper-regions $r_{n,i}$ and $r_{n,j}$) to be neighbors when newspaper n is circulating in both of them.

$$w_{i,j} = \begin{cases} 1 & , & RS_{n,i} > 0 \ \text{ and } \ RS_{n,j} > 0 \\ 0 & , & otherwise \end{cases} \tag{8}$$

In a similar manner, the matrix is row-standardized before being transformed into spatial weights. The LM-tests revealing the presence of this dependency structure in the data confirm the appropriateness of this specification. Notably, our first dependent variable is a count variable ($INNOV_{n,r}$). Unfortunately, methods for dealing with spatial auto-correlation and count data are not yet sufficiently developed (Glaser, 2017). Consequently, we apply the second-best solution, namely, log-transforming the variable.[10]

---

[10]We also add 1 to the value before the transformation to ensure finite values.

4 Results and discussion

4.1 Regional variation in innovation news coverage

To explore the determinants of regional variation in the frequency with which newspapers cover innovation and new technology-related information, we use INNOV as our dependent variable. It is regressed onto the previously described set of variables using the spatial error model.[11] Table 3 shows the according results. In Section 3.1, we pointed out that our construction of readership shares that are used to allocate news articles across regions, are based on a number of assumptions. Therefore, we must explore the robustness of our findings, with respect to the specification of this allocation procedure. We do this by repeating our regression analysis for subsamples of our data, defined on the basis of different threshold values of regional readership shares. Newspapers are required to exceed these values before allocating their articles to a region. More precise, in addition to the baseline scenario with no threshold on the readership share, we construct subsamples of the data that include only articles of newspapers that exceed regional readership shares of at least 0.1%, 0.5%, and 1%, respectively. We cannot test for greater thresholds, as the numbers of remaining observations fall considerably. Since increasing the thresholds implies lower chances of newspaper misallocation, we interpret findings that hold in multiple scenarios, as being more reliable and less conditional on the allocation procedure. However, it turns out that almost all variables with significant coefficients in the baseline scenario remain significant in the other scenarios as well. Consequently, our results appear to be very robust, with respect to alterations in the matching of newspapers to regions.

Before looking at the variables representing regional characteristics, the control variables are worth discussing. The variable log(NWORDS) obtains a positive and statistically significant coefficient, confirming our expectations. Innovation news items are more likely to appear in newspapers with longer articles. For similar reasons, log(NNEWS) becomes significantly positive as well; The more news articles newspapers publish, the more likely it is that some of them will refer to innovation and new technologies. Accordingly, both variables' importance is of a rather technical nature, as they control for "size-effects" or features of the data-collection process.

Insightful results are obtained for NATIONAL. Its coefficients are also significantly positive. Accordingly, national newspapers are more likely to feature articles dealing with innovation and new technologies. This finding most likely reflects the fact that some larger national newspapers have dedicated news sections to technological issues, missing from regional newspapers. For instance, the prominent German newspaper *Frankfurter Allgemeine* has a section called "Technik und Motor" (*Technology and Engine*). However, it may also be due to national newspapers seeking to cover more general newsworthy issues of interest to the whole country.

---

[11]As discussed in Section 3.7, the choice of employed spatial model is based on the Lagrange multiplier (LM) diagnostics. Test-statistics and corresponding p-values for each regression appear in C.

Table 3: Frequency of innovation news coverage

RS$_{n,r}$	log(INNOV)			
	> 0	> 0.001	> 0.005	> 0.01
log(POPDEN)	0.058*	0.085*	0.091*	0.116**
	(0.027)	(0.035)	(0.038)	(0.040)
log(UNEMP)	−0.072	−0.199**	−0.184*	−0.116
	(0.057)	(0.077)	(0.082)	(0.090)
log(GDPC)	0.063	−0.205	−0.094	−0.016
	(0.109)	(0.156)	(0.163)	(0.174)
EAST	0.162***	0.195**	0.185**	0.198**
	(0.046)	(0.060)	(0.066)	(0.073)
log(NWORD)	1.153***	1.245***	1.203***	1.174***
	(0.025)	(0.035)	(0.036)	(0.039)
NATIONAL	0.901***	1.068***	0.868***	0.673***
	(0.032)	(0.045)	(0.048)	(0.056)
log(NNP)	−0.128	0.059	0.090	0.019
	(0.073)	(0.073)	(0.060)	(0.061)
log(NNEWS)	1.033***	1.012***	1.019***	1.028***
	(0.013)	(0.019)	(0.021)	(0.022)
Constant	−9.830***	−9.493***	−9.926***	−10.286***
	(0.407)	(0.582)	(0.612)	(0.650)
λ	0.061	−0.047	0.026	−0.006
	(0.046)	(0.056)	(0.055)	(0.059)
Observations	2002	1183	863	679
Max VIF	3.130	3.160	2.970	2.990
Log Likelihood	−1898.674	−1205.481	−792.831	−598.305
AIC (Linear)	3818.998	2431.674	1605.875	1216.619
AIC (Spatial)	3819.348	2432.961	1607.661	1218.609
LR test: statistic	1.650	0.713	0.214	0.010
LR test: p-value	0.199	0.398	0.644	0.922

Note: *p<0.05; **p<0.01; ***p<0.001

The first regional variable for which we obtain a significant coefficient in the baseline model and in all threshold-scenarios is population density (POPDEN). Its coefficient is significantly positive suggesting that news articles on innovation are more likely to appear in newspapers circulating in urbanized regions. The explanation for these findings may reside in a generally greater interest among urban readership in topics related to innovation and new technologies, i.e., the relevance of this topic to the urban audience; or in the availability of more unexpected and bigger events in these areas that regional newspapers pick up. Given the substantial concentration of innovation and technological activities in urban regions (Feldman and Audretsch, 1999; Broekel and Brenner, 2011), innovations mostly emerge in cities. Of these innovations, only a few may have a substantial impact on society in general (e.g., "COVID-19 vaccine") and, thus, interest for readers in multiple regions. However, many innovations may be of relatively greater importance to individual regions, either representing success events of local businesses or research institutions, or having noteworthy consequences, such as securing or expanding local employment. Consequently, regions where more

such events occur (urban areas), offer a larger pool of newsworthy events to report. According to our results, this also translates into a higher share of innovation news in the newspapers circulating in these places. An alternative to this supply-side (or news factors) argument, is a greater demand for such news in urban areas. Cities tend to have more specialized in high-tech industries and high-skilled jobs (Gomez-Lievano et al., 2018). This creates an audience that is likely to be more interested in news about innovation. Accordingly, newspapers that seek to appeal to this audience will feature more articles of this kind, in urban areas. This greater exposure to innovation-related news in urban regions will contribute to the formation of a collective expectation generally (Konrad, 2006), and more positive ones in particular. Thus, it paves the way for quicker adoption (Budde et al., 2012), which allocates the early-adopter advantages of innovations to cities at the expense of rural regions.

The second variable with a consistently significant coefficient in all scenarios is EAST, i.e., the indication of a region being located in the territory of the former GDR. The variable's coefficient is positive, highlighting that newspapers circulating in these regions tend to have more innovation news. Interestingly, in contrast to POPDEN, a supply-side explanation seems unlikely to underlie this finding. Even thirty years after the reunification, on average, innovation activities are still not at the level they are in the western part of the country (Gomez-Lievano et al., 2018). Consequently, we can only speculate about factors on the demand side, i.e., the relevance of innovation and new technology related discussions to individuals living in the territory of the former GDR. It might be related to the higher average age of media consumers in East Germany (Gomez-Lievano et al., 2018), or that the history of the GDR, with its stronger focus on technologies and natural science in education, still shaping the perception of and interest in these issues (Gensicke, 1995). This clearly deserves more attention in future research.

Restricting articles to newspapers exceeding a minimum regional readership (RS> 0.001 and RS> 0.005), the coefficient of the rate of unemployment is becoming significantly negative. However, that does not hold for the strictest scenario, a minimum readership share of at least 1%. Accordingly, it can be seen as a weak indication of regions with higher unemployment having lower shares of innovation news in newspapers circulating there. Again, both supply- and demand-side explanations are possible. On the supply side, high unemployment tends to go along with less dynamic economic and technological developments. On the demand side, the local newspaper audience may be less interested in such topics, as there is at least one much more pressing issue - the high unemployment, which leaves less room for news on innovation. Moreover, and somewhat more likely, high unemployment in regions goes hand in hand with lower levels of highly-skilled human capital and high-tech industries. In this sense, high-unemployment may indicate a less technology-interested audience in general, and these topics are less relevant to the audience that our model picked up.

Together with the previous finding concerning urbanization and its strong correlation to the number of patents (see footnote 7), these observations imply that from a regional perspective, rarity does not drive the newsworthiness of innovations. Unexpectedness is a driver of newsworthiness (Galtung and Ruge, 1965), so the relatively greater rarity of innovation events in rural areas and regions with higher unemployment should have increased the likelihood of their being reported. We observe the opposite: Regions with fewer innovations are not associated with relatively greater

newsworthiness or unexpectedness of individual innovation events. In contrast, the finding can be better explained by the *relevance* and *proximity* effect (Harcup and Oneill, 2017; Shoemaker et al., 2007), in regions where innovation more closely relates to or more strongly shapes socioeconomic conditions.

4.2 Regional variation in innovation news sentiment

The frequency with which innovations appear in the news is a crucial condition for the general public to learn about them. However, for adoption or, at least gaining legitimacy,their presentation in a positive light is also essential. The aforementioned example of the media presentation of bioenergy in Sweden and Norway leading to different diffusion patterns in each country (Skjølsvold, 2012) highlights this likelihood. Thus, it is important to explore the representation of technological developments and innovations in different regions. A spatial regression model, using SENT as a dependent variable, explores the degree to which sentiments on innovation differ between regions. Table 4 presents the corresponding regression results.

As in the previous subsection, a look at the robustness of our results, with respect to specification in the regional allocation procedure, is worthwhile. Given that the models show little sensitivity to alternative specifications, our empirical approach appears to work well when it comes to exploring regional variations in sentiments as well.

The first control variable of interest in this model is NPSENT, i.e., the average sentiment of articles in the focal newspaper. The variable obtains a significantly positive coefficient in the baseline and all other scenarios. This suggests that newspapers that generally have a relatively more positive tone in their articles show it in news on innovation as well. In itself, this is not surprising. However, it has substantial consequences for the interpretation of the other variables because it implies that we are controlling for newspaper-level effects. Put differently, our findings show the relation between (regional) variables and sentiments on innovation news that go beyond the general tone of newspapers.

With respect to other control variables, the models indicate a significantly negative relationship with the length of articles (NWORD), which is primarily technical, as the number of (sentiment-bearing) words serves as the denominator in our sentiment index and the number of these words tend to increase with the length of articles. The second strongly significantly negative control variable is the number of innovation news items (INNOV) in the respective newspaper. Accordingly, newspapers that report less about innovation, tend to focus on a more positive representation of these topics[12]. In contrast, newspapers with a greater dedication towards this topic seem to take a more critical stand or give more room to less-positive evaluations of innovations and new technologies.

As in the case of innovation news' frequency, we observe significantly positive coefficients for NA-TIONAL, which substantiates the structural difference between regional and national newspapers. Not only do the latter report more frequently about innovations and new technologies; they also do this in a more positive manner. One potential but purely speculative explanation for this might be that national newspapers rather focus on the overall impact of innovations on societal develop-

---

[12]Note that the newspapers with no innovation news are excluded in this analysis.

ment. In contrast, regional newspapers' focus on particular regions that may be more concerned about their potential negative (local) consequences. The literature on regional innovation increasingly recognizes that many innovations may actually contribute to the growing spatial inequalities of economic and social status (see for a recent review, see Biggi and Giuliani, 2020). Consequently, innovations might endanger the economic development of many regions, while society, as a whole, benefits from them. Our findings on the difference between national and regional newspapers may reflect this. Again, we have no reliable empirical support for this claim at this stage, which calls for more research on the issue.

Table 4: Tone of innovation news coverage

	log(SENT)			
RS$_{n,r}$	> 0	> 0.001	> 0.005	> 0.01
log(POPDEN)	0.007	0.003	0.007	0.002
	(0.004)	(0.005)	(0.006)	(0.007)
log(UNEMP)	−0.030***	−0.029*	−0.029*	−0.029
	(0.009)	(0.011)	(0.014)	(0.017)
log(GDPC)	−0.015	−0.015	−0.029	−0.024
	(0.018)	(0.022)	(0.027)	(0.033)
EAST	0.013	0.013	0.012	0.017
	(0.007)	(0.009)	(0.011)	(0.013)
log(NWORD)	−0.045***	−0.049***	−0.047***	−0.054***
	(0.004)	(0.006)	(0.007)	(0.008)
NATIONAL	0.033***	0.028***	0.050***	0.054***
	(0.006)	(0.008)	(0.010)	(0.012)
log(NNP)	−0.022	−0.019	−0.019	0.004
	(0.012)	(0.011)	(0.010)	(0.012)
log(INNOV)	−0.037***	−0.039***	−0.038***	−0.034***
	(0.002)	(0.002)	(0.003)	(0.004)
log(NPSENT)	0.736***	0.727***	0.820***	0.845***
	(0.028)	(0.037)	(0.046)	(0.054)
Constant	0.690***	0.718***	0.719***	0.686***
	(0.061)	(0.077)	(0.094)	(0.112)
λ	−0.130*	−0.044	−0.120	−0.157*
	(0.057)	(0.057)	(0.064)	(0.069)
Observations	1985	1171	856	672
Max VIF	3.170	3.160	2.970	2.990
Log Likelihood	1748.783	1075.972	746.022	533.993
AIC (Linear)	−3470.071	−2129.323	−1466.319	−1040.538
AIC (Spatial)	−3473.565	−2127.944	−1468.043	−1043.987
LR test: statistic	5.494	0.622	3.725	5.449
LR test: p-value	0.019	0.430	0.054	0.020

Note: *p<0.05; **p<0.01; ***p<0.001

With respect to regional variables, we primarily identify the unemployment rate (UNEMP) as having a robust and nearly consistent negative relation with the sentiments of news on innovation. Only in the models with a threshold of 1%, the coefficient remains insignificant at the 0.5 level

(though it is significant at the 0.1 level). Accordingly, the interpretation calls for exercising caution, as the result is conditional on considering even relatively small readership shares in the news articles' spatial allocation.

The finding on UNEMP implies that regions with higher levels of unemployment tend to have newspapers in circulation that report relatively more negatively about innovation and new technologies. It is important to underline that, as Section 3.3 shows, news on these issues generally has a positive tone, implying that this is a strictly relative perspective. Nevertheless, this raises the question of why newspapers in regions with higher rates of unemployment report more negatively about innovation and new technologies (and, as shown in the previous section, also less frequently) than regions with lower rates of unemployment? An explanation that fits to this finding is the potential regional variations in the sentiments towards specific technologies. Clearly, many have great potentials to revolutionize our ways of living and promise prosperity (see, e.g., the discussion on AI in Agrawal et al., 2019; Korinek and Stiglitz, 2017). However, these potential benefits are unlikely to spread uniformly geographically. Put differently, only some regions will experience these benefits, while others rather face negative consequences when these technologies become widely adopted. Moreover, most contemporary new and revolutionizing technologies will materialize in regions with the necessary infrastructure and related economic structures in place, e.g., advanced ICT, biotechnology, and industry 4.0 (Iammarino et al., 2019). It seems reasonable that these technologies would also be the ones most discussed in today's newspapers. Given their potential to increase spatial inequality and challenge existing regional economic comparative advantages, for many less-developed regions, these technologies represent a threat rather than bright promises. In this case, we can expect the news media to focus more strongly on such technologies' "dark sides" in regions where their effects are envisioned as rather negative. Our empirical findings seem to support that this might be the case. Yet admittedly, the findings do not provide any direct empirical proof.

To get a somewhat more detailed picture of the matter, we repeat the analysis with a focus on two technologies. They fall into the category of technologies with the potential to boost economic prosperity in some regions that already possess an advanced technological infrastructure and economic basis, namely, AI and automation (Iammarino et al., 2019). On the other hand, they challenge the foundation of many less-advanced regional economies. Section 3.2 describes the identification of the corresponding sets of news articles. All variables are the same as in the previous models. The only difference is that the dependent variable log(SENT), for AI and automation exclusively indicates the average sentiment of the news articles that contain the words AI and automation, respectively. The regression results based on the two distinct subsamples appear in Table 5. Interestingly, the results hardly differ from what we obtained for the total set of news on innovation. This is somewhat surprising, given that AI-related and automation-related news only account for 8.5% and 18% of all innovation news, respectively.

Besides the control variables, unemployment still obtains a significantly negative coefficient in all specifications. Crucially, it becomes significant in all specifications for automation, while in the case of AI, the coefficient remains insignificant in the most restrictive model requiring at least 1% readership. In any case, these findings clearly show that in regions where the unemployment rate is higher, when reporting about the developments in AI and automation, newspapers focus on the

potential threats more than the opportunities these technologies may unfold. This adds to our argument that technologies threatening existing economic advantages are discussed relatively less positively in regions with high unemployment.

The observed relationships between unemployment and sentiments towards innovations and new technologies reinforces the arguments concerning frequency: High *relevance* of innovations for the regional audience that originates in an urban, technology-driven economy accompanies an audience interested in a positive presentation of such topics and newspapers satisfying this demand. Crucially, since such news not only appears with greater frequency but also in a more positive fashion, newspapers are likely contributing strongly to the building of positive collective expectations (Konrad, 2006; Budde et al., 2012). Consequently, newspapers seem to facilitate innovation diffusion, which particularly works to the benefit of already well-developed regions. Also, this example of a link between regional socio-economic structures and (regional) news, supports the view that newspaper data gives approximate insights into local public opinions and attitudes (see, also Fenn and Raskino, 2008; Melton et al., 2016).

Table 5: Tone of AI and automation news coverage

	log(SENT)							
	Artificial Intelligence				Automation			
$RS_{n,r}$	> 0	> 0.001	> 0.005	> 0.01	> 0	> 0.001	> 0.005	> 0.01
log(POPDEN)	0.008*	−0.008	−0.011	−0.010	0.027*	0.021	0.034*	0.030
	(0.004)	(0.010)	(0.011	(0.013)	(0.011)	(0.012)	(0.014)	(0.016)
log(UNEMP)	−0.064***	−0.047***	−0.028**	−0.022	−0.036***	−0.035**	−0.036**	−0.042*
	(0.017)	(0.011)	(0.021)	(0.028)	(0.009)	(0.011)	(0.013)	(0.017)
log(GDPC)	−0.008	−0.005	−0.015	−0.002	−0.018	−0.016	−0.026	−0.031
	(0.015)	(0.018)	(0.022)	(0.025)	(0.017)	(0.022)	(0.026)	(0.032)
EAST	0.018**	0.012	0.010	0.012	0.014*	0.019*	0.020	0.026
	(0.006)	(0.007)	(0.009)	(0.010)	(0.007)	(0.009)	(0.011)	(0.013)
log(NWORD)	−0.044***	−0.048***	−0.045***	−0.047***	−0.045***	−0.049***	−0.047***	−0.051***
	(0.004)	(0.005)	(0.005)	(0.006)	(0.004)	(0.005)	(0.006)	(0.008)
NATIONAL	0.019***	0.020**	0.039***	0.043***	0.030***	0.025**	0.046***	0.052***
	(0.005)	(0.007)	(0.008)	(0.010)	(0.006)	(0.008)	(0.010)	(0.012)
log(NNP)	−0.040*	−0.059***	−0.047**	−0.062***	−0.026*	−0.025*	−0.025*	0.000
	(0.023)	(0.020)	(0.019)	(0.021)	(0.011)	(0.011)	(0.010)	(0.012)
log(INNOV)	−0.038***	−0.046***	−0.047***	−0.047***	−0.035***	−0.034***	−0.032***	−0.028***
	(0.002)	(0.002)	(0.003)	(0.004)	(0.002)	(0.002)	(0.003)	(0.004)
log(NPSENT)	0.606***	0.648***	0.747***	0.760***	0.705***	0.739***	0.822***	0.840***
	(0.030)	(0.038)	(0.047)	(0.051)	(0.029)	(0.039)	(0.049)	(0.057)
Constant	0.695***	0.752***	0.738***	0.701***	0.689***	0.707***	0.694***	0.672***
	(0.054)	(0.064)	(0.074)	(0.085)	(0.058)	(0.074)	(0.091)	(0.109)
λ	0.066	−0.078	0.006	−0.104	0.001	0.006	0.062	0.043
	(0.056)	(0.068)	(0.066)	(0.076)	(0.054)	(0.057)	(0.057)	(0.060)
Observations	1581	991	712	553	1892	1118	810	631
Max VIF	3.360	3.130	2.880	2.940	3.240	3.170	2.980	3.010
Log Likelihood	1775.918	1192.577	851.087	643.267	1821.544	1092.333	751.058	535.133
AIC (Linear)	−3528.509	−2361.818	−1680.164	−1262.618	−3621.088	−2162.655	−1478.989	−1047.785
AIC (Spatial)	−3527.836	−2361.153	−1678.173	−1262.535	−3619.088	−2160.666	−1478.117	−1046.265
LR test: statistic	1.327	1.335	0.009	1.917	0.001	0.011	1.128	0.480
LR test: p-value	0.249	0.248	0.924	0.166	0.980	0.915	0.288	0.488

Note: *p<0.05; **p<0.01; ***p<0.001

Two further results are worth pointing out. EAST is positively significant in the case of AI and automation in the baseline models. This suggests a tendency toward somewhat more positive presentations of both technologies in the eastern part of Germany. Besides confirming the (still) existing structural differences between the two parts of Germany, we interpret this finding as newspapers in East Germany reflecting a more positive attitude toward these technologies. We also find a weak indication (i.e., coefficient of POPDEN being significantly positive at the 0.5 level in the baseline models) that sentiments of AI- and automation-related news are more positive in urban regions. A potential explanation might be that the impacts of automation on employment are expected to be lower in cities (Frank et al., 2018). Accordingly, in these places, people might feel less threatened by automation, and news outlets might be less inclined to cover specifically negative aspects or they focus on these technologies' positive side.

5   Implications and conclusion

Aggregate expectations play an important role in the development, diffusion, and use of new technologies. The frequency of exposure to innovation- and new technology-related information and the tone of the presentation are crucial for the formation of collective expectations and public opinion. One important channel by which information and opinions about innovation are diffused is the news media. They have a strong geographical dimension, and this may have substantial consequences, in the sense that the heterogeneity in the media reporting is likely to translate into unequal exposure and, thus, differing opinions. Studies widely confirm this at the national level, showing that media differ between countries in frequency and tone of coverage about innovation and new technologies.

The present paper adds to this literature with a complementary study at the subnational (regional) level. More precisely, we argue that so far, much of the literature neglects the realization that innovation processes and technological diffusion are at least as much a subnational process as a national one. In addition, most of the media are also strongly organized at the subnational level. With few nationwide outlets, regional and local newspapers strongly shape the newspaper landscape in many countries. For instance, in Germany, about 34% of households consume regional news media weekly (Newman et al., 2019). The present paper links this regional dimension of the news media to that of innovation (diffusion) processes, by exploring whether there is a significant variation in the frequency and sentiments of innovation news at the subnational level. Thus, the paper not only addresses an issue hardly addressed until now, but it also brings together two literature streams (the geography of media and the geography of innovation studies) that are (still) relatively loosely connected.

To address this research gap, our empirical study employs a novel data set on national and regional news in Germany, recently established *RegNeS* database. This database covers headlines and snippets of more than 300 news sources from the German-speaking world. Within this data, we identified innovation news, by employing string matching algorithms and topic modeling. Their sentiments are quantified using the polarity classification tool developed by Rauh (2018). The data has further been enriched by information on newspapers' spatial circulation, which allows for approximating where news articles are most likely read.

27

Our results highlight that newspapers circulating in urban areas are more likely to feature news on innovation, implying more frequent exposure of readers to such information. We also identify notable differences between East and West Germany; news on innovation appears more frequently in newspapers in regions located in the territory of the former GDR. Some weaker evidence of a difference between the two parts of the country are also observed with respect to the sentiments with which such news are presented. On average, AI- and automation-related news tends to be written in a more positive tone in East, than in West Germany. Given that the reunification of the country did take place more than 30 years ago, these are somewhat surprising findings, which clearly call for more research in the future. Our study also hints at that the frequency and sentiments of innovation news are negatively associated with the levels of regional unemployment. Newspapers circulating in regions with less favorable labor markets appear to feature fewer articles on this topic, and if they do feature them, it has a relatively less-positive tone. A potential explanation for this finding is that many contemporary technologies (e.g., AI and automation) represent further severe threats (Frey and Osborne, 2017; Acemoglu and Restrepo, 2019) to the regions that are already economically weaker. This may lead to less-positive sentiments and generally less interest in these types of technologies, which carry into the editorial rooms of regional newspapers.

Our analyses show that newspaper coverage of innovation related events and discussions systematically varies across regions. This is likely to be explained by what is frequently referred to as *reverse agenda setting theory* (Watt Jr and Van Den Berg, 1978), i.e., prominence of news stories about innovation is driven by the already existing interest on the issue by the regional newspapers' readership. In turn, this relates to the regional factors considered. Accordingly, our study adds to prior research findings that the characteristics of regional news markets affect the news (Althaus et al., 2009) and that regional news media's tone reflects existing beliefs and attitudes of their consumers (Gentzkow and Shapiro, 2006). We show that regional socio-economic characteristics explain the available information set about certain topics in each region. For instance, relatively more-negative (or less-positive) sentiment regarding innovation news coverage in regions with higher unemployment rates, demonstrates this alignment of news media's attitude and regional macroeconomic indicators. It suggests that, in particular, the *relevance* and *proximity* effect drive the alignment between news and regional characteristics.

Yet, our study does not deliver causal empirical evidence. Theoretically, processes such as newspapers' having the power of agenda-setting (McCombs and Shaw, 1972) may fuel a reverse relationship. In addition, more omitted variables (e.g., newspaper ownership) might play a role in this setting. Nevertheless, given the rather fundamental and time-invariant nature of the considered regional characteristics, as well as the usually unemotional and politically cold discussions surrounding most innovations, there is hardly any support for such interpretation.

Our findings appear against the backdrop of some empirical limitations. Our data exclusively feature regional and national newspapers, ignoring other news outlets that might be of even greater relevance, such as tech-magazines, radio shows, and TV programs. We also concentrate on a single country (Germany). The employed data is limited to only headlines and text snippets. Add to this its cross-sectional nature, implying that specific events and short-term trends may affect some of

our findings. Collecting long-term news information and having access to full-text news archives are the obvious solutions to these limitations, for future research.

Admittedly, our study is also limited in its ability to disentangle and identify the mechanisms underlying our findings. However, it clearly confirms the existence of substantial and systematic heterogeneity in the presentation of innovation and new technologies at the subnational (regional) level. It also shows that this heterogeneity is related to the fundamental regional socioeconomic characteristics, implying that the news media does not stand apart from the general spatial organization of countries. In this sense, it raises the question of whether and to what extent the general path-dependent development of regions shapes the news and may contribute to it. Consequently, regarding the specific set of news in the focus of the present paper, looking at news at the subnational level appears to be a fruitful avenue for gaining a better understanding of technologies' spatial diffusion, in general, and for the development of regional innovation systems, in particular.

Keeping the limitations of the study in mind, our study suggests some implications. For instance, our observation of innovation news being relatively more frequent and positive in urban areas might contribute to the easier adoption and diffusion of new technologies in such places, in the long run. Consequently, this difference might be one factor fueling the increasing concentration of innovation activities in urban areas (Balland et al., 2020). Similarly, more negative news on innovation in regions with weaker economic developments, e.g., higher unemployment, may severely reduce public expectations and sentiments in this context. It can reinforce a non-supportive or even absent innovative culture, which further lowers local aspirations to positively engage with technological change and entrepreneurial activities (on the latter aspect, see von Bloh et al., 2019). In this case, the discussion of the local news media may occur in the context of regional innovation policies. However, this requires a better understanding of the geographical dimension of the news media and how this shapes and is shaped by other socioeconomic spatial structures.

References

Acemoglu, D. and P. Restrepo (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review 108*(6), 1488–1542.

Acemoglu, D. and P. Restrepo (2019). Automation and new tasks: how technology displaces and reinstates labor. *Journal of Economic Perspectives 33*(2), 3–30.

Acemoglu, D. and P. Restrepo (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy 128*(6), 2188–2244.

Aghion, P., B. F. Jones, and C. I. Jones (2017). Artificial intelligence and economic growth. Technical report, National Bureau of Economic Research.

Agrawal, A., J. S. Gans, and A. Goldfarb (2019). Exploring the impact of artificial intelligence: Prediction versus judgment. *Information Economics and Policy 47*, 1–6.

Aldrich, H. E. and C. M. Fiol (1994). Fools rush in? The institutional context of industry creation. *Academy of Management Review 19*(4), 645–670.

Alkemade, F. and R. A. Suurs (2012). Patterns of expectations for emerging sustainable technologies. *Technological Forecasting and Social Change 79*(3), 448–456.

Allern, S. (2002). Journalistic and commercial news values: News organizations as patrons of an institution and market actors. *Nordicom Review 23*(1-2), 137–152.

Althaus, S. L., A. M. Cizmar, and J. G. Gimpel (2009). Media supply, audience demand, and the geography of news consumption in the united states. *Political Communication 26*(3), 249–277.

Anselin, L. (2013). *Spatial econometrics: methods and models*, Volume 4. Springer Science & Business Media.

Anselin, L., A. K. Bera, R. Florax, and M. J. Yoon (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics 26*(1), 77–104.

Arun, R., V. Suresh, C. V. Madhavan, and M. N. Murthy (2010). On finding the natural number of topics with latent dirichlet allocation: Some observations. In *Pacific-Asia conference on knowledge discovery and data mining*, pp. 391–402. Springer.

Bakker, S. (2010). The car industry and the blow-out of the hydrogen hype. *Energy Policy 38*(11), 6540–6544.

Balland, P.-A., C. Jara-Figueroa, S. G. Petralia, M. P. Steijn, D. L. Rigby, and C. A. Hidalgo (2020). Complex economic activities concentrate in large cities. *Nature Human Behaviour 4*(3), 248–254.

Bathelt, H., A. Malmberg, and P. Maskell (2004). Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography 28*(1), 31–56.

Bednarek, M. and H. Caple (2017). *The discourse of news values: How news organizations create newsworthiness.* Oxford University Press.

Bednarz, M. and T. Broekel (2020). Pulled or pushed? The spatial diffusion of wind energy between local demand and supply. *Industrial and Corporate Change*.

Bergek, A., S. Jacobsson, and B. A. Sandén (2008). Legitimation and development of positive externalities: two key processes in the formation phase of technological innovation systems. *Technology Analysis & Strategic Management 20*(5), 575–592.

Berkhout, F. (2006). Normative expectations in systems innovation. *Technology Analysis & Strategic Management 18*(3-4), 299–311.

Biggi, G. and E. Giuliani (2020). The noxious consequences of innovation: what do we know? *Industry and Innovation*, 1–23.

Bivand, R. and G. Piras (2015). Comparing implementations of estimation methods for spatial econometrics. American Statistical Association.

Blasini, B., R. J. Dang, T. Minshall, and L. Mortara (2013). The role of communicators in innovation clusters. In *Strategy and Communication for Innovation*, pp. 119–137. Springer.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent dirichlet allocation. *Journal of machine Learning research 3*(Jan), 993–1022.

Bogart, L. (1989). *Press and public: Who reads what, when, where, and why in American newspapers.* Psychology Press.

Borup, M., N. Brown, K. Konrad, and H. Van Lente (2006). The sociology of expectations in science and technology. *Technology Analysis & Strategic Management 18*(3-4), 285–298.

Boukes, M. and R. Vliegenthart (2020). A general pattern in the construction of economic newsworthiness? Analyzing news factors in popular, quality, regional, and financial newspapers. *Journalism 21*(2), 279–300.

Boykoff, M. T. (2009). We speak for the trees: Media reporting on the environment. *Annual review of Environment and Resources 34*, 431–457.

Brenner, T. and T. Broekel (2011). Methodological issues in measuring innovation performance of spatial units. *Industry and Innovation 18*(1), 7–37.

Broekel, T. and M. Binder (2007). The regional dimension of knowledge transfers - a behavioral approach. *Industry and Innovation 14*(2), 151–175.

Broekel, T. and T. Brenner (2011). Regional factors and innovativeness: an empirical analysis of four german industries. *The Annals of Regional Science 47*(1), 169–194.

Buarque, B. S., R. B. Davies, D. F. Kogler, and R. M. Hynes (2019). Ok computer: The creation and integration of AI in europe. Technical report.

Budde, B., F. Alkemade, and K. M. Weber (2012). Expectations as a key to understanding actor strategies in the field of fuel cell and hydrogen vehicles. *Technological Forecasting and Social Change 79*(6), 1072–1083.

Cao, J., T. Xia, J. Li, Y. Zhang, and S. Tang (2009). A density-based method for adaptive lda model selection. *Neurocomputing 72*(7-9), 1775–1781.

Caple, H. and M. Bednarek (2016). Rethinking news values: What a discursive approach can tell us about the construction of news discourse and news photography. *Journalism 17*(4), 435–455.

Caulfield, T. (2004). Popular media, biotechnology, and the cycle of hype. *Houston Journal of Health, Law & Policy 5*, 213.

Cooke, P., M. G. Uranga, and G. Etxebarria (1997). Regional innovation systems: Institutional and organisational dimensions. *Research Policy 26*(4-5), 475–491.

Deuten, J., A. Rip, and J. Jelsma (1997). Societal embedding and product creation management. *Technology Analysis & Strategic Management 9*(2), 131–148.

Deveaud, R., E. SanJuan, and P. Bellot (2014). Accurate and effective latent concept modeling for ad hoc information retrieval. *Document Numérique 17*(1), 61–84.

Dudo, A., S. Dunwoody, and D. A. Scheufele (2011). The emergence of nano news: Tracking thematic trends and changes in US newspaper coverage of nanotechnology. *Journalism & Mass Communication Quarterly 88*(1), 55–75.

Eames, M., W. Mcdowall, M. Hodson, and S. Marvin (2006). Negotiating contested visions and place-specific expectations of the hydrogen economy. *Technology Analysis & Strategic Management 18*(3-4), 361–374.

Eilders, C. (2006). News factors and news decisions. theoretical and methodological advances in germany. *Communications 31*(1), 5–24.

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication 43*(4), 51–58.

Ewart, J. (2000). Capturing the heart of the region: How regional media define a community. *Transformations 1*(1), pp–1.

Fast, E. and E. Horvitz (2017). Long-term trends in the public perception of artificial intelligence. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Feldman, M. P. and D. B. Audretsch (1999). Innovation in cities:: Science-based diversity, specialization and localized competition. *European Economic Review 43*(2), 409–429.

Fenn, J. and M. Raskino (2008). *Mastering the hype cycle: how to choose the right innovation at the right time.* Harvard Business Press.

Frank, M. R., L. Sun, M. Cebrian, H. Youn, and I. Rahwan (2018). Small cities face greater impact from automation. *Journal of The Royal Society Interface 15*(139), 20170946.

Frey, C. B. and M. A. Osborne (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change 114*, 254–280.

Furman, J. and R. Seamans (2019). AI and the economy. *Innovation Policy and the Economy 19*(1), 161–191.

Galtung, J. and M. H. Ruge (1965). The structure of foreign news: The presentation of the congo, cuba and cyprus crises in four norwegian newspapers. *Journal of Peace Research 2*(1), 64–90.

Gamson, W. A. and A. Modigliani (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *American Journal of Sociology 95*(1), 1–37.

Gaskell, G., M. W. Bauer, J. Durant, and N. C. Allum (1999). Worlds apart? the reception of genetically modified foods in europe and the us. *Science 285*(5426), 384–387.

Geels, F. W. (2002). Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. *Research Policy 31*(8-9), 1257–1274.

Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. *Research Policy 33*(6-7), 897–920.

Geels, F. W. and B. Verhees (2011). Cultural legitimacy and framing struggles in innovation journeys: a cultural-performative perspective and a case study of dutch nuclear energy (1945–1986). *Technological Forecasting and Social Change 78*(6), 910–930.

Gensicke, T. (1995). Pragmatisch und optimistisch. In *Ostdeutschland im Wandel: Lebensverhältnissepolitische Einstellungen*, pp. 127–154. Springer.

Gentzkow, M. and J. M. Shapiro (2006). Media bias and reputation. *Journal of political Economy 114*(2), 280–316.

Gentzkow, M. and J. M. Shapiro (2010). What drives media slant? Evidence from us daily newspapers. *Econometrica 78*(1), 35–71.

Gibbons, M. (1999). Science's new social contract with society. *Nature 402*.

Glaser, S. (2017). A review of spatial econometric models for count data. Technical report, Hohenheim Discussion Papers in Business, Economics and Social Sciences.

Goldfarb, A. and D. Trefler (2018). AI and international trade. Technical report, National Bureau of Economic Research.

Gomez-Lievano, A., O. Patterson-Lomba, and R. Hausmann (2018). Explaining the prevalence, scaling and variance of urban phenomena. *Nature Energy*, 1–9.

González-Romá, V. and A. Hernández (2016). Uncovering the dark side of innovation: The influence of the number of innovations on work teams satisfaction and performance. *European journal of work and organizational psychology 25*(4), 570–582.

Griffiths, T. L. and M. Steyvers (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences 101*(suppl 1), 5228–5235.

Grün, B. and K. Hornik (2011). topicmodels: An R package for fitting topic models. *Journal of Statistical Software 40*(13), 1–30.

Haller, M. (2012). *Lokaljournalismus in den neuen Bundeslndern.* `http://www.bpb.de/gesellschaft/medien-und-sport/lokaljournalismus/151237`.

Hamilton, J. (2004). *All the news that's fit to sell: How the market transforms information into news.* Princeton University Press.

Harcup, T. and D. Oneill (2017). What is news? News values revisited (again). *Journalism Studies 18*(12), 1470–1488.

Hester, J. B. and R. Gibson (2003). The economy and second-level agenda setting: A time-series analysis of economic news and public opinion about the economy. *Journalism & Mass Communication Quarterly 80*(1), 73–90.

Humprecht, E. and F. Esser (2018). Diversity in online news: On the importance of ownership types and media system types. *Journalism Studies 19*(12), 1825–1847.

Hutchins, B. (2004). Castells, regional news media and the information age. *Continuum 18*(4), 577–590.

Iammarino, S., A. Rodríguez-Pose, and M. Storper (2019). Regional inequality in europe: evidence, theory and policy implications. *Journal of Economic Geography 19*(2), 273–298.

Inhoffen, L. (2018). *Künstliche Intelligenz: Deutsche sehen eher die Risiken als den Nutzen.* `https://yougov.de/news/2018/09/11/kunstliche-intelligenz-deutsche-sehen-eher-die-ris/`.

Jelodar, H., Y. Wang, C. Yuan, X. Feng, X. Jiang, Y. Li, and L. Zhao (2019). Latent dirichlet allocation (lda) and topic modeling: Models, applications, a survey. *Multimedia Tools and Applications 78*(11), 15169–15211.

Kepplinger, H. M. and S. C. Ehmig (2006). Predicting news decisions. an empirical test of the two-component theory of news selection. *Communications 31*(1), 25–43.

Klapper, J. T. (1960). The effects of mass communication.

Konrad, K. (2006). Shifting but forceful expectations: structuring through the prospect of materialisation. In *Twente VII workshop Material Narratives–of Technology in Society, Enschede.*

Korinek, A. and J. E. Stiglitz (2017). Artificial intelligence and its implications for income distribution and unemployment. Technical report, National Bureau of Economic Research.

LeSage, J. and R. K. Pacey (2009). *Introduction to Spatial Econometrics.* Chapman and Hall/CRC.

Lippmann, W. (1922). *Public opinion.* New York: Macmillan.

Makridakis, S. (2017). The forthcoming artificial intelligence (AI) revolution: Its impact on society and firms. *Futures 90*, 46–60.

Mangold, F., J. Vogelgesang, and M. Scharkow (2017). Nachrichtennutzung in deutschland. eine nutzerzentrierte repertoireanalyse. *M&K Medien & Kommunikationswissenschaft 65*(4), 704–723.

Marks, L. A., N. Kalaitzandonakes, L. Wilkins, and L. Zakharova (2007). Mass media framing of biotechnology news. *Public Understanding of Science 16*(2), 183–203.

Mast, C., S. Huck, and A. Zerfass (2005). Innovation communication. *Innovation Journalism 2*(4), 165.

Mazur, A. (2006). Risk perception and news coverage across nations. *Risk Management 8*(3), 149–174.

McCombs, M. E. and D. L. Shaw (1972). The agenda-setting function of mass media. *Public opinion quarterly 36*(2), 176–187.

McCormick, K. (2010). Communicating bioenergy: A growing challenge. *Biofuels, Bioproducts and Biorefining 4*(5), 494–502.

Mejía, C. and Y. Kajikawa (2019). Technology news and their linkage to production of knowledge in robotics research. *Technological Forecasting and Social Change 143*, 114–124.

Melton, N., J. Axsen, and D. Sperling (2016). Moving beyond alternative fuel hype to decarbonize transportation. *Nature Energy 1*(3), 16013.

Negro, S. O., F. Alkemade, and M. P. Hekkert (2012). Why does renewable energy diffuse so slowly? A review of innovation system problems. *Renewable and Sustainable Energy Eeviews 16*(6), 3836–3846.

Newman, N., R. Fletcher, A. Kalogeropoulos, and R. Nielsen (2019). *Reuters institute digital news report 2019*, Volume 2019. Reuters Institute for the Study of Journalism.

Nordfors, D. (2004). The role of journalism in innovation systems. *Innovation Journalism 1*(7), 1–18.

Petersen, A. (2001). Biofantasies: genetics and medicine in the print news media. *Social Science & Medicine 52*(8), 1255–1268.

Price, V. and J. Zaller (1993). Who gets the news? alternative measures of news reception and their implications for research. *Public opinion quarterly 57*(2), 133–164.

Priest, S. H. (1994). Structuring public debate on biotechnology: Media frames and public response. *Science Communication 16*(2), 166–179.

Rauh, C. (2018). Validating a sentiment dictionary for german political languagea workbench note. *Journal of Information Technology & Politics 15*(4), 319–343.

Remus, R., U. Quasthoff, and G. Heyer (2010). Sentiws-a publicly available german-language resource for sentiment analysis. In *LREC*. Citeseer.

Ruef, A. and J. Markard (2006). What happens after a hype. In *Changing expectations and their effect on innovation activities, EASST Conference*, pp. 23–26.

Schot, J. and F. W. Geels (2008). Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy. *Technology Analysis & Strategic Management 20*(5), 537–554.

Shoemaker, P. J., J. H. Lee, G. Han, and A. A. Cohen (2007). Proximity and scope as news values. *Media Studies: Key Issues and Debates*, 231–248.

Skjølsvold, T. M. (2012). Curb your enthusiasm: On media communication of bioenergy and the role of the news media in technology diffusion. *Environmental Communication: A Journal of Nature and Culture 6*(4), 512–531.

Staab, J. F. (1990). The role of news factors in news selection: A theoretical reconsideration. *European Journal of Communication 5*(4), 423–443.

Stephens, J. C., G. M. Rand, and L. L. Melnick (2009). Wind energy in US media: A comparative state-level analysis of a critical climate change mitigation technology. *Environmental Communication 3*(2), 168–190.

Stone, P., R. Brooks, E. Brynjolfsson, R. Calo, O. Etzioni, G. Hager, J. Hirschberg, S. Kalyanakrishnan, E. Kamar, S. Kraus, et al. (2016). Artificial intelligence and life in 2030. One hundred year study on artificial intelligence: Report of the 2015-2016 study panel. *Stanford University, Stanford, CA. 6*.

Van Lente, H. (1995). Promising technology: The dynamics of expectations in technological developments.

Van Lente, H., C. Spitters, and A. Peine (2013). Comparing technological hype cycles: Towards a theory. *Technological Forecasting and Social Change 80*(8), 1615–1628.

von Bloh, J., T. Broekel, B. Ozgun, and R. Sternberg (2019). New (s) data for entrepreneurship research? An innovative approach to use big data on media coverage. *Small Business Economics*, 1–22.

Waltinger, U. (2010). Germanpolarityclues: A lexical resource for german sentiment analysis. In *LREC*, pp. 1638–1642.

Watt Jr, J. H. and S. A. Van Den Berg (1978). Time series analysis of alternative media effects theories. *Annals of the International Communication Association 2*(1), 215–224.

Webster, J. G. and L. Lichty (1991). Ratings analysis: Theory and practice.

Zucker, H. G. (1978). The variable nature of news media influence. *Annals of the International Communication Association 2*(1), 225–240.

Appendix A   Topic modeling

To find the optimal number of topics to identify by the LDA, we relied on the metrics developed by Griffiths and Steyvers (2004); Cao et al. (2009); Deveaud et al. (2014) and Arun et al. (2010).[13] In our case, they suggest to consider between 50 and 70 topics. Qualitative assessments of the results suggest 60 topics to deliver the most coherent and meaningful groupings. On this basis, the LDA parameters were estimated using Gibbs sampling (Grün and Hornik, 2011). The outcome of the topic modeling is that each document is assigned to a topic, and each topic is described by terms with varying probabilities. We relied on qualitative (manual) assessment to identify those topics most likely to relate to innovation and new technologies. Table A1 shows a section of the obtained term-topic matrix. The red-colored topics are examples of the ones we considered to be of relevant in the context of the paper. In total, of the 60 topics, we classified 50 as being related to innovation and new technologies.

Table A1: The most common 5 terms within each topic

1	fahrzeug	technologi	eautos	elektromobilitat	zukunft
2	klimawandel	weltweit	thunberg	global	klimaschutz
3	unternehm	impfstoff	biontech	curevac	corona
4	digital	automatisier	technologi	zukunft	wirtschaft
5	spiel	sport	fussball	team	saison
...	...	...	...	...	...
56	autonom	fahr	bus	fahrend	automatisiert
57	cdu	spd	bundesregier	grun	polit
58	startup	unternehm	innovation	wirtschaft	investor
59	robot	international	iss	humanoid	raumstation
60	innovation	kunstlich	intelligenz	technologiezentrum	eroffnet

---

[13]Implemented in the *ldatuning* R-package.

Appendix B   Correlation table

Correlation of the variables used in the regressions appears in Table B1.

Table B1: Correlation matrix

	POPDEN	GDPC	UNEMP	EAST	NNP	NNEWS	INNOV	SHARE	NWORD	NATIONAL	SENT
GDPC	0.35***										
UNEMP	0.37***	-0.38***									
EAST	0.07*	-0.46***	0.46***								
NNP	0.58***	0.51***	0.17***	-0.28***							
NNEWS	-0.02	0.00	-0.05*	-0.03	-0.07**						
INNOV	-0.03	-0.03	-0.03	0.02	-0.09***	0.29***					
SHARE	-0.03	-0.03	0.00	0.00	-0.04	-0.06*	0.81***				
NWORD	-0.01	0.00	-0.01	-0.06**	0.04	-0.17***	0.10***	0.50***			
NATIONAL	-0.10***	-0.11***	-0.01	0.08**	-0.20***	0.18***	0.34***	0.23***	-0.21***		
SENT	-0.02	0.03	-0.08**	0.00	-0.04	-0.10***	-0.32***	-0.28***	-0.10***	-0.19***	
NPSENT	-0.03	0.04	-0.10***	-0.08**	0.02	-0.01	0.01	0.05	0.16***	-0.40***	0.47***

*p<0.05; **p<0.01; ***p<0.001

Appendix C   Lagrange multiplier diagnostics for spatial dependence

Table C1 shows the LM test results for each regression analysis. Of all models, the spatial error model seems to be the most appropriate.

Table C1: LM test results

	$\mathrm{RS}_{n,r}$	RLMerr		RLMlag	
		test-statistic	p-value	test-statistic	p-value
log(INNOV)					
	> 0	87.737	0.000	0.054	0.815
	> 0.001	71.176	0.000	0.759	0.383
	> 0.005	70.866	0.000	0.083	0.773
	> 0.01	58.926	0.000	0.584	0.445
log(SENT)					
	> 0	22.118	0.000	0.153	0.695
	> 0.001	23.627	0.000	0.063	0.802
	> 0.005	27.465	0.000	1.293	0.256
	> 0.01	3.073	0.008	3.506	0.061
AI log(SENT)					
	> 0	12.771	0.001	0.802	0.371
	> 0.001	6.213	0.013	0.488	0.485
	> 0.005	8.789	0.003	0.032	0.858
	> 0.01	12.723	0.000	0.428	0.513
Automation log(SENT)					
	> 0	14.711	0.000	0.024	0.869
	> 0.001	5.885	0.015	0.006	0.938
	> 0.005	4.114	0.042	0.078	0.779
	> 0.01	14.771	0.000	4.850	0.076