WORKING PAPERS

N° 1693

November 2025

"The Police as Gatekeepers of Information: Immigration Salience and Selective Crime Reporting"

Ashrakat Elshehawy, Arun Frey, Violeta I. Haas,
Sascha Riaz and Tobias Roemer

Toulouse School of Economics

The Police as Gatekeepers of Information: Immigration Salience and Selective Crime Reporting*

Ashrakat Elshehawy[1] Arun Frey[2] Violeta I. Haas[3]
Sascha Riaz[4], and Tobias Roemer[5]

[1] *Assistant Professor, University College London*
[2] *Independent Researcher*
[3] *Postdoctoral Fellow, Institute for Advanced Study in Toulouse, Toulouse School of Economics*
[4] *Assistant Professor, Department of Political and Social Sciences, European University Institute*
[5] *PhD Candidate, Nuffield College, University of Oxford*

November 11, 2025

Abstract

What drives the supply of crime news? While prior research focuses on news media reporting, we study a crucial upstream gatekeeper of information: the police. We argue that the police act as strategic bureaucrats who increase the disclosure of out-group cues (ethnicity, nationality) when immigration is salient to signal transparency to the public. To test this, we use an LLM annotation pipeline to code out-group cues in about one million press releases published by local police stations across Germany (2014–2024). Using a regression discontinuity in time design, we demonstrate an increase in out-group cues in police communications (1) following a nationwide shock to immigration salience (the 2015/16 Cologne New Year's Eve sexual assaults), and (2) in the days before regional state elections in which immigration is a salient campaign issue. Our findings demonstrate how bureaucratic discretion shapes the supply of politically charged information.

*We thank Ran Abramitzky (& group), Alexander de Juan, Elias Dinas, Dan Jurafsky, Hanno Hilbig, Federico Nanni, Nini Petriashvili, Thomas Tichelbaecker, Elisa Maria Wirsching, Delia Zollinger, and audiences at the European University Institute, EPSA 2025 LLM Pre-Workshop, Institute for Advanced Study in Toulouse, and University of Mannheim for helpful comments on this project. We thank Thomas Hestermann for sharing data. We also thank Luca Schenk and Greta Plattner for excellent research assistance. We gratefully acknowledge financial support from Nuffield College, University of Oxford. Violeta I. Haas further acknowledges support from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

1 Introduction

Crime is one of the most popular and widely consumed types of news: related stories consistently rank among the most-read articles on national news websites and feature prominently on television and social media. In surveys, crime emerges as the second most widely reaching news topic among the public, surpassed only by weather (Pew Research Center 2024*a*). Across Western democracies, crime is closely linked with immigration in the public mind (Fitzgerald, Curtis and Corliss 2012). The belief that immigration increases crime is widespread and fuels anti-immigrant sentiment (Pew Research Center 2024*b*; YouGov 2019, 2025). Accordingly, crime reporting does more than just inform; it shapes public opinion in politically consequential ways. Prior research shows that media coverage of crime involving ethnic-minority suspects can have far-reaching consequences – raising electoral support for the radical right (Giavazzi et al. 2024; Alizade 2025; Sabet, Liebald and Friebel 2025) and, in some cases, even precipitating hate crimes (Frey 2020; Riaz, Bischof and Wagner 2024).

Who decides which crimes become news in the first place? We argue that previous research has largely overlooked a crucial gatekeeper in this context: the police. While most prior work treats the news media as the key agent in selecting and framing crime stories (e.g., Couttenier et al. 2021; Keita, Renault and Valette 2024; Berk 2025), it is often a causally upstream actor – the police – that makes similarly decisive choices. Police press offices exercise discretion on two levels: first, they select which incidents to publicize from the vast number of recorded offenses; second, they also decide what information to include or exclude in police communications, such as the characteristics of a suspect.[1] This dual selection grants police substantial power to shape the supply of politically charged information long before it reaches a newsroom.

This paper empirically examines how this discretion is exercised. Specifically, we study the disclosure of *out-group cues* – explicit or implicit references to a suspect's foreign na-

[1]Based on our data, we estimate that in the German context, less than 2% of recorded offenses are ever publicized in a press release.

tionality or non-German ethnicity – in police press releases. Out-group cues are political because they speak directly to salient public anxieties about immigration and crime and are readily amplified by local news outlets that draw on information shared by the police in their reporting. We argue that the disclosure of suspect ethnicity is not a mechanical byproduct of case facts or neutral press rules; it is a strategic communication choice that varies with political salience.

Building on classic insights about bureaucratic politics and reputation (Carpenter 2001, 2010; Carpenter and Krause 2012), we argue that local police agencies act as strategic bureaucrats who leverage their near-monopoly on crime information to manage public legitimacy. Press officers face a reputational trade-off: disclosing nationality/ethnicity risks charges of ethnicization and bias; withholding it risks accusations of opacity or cover-ups. We argue that, under heightened immigration salience, police *dial up* the disclosure of out-group cues to signal transparency to the public. Disclosing a suspect's foreign nationality is a low-cost action that allows the police to signal transparency and responsiveness.

We study the case of Germany, which provides an ideal setting to study this supply-side channel. Policing is decentralized to the sixteen states; small press offices embedded in local police stations screen incident logs, decide which events are "press-relevant," and select which attributes – such as nationality or ethnicity – to include. Professional and internal codes permit disclosure of nationality when there is a "justified public interest," creating a large degree of bureaucratic discretion. Local news outlets frequently republish police texts verbatim, replicating upstream choices by police press offices.[2]

We test our argument using a novel corpus of approximately one million police press releases issued by local police stations across Germany over a ten-year period (2014–2024). The task of extracting structured information from an unstructured textual dataset of this

[2]Examples include the "Blaulicht/Polizei" sections of: Rhein-Zeitung; Donaukurier; Mittelbayerische Zeitung; WAZ (NRW); Kieler Nachrichten (Polizeiticker Kiel); Neue Westfälische; and Ostthüringer Zeitung.

scale raises substantial challenges. As the corpus consists of texts written by hundreds of different individuals, it exhibits substantial variation in style, vocabulary, and rhetorical strategies, further complicating the use of standard NLP tools that rely on consistent lexical patterns. We address these challenges by employing recent advances in large language models (LLMs). To extract information from this unstructured text, including the subtle cues that indicate the out-group status of suspects, we use a LLM (GPT-4o) after extensive prompt optimization. This approach helps us move beyond widely used Natural Language Processing methods, such as keyword and dictionary-based approaches, to be able to capture implicit and contextual cues that conventional text analysis methods would miss. We employ multiple validation strategies to systematically assess the reliability of the LLM's output. We rigorously validated our approach against human coders and achieved a high level of classification performance (F1 score ≈ 0.9).

Our empirical strategy leverages two sources of variation in immigration salience. First, we analyze a nationwide focusing event: the 2015/16 Cologne New Year's Eve assaults. Following this shock in immigration salience, we document a sharp, discontinuous rise in out-group cue disclosure that persists for roughly a year. Second, we examine predictable cycles of immigration salience around state (*Landtag*) elections using a regression discontinuity in time design. In the days immediately preceding an election, the share of press releases containing out-group cues increases by about three percentage points (32% relative to the mean). This effect is strongest in elections where immigration is a particularly salient campaign issue, as measured by pre-election survey data. Taken together, these findings from two distinct settings provide evidence for our core argument: police strategically increase the disclosure of out-group cues when the political salience of immigration is high.

Our study contributes to multiple lines of research. First, we introduce a supply-side perspective to the literature on media and politics. While most research examines the downstream effects of crime news coverage on public opinion and behavior (e.g., Couttenier et al. 2021; Keita, Renault and Valette 2024; Berk 2025), we analyze a crucial upstream actor:

the police, whose discretionary choices about what crime information to disclose shape the supply of politically salient information that newsrooms and the public receive.[3]

Second, we contribute to the study of bureaucratic discretion and its role in producing biased outcomes (Alizade, Dancygier and Homola 2024; Wirsching 2025). Prior work has shown that discretion allows legal decisions to be responsive to the political environment, amplifying racial biases and shifting with public discourse (Shayo and Zussman 2011; Riaz and Hamjediers 2025). We demonstrate an analogous dynamic in police communication, where press offices use their discretion to adjust the disclosure of out-group cues as immigration becomes politically salient. In both settings, discretion renders bureaucratic decisions responsive to shifts in the political climate.

Third, our paper contributes to a large and growing literature on racial bias in policing. While previous research has documented how bias materializes in police-civilian interactions (Voigt et al. 2017; Hoekstra and Sloan 2022; Rho et al. 2023; Xu et al. 2024), we show that it can also materialize in reporting, especially during periods of increased public scrutiny (see also Grunwald, Nyarko and Rappaport 2022).

Finally, we complement recent work on the politics of police transparency, in particular Cook and Fortunato (2023), who show that higher legislative capacity and oversight can compel police agencies to increase transparency. Our argument highlights a functional substitute: public pressure during periods of high immigration salience. When scrutiny intensifies, the police face incentives to proactively project transparency. Crucially, however, the resulting transparency is biased: rather than uniformly expanding reporting, the police selectively amplify out-group cues in press releases. This supply-side response thus risks creating a feedback loop, where the selective ethnicizing of crime news further feeds into the very salience that drives the initial disclosure.

[3]To our knowledge, the only previous analysis of the disclosure of ethnic markers in police communications is that of Grunwald, Nyarko and Rappaport (2022), who studied the social media posting behavior of police agencies across the United States.

2 Bureaucratic Power and the Politics of Information

Following the literature on bureaucratic politics, we conceive of public agencies as strategic organizations with their own distinct interests, such as institutional autonomy, budgetary maximization, and public legitimacy, that can diverge from those of their political principals (Carpenter 2001, 2010; Carpenter and Krause 2012). This divergence complicates simple models of political control, as it implies that bureaucrats' behavior is guided not just by formal mandates but by a strategic need to manage their external political environment. The police, as one such street-level bureaucrat (Lipsky 1980), for instance, have been shown to manipulate operational functions, such as slowing emergency response times, to exert pressure on politicians who adopt policies they oppose (Wirsching 2025).[4] Beyond manipulating service provision, however, we theorize that a less costly instrument of bureaucratic power is the police's control over the narrative of public safety.

We argue that a key instrument in the police's management of their political environment is their position as the near-monopolist of crime information. While foundational work like that of Banks and Weingast (1992) models how political principals struggle to overcome this informational asymmetry, our argument focuses on how the agent proactively leverages this advantage to shape their audience's perceptions and control reputation. Control over the initial collection, interpretation, and dissemination of information about crime confers significant agenda-setting power. Through formal communication channels, police departments construct the primary narrative of public safety, thereby shaping the informational environment in which citizens form attitudes and behavior. We argue that this strategic communication is a crucial mechanism for signaling institutional legitimacy.

However, managing public perceptions of legitimacy is not merely an abstract reputational goal. It is an instrumental objective directly linked to the police's material interests.

[4]Similarly, Nowacki and Thompson (2021) present evidence that replacing appointed police authorities with directly elected commissioners made policing more responsive to local voter preferences, illustrating how electoral accountability can shape policing priorities.

By proactively communicating about a perceived threat, such as out-group crime, police can construct a public narrative that justifies requests for increased budgets, expanded staffing, or greater operational authority from their political principals. In this view, signaling legitimacy is functionally equivalent to demonstrating the necessity of police resources in tackling salient public safety challenges.

The incentive to signal legitimacy becomes particularly acute during periods of heightened immigration salience. Public discourse routinely links immigration to crime, even though empirical evidence shows little basis for this association (Marie and Pinotti 2024). Across countries, the public systematically overestimates the share of crimes committed by foreigners, the proportion of immigrants in prisons, and their involvement in gangs (Semyonov, Raijman and Gorodzeisky 2008; Stansfield and Stone 2018; Rumbaut and Ewing 2007; Moore-Berg, Hameiri and Bruneau 2022). Immigration is also frequently associated with terrorism, despite little evidence of a direct link (Skinner and Gottfried 2017; Helbling and Meierrieks 2022). These misperceptions are consequential. Majorities in the United States believe immigration increases crime (Pew Research Center 2024*b*), and across Western Europe large shares – e.g., Germany (81%), Spain (80%), Sweden (73%), Italy (71%), and the United Kingdom (71%) – judge immigration over the past decade to have been "too high," often citing crime among the main reasons (YouGov 2019, 2025). Populist and radical-right actors amplify this immigration-crime framing and mobilize around it (Alizade 2025; Schneider-Strawczynski and Valette 2025; Riaz and Roemer 2025). In this high-pressure environment, the police – as upstream gatekeepers of crime information – face intense accountability pressures.

We argue that this creates a complex trade-off for the police. On the one hand, disclosing suspect nationality maximally at all times risks violating internal guidelines and fueling accusations of racial profiling. On the other hand, non-disclosure, especially when public attention is high, invites charges of opacity and political cover-ups, making the police vulnerable to criticism from the political right and their supporters, threatening their

reputation. This trade-off was a recurring theme in our interviews with police press officers across multiple German states. We argue that local police stations navigate this dilemma strategically. Because they have discretion over what incidents to publicize and which attributes to disclose, they *dial up* out-group cue disclosure when immigration is especially salient, i.e., when the reputational returns from signaling transparency are highest.

The signal sent through this strategic disclosure is twofold. First, the very act of releasing information on a sensitive topic signals transparency, directly countering the damaging accusations of opacity and cover-ups that often arise during high-salience periods (see interview evidence described in Section 3.2). Second, the content of the disclosure, i.e., reporting on out-group crime, projects an image of proactive enforcement and control. While the public reception of this signal, often mediated through the news media, may be complex, the immediate goal from the bureaucrats' perspective is to demonstrate that they are actively responding to a perceived threat, thereby reinforcing their legitimacy.

This form of bottom-up, strategic transparency stands in stark contrast to the top-down, coerced transparency compelled by formal legislative oversight. While strong legislatures may force police agencies to report more accurately on their misconduct, such as the use of lethal force (Cook and Fortunato 2023), the public-pressure mechanism we identify creates a different set of incentives. Instead of being forced to reveal politically costly information about their own failures, police can strategically respond to public demand by increasing selective transparency about a different target: out-group crime.

Disclosing the foreign nationality of suspects offers a uniquely powerful signal in this context. It is a low-cost action, requiring only a minor change in a press release, that can have a high political impact. Selective disclosure taps into pre-existing public anxieties and political narratives around immigration, allowing the police to frame a simple narrative with a clear "threat" (out-group crime) that they are visibly "solving." This projects an image of control and is likely to generate significant media attention, reinforcing the signal of transparency and legitimacy. Indeed, large-scale descriptive work on police communications in the United

States shows that law enforcement agencies systematically overreport on minority suspects in their social media posts relative to local arrest rates (Grunwald, Nyarko and Rappaport 2022). Furthermore, the public and political debate over whether to disclose suspects' nationality is a salient issue in many democracies (Reuters 2025; The Independent 2016; SWI 2021; The Guardian 2018), demonstrating that the decision to release this information is a politically charged act over which police can signal transparency.

This argument yields a key testable implication: police disclosure of out-group cues should rise and fall with the political salience of immigration. We test this expectation across two distinct settings. The first is a sudden, nationwide shock: the 2015/16 New Year's Eve assaults in Cologne. The sexual assaults and robberies committed by groups of men, widely reported to be of North African or Middle Eastern origin, ignited a national scandal over immigration, public safety, and police communication (Frey 2020). Such focusing events create an acute crisis, putting police on the defensive and generating immense pressure to demonstrate legitimacy and control reputation. The second context involves predictable, institutionalized cycles of immigration salience around regional state (*Landtag*) elections. We focus on state elections because political authority over the police is vested at the state level in Germany (see Section 3). During electoral campaigns, issues like crime and immigration are frequently at the forefront of political debate, attracting heightened attention from voters, the media, and politicians. In both scenarios, we argue, police face a powerful incentive to signal transparency to an attentive and concerned public.

3 Context

3.1 Organizational Structure of the German Police

Policing is highly decentralized in Germany. Primary authority rests with the sixteen federal states (*Länder*). Each state maintains its own police force (*Landespolizei*) governed by state-specific law and overseen by the state's Interior Ministry. The state interior minister supervises police operations within that jurisdiction.

At the state level, the state police handle the vast majority of policing duties and collectively employ the bulk of Germany's police personnel (Statistisches Bundesamt 2024).[5] Although each state operates under its own police law, creating sixteen distinct legal frameworks, these laws share many similarities. Traditionally, state forces are organized into a uniformed branch (*Schutzpolizei*) responsible for patrol, emergency response, and order maintenance, and a detective branch (*Kriminalpolizei* for investigative work, often centralized in a State Criminal Police Office).

Federal police agencies, in contrast, have more circumscribed jurisdictions. The largest is the federal police (*Bundespolizei*), subordinate to the Federal Interior Ministry. With approximately 41,000 total staff (including 33,000 police officers), they are responsible for border protection, security at airports and major railway stations, and the safeguarding of federal buildings and embassies. It also includes specialized units for counter-terrorism. Another key agency is the Federal Criminal Police Office (*Bundeskriminalamt*, BKA), a civilian investigative bureau. The BKA functions as Germany's central criminal investigative authority, coordinating complex or transnational cases and maintaining national crime databases. It conducts investigations into serious crimes that span multiple states or have international dimensions, in cooperation with the state police forces. Importantly, neither the Bundespolizei nor the BKA has direct command over the state police forces. Press releases by the federal police are not included in our main analyses (see section 4). We do, however, leverage press releases issued by the federal police for a placebo test (Table A.22).

3.2 Police Press Releases

The police communicate with the public through multiple channels, including social media, online portals for filing reports, and official press releases. While social media posts may serve rapid-response, promotional, and image management purposes (Hu, Rodgers and

[5] In 2023, around 300,000 full-time equivalent police personnel, of which around 80% belonged to state police (Statistisches Bundesamt 2024). Police personnel include police officers, but also, e.g., administrative personnel.

Lovrich 2018; O'Connor 2017; O'Connor and Zaidi 2020), press releases are the most institutionally vetted and detailed source of public information about police operations and crime.[6] Although press communication is decentralized, it follows a structurally similar process across jurisdictions. Drawing on extensive interviews with officers from three local police stations, we outline the common organizational structures, workflows, decision-making criteria, and communication routines that shape how and when crime is reported to the public below.

Press releases are issued by specialized public relations divisions within regional police headquarters. These units are responsible for external communication and often maintain direct contact with local and national media outlets. While they draw on information from field officers and investigators, press offices exercise substantial autonomy in determining whether and how incidents are communicated to the public. The composition of these units varies. The core staff typically consists of sworn police officers with field or investigative experience who have received specialized training in media communication. Many rotate into press roles as part of internal career development. In some jurisdictions, press teams also include specialists with backgrounds in journalism. In more politically sensitive or high-profile cases, department heads may be involved in reviewing press drafts before publication. This oversight, however, occurs on an ad hoc basis rather than as part of the standard workflow.[7]

Table 1 provides an overview of the standard workflow from crime event to the final press release. As shown, press releases are not a comprehensive record of criminal activity. They reflect a highly selective process. Across Germany, only a small fraction of all recorded crime ($\approx 2\%$ based on our data) is ever the subject of a public press release. This means that what the public and media perceive as "crime" is largely shaped by the discretionary choices

[6] Police social media updates routinely cross-post brief summaries and link back to the full press release.

[7] Because they are written by humans, press releases are capacity-constrained. German police personnel increased from 310,900 to 355,800 between 2014 and 2024, i.e., at a compound annual growth rate of about 1.36% (Statistisches Bundesamt (Destatis) 2025).

Stage	Description
1. Incident	Dispatched officers record details (event, witnesses, location, etc.).
2. Documentation	Officers enter data for each event into a centralized system.
3. Flagging	Local police stations review incident logs and flag high-profile events, incidents with investigative necessities, or public-awareness campaigns as "press-relevant."
4. Selection	Regional Public Relations Division receives flagged and unflagged cases and decides which incidents to release.
5. Drafting	A press officer prepares the release and selects the information to include, while considering ethical codes.
6. Review	Drafts are reviewed internally, typically by multiple staff members, sometimes by senior leadership, before approval.
7. Dissemination	Approved releases are published via the police website, Presseportal.de, and email to subscribed media. Social media is used selectively.
8. Media Interaction	Journalists may follow up with questions or requests for clarification. The press office may issue corrections or updates.

Table 1: Standard Workflow from Crime Incident to Press Release

of local police units. This selection is applied at two levels. First, police decide whether an incident merits a press release. Second, if a press release is issued, police decide which details to include, such as the nature of the offense and the suspect's characteristics. These choices are not mechanically determined by protocols but reflect professional judgment, media considerations, and organizational priorities.

At the first level of selection, police evaluate whether an incident should be made public. This decision is guided by three commonly referenced criteria identified during our interviews and in guidelines (Ministerium des Inneren, für Digitalisierung und Kommunen Baden-Württemberg 2022): (1) *public interest*, which considers media inquiries and the

visibility of the incident; (2) *investigative value*, which relates to calls for witnesses or clarifications of conflicting facts; and (3) *proactive communication*, which includes managing the narrative in high-profile cases, such as use of police force, or implementing crime prevention campaigns against recurring events, such as burglary series. Additionally, certain events, like suicides, are generally not reported.[8]

Once an incident is selected for reporting, the press office, in a second level of selection, decides which details to disclose. This includes not only the factual core of the communication, such as the nature, time, and location of the incident, but also whether to include characteristics of those involved. While ethics codes, both internal and more general journalistic ones, emphasize neutrality, truthfulness, and non-discrimination, they also allow for the inclusion of sensitive attributes such as nationality, "if there is a justified public interest or upon media inquiry. [...] Other than this, no mention shall be made if it is not relevant to the understanding of a case." (Ministerium des Inneren, für Digitalisierung und Kommunen Baden-Württemberg 2022, Richtlinie 6.2).

Thus, in principle, all suspects must be treated equally before the law, and press coverage should avoid reinforcing stereotypes or biases. However, because the threshold for "public interest" is not clearly defined, officers retain wide discretion. Our interviews revealed that decisions to disclose nationality are often shaped by informal heuristics, such as media expectations, political climate, and reputational concerns, rather than consistent application of formal rules. While officers uniformly emphasized "neutrality," they also acknowledged that reporting decisions are shaped by workload, media demand, and politics. One officer, for example, explained that during the heightened political focus on asylum seekers around 2019, his unit adopted a policy of "maximum transparency." They deliberately communicated incidents involving this group to preempt accusations from right-wing voices that they

[8]The "public interest" criterion is often presumptively satisfied for serious crimes, such as homicide and terrorism, which represent a very small share of all criminal offenses. For example, various forms of homicide (*Mord, Totschlag und Tötung auf Verlangen*) account for less than 0.1% of all crimes.

were concealing information about crimes committed by asylum seekers.

A similar logic was applied in response to geopolitical events. Another officer noted that following Russia's war in Ukraine, they became more likely to disclose the nationality in conflicts involving Russians or Ukrainians. He clarified this was not due to any rule change, but because the war had made the topic a matter of public concern, transforming a detail that would have been ignored five years ago into one the public now expected to see reported.

Taken together, our interviews and review of internal ethical codes show that police press work is governed by a combination of formal guidelines and informal practices, granting officers substantial leeway over both whether an incident is reported and how it is presented. This discretionary power also speaks against a simple media-demand explanation. As first movers, police hold an informational monopoly. Journalists generally cannot access crime scenes, suspect details, or investigation progress, and must rely on the information shared by the police. Our interviews confirm this proactive, supply-driven mindset: one officer explicitly described adopting "maximum transparency" on crimes involving asylum seekers to preempt accusations from the right, a strategic act of reputation management. Local media's structural constraints amplify this supply-side power. Lacking investigative resources, many outlets publish police press releases verbatim, often through direct feeds (see footnote 2). Thus, as causally upstream gatekeepers, the police not only determine which stories become part of the public record but also actively shape perception through the selective disclosure of suspect characteristics.

4 Data

Our dataset comprises about one million police press releases issued by local police stations across 15 German states between 2014 and 2024. These press releases represent the primary channel through which police departments in Germany communicate with the public and media about criminal incidents. To construct this dataset, we combined data from multiple sources. For the period 2020 to August 2024, we scraped all available police releases from

Presseportal.de, a centralized platform operated by Germany's largest newswire service (dpa) and used by nearly all state and federal police units to disseminate press releases. For the period 2014 to 2019, we rely on data collected and generously shared by Hestermann (2025). We rely on this dataset because *Presseportal* deletes older records after several years, making it impossible to retrieve a full archive of earlier press releases through direct scraping. The dataset provides press releases retrieved from *Presseportal* prior to 2020, preserving material that is no longer publicly available online today. For a few states – including Berlin, Brandenburg, and Saxony-Anhalt – coverage on *Presseportal* was initially incomplete or limited to federal agencies. In these cases, we supplemented the dataset with releases scraped directly from the official websites of the relevant state police whenever possible (see Appendix section A.1.1 for additional details).

To isolate a clean sample of single, crime-related events suitable for testing our theory of discretionary disclosure, we apply multiple filters to the raw data. We exclude traffic incidents and releases unrelated to criminal events (e.g., community outreach or safety campaigns). This filtering was achieved through headline-based keyword matching (e.g., removing headlines containing the word "Unfall") and LLM-based classification (Appendix A.1.2 and A.1.3). To ensure each unit of observation corresponds to a single criminal event, we also exclude releases that aggregate multiple incidents into a single report using our LLM-based classification (Appendix Sections A.1.3). We then restrict the sample to cases where our theoretical mechanism of bureaucratic discretion can plausibly operate. We therefore exclude press releases from North Rhine-Westphalia (after July 2024) and Mecklenburg-Vorpommern (after February 2020) from the analysis, because state-level directives in those jurisdictions required police to always disclose offender nationality in press releases, removing the discretion central to our design (Ministerium für Inneres und Bau Mecklenburg-Vorpommern 2020; Evangelische Zeitung 2024). We also exclude federal police releases, whose centralized jurisdiction (e.g., border, rail, airports) and reporting chains are not electorally accountable at the state level and thus not comparable to state police in our election-based analyses. The

final dataset consists of about one million (972,231) press releases from 324 unique police stations spanning over a decade (01/2014 – 08/2024).

SI Figure A.3 displays the publication frequency of police press releases across all police stations in our dataset, disaggregated by state and including federal police units. The completeness of coverage varies substantially across states. In some states – such as Bavaria – many local police stations are underrepresented on *Presseportal*, resulting in sparser coverage. Temporal coverage also differs: in several cases, police stations began uploading to *Presseportal* only partway through the study window (e.g., Saarland), or experienced temporary interruptions in publication. These differences reflect decentralized reporting practices and variation in institutional adoption of centralized digital platforms. These over-time changes do not impact the validity of our empirical estimates (see section 5), which rely on short-term temporal changes in reporting practices.

Each press release contains the date of publication, the issuing police station, and the full report text (an example is provided in Figure 1). The text is highly unstructured, as key information – such as suspect ethnicity or nationality – is not consistently formatted and can be conveyed through various explicit or implicit language cues.

To transform this unstructured corpus into a structured tabular dataset, we implemented an annotation pipeline using GPT-4o. We prompted the model to extract and classify information across several key dimensions: (see Table 2 for the variables extracted for the example press release shown in Figure 1).

- Type of crime (e.g., violent, property, traffic-related)
- Investigation status (ongoing or concluded)
- Presence of out-group cues (e.g., foreign nationality or ethnicity)
- Whether the release describes multiple events

Figure 1: Example press release

Notes : Each press release includes publication date, issuing station, and full text. Content is unstructured, and out-group cues can appear explicitly or implicitly.
Translated excerpt: "Three Suspected children and adolescents arrested. On August 1, 2016, at around 6:25 p.m., three male children and adolescents entered the sales area of a pharmacy on Heggerstraße. While two of them distracted the employee by asking questions in English and broken German, the third person secretly hid in a side room. The two accomplices initially left the store without purchasing anything but returned a short time later. At the same time, another employee entered the side room and discovered the hidden third person. He informed the owner, who immediately locked the entrance door and notified the police. The officers temporarily arrested the three youths aged 13 and 14, from Duisburg and of Romanian descent on suspicion of attempted theft [...]"

We developed our prompt through an iterative, evaluation-driven process (detailed in

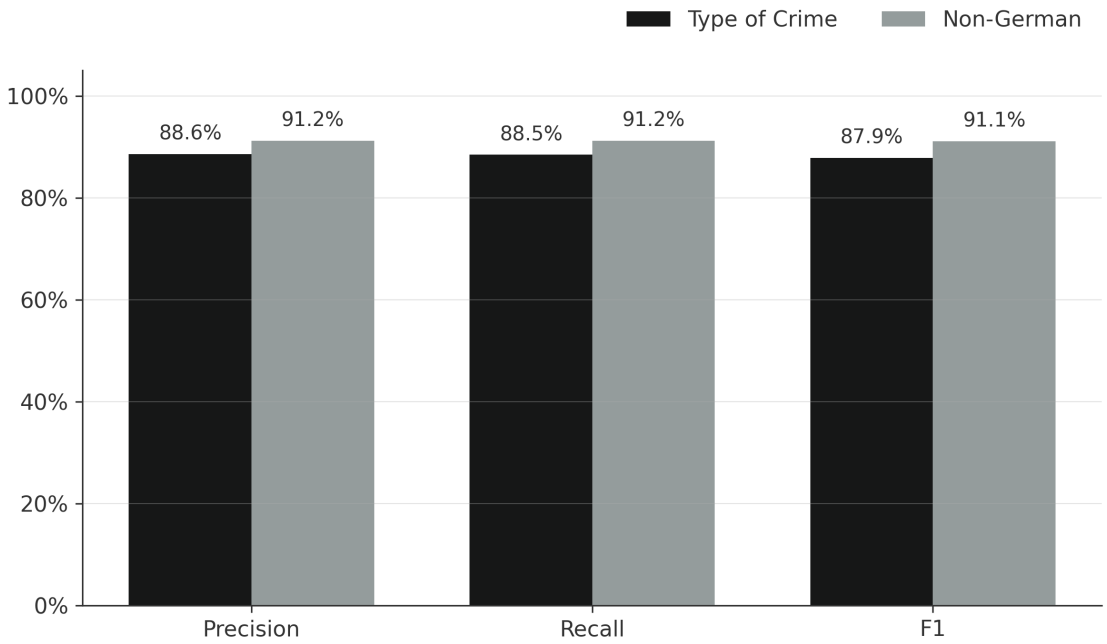
Table 2: Labeling of example press release

Variable	**Coding**
Publication Date	August 2nd, 2016
Location	Hattingen
Type of crime	Property
Investigation status	Concluded
Out-group cue	Yes
Out-group markers (verbatim)	"gebrochene deutsche Sprache; rumänischer Abstammung"
Multiple events	No

Appendix A.1.3, A.1.4, and A.1.8), refining it against a subset of our gold-standard, human-labeled press releases. Key refinements included: 1) structured `Markdown` formatting with variable-by-variable coding rules to ensure reliable instruction-following, and 2) strict output constraints to minimize hallucinations. Our pipeline also incorporates post-processing corrections, such as preventing the model from inferring German ethnicity from residence alone. To guard against overfitting, we validated our approach using a hold-out evaluation subset, confirming stable model performance across random splits of the labeled data.

We evaluate the performance of our LLM-based annotation pipeline against the human-labeled dataset to assess its reliability and performance (Baumann et al. 2025). The final model achieves high precision and recall for all key variables (Figure 2), with a full performance report available in Appendix A.1.6. This high performance using zero-shot learning is consistent with findings that explicit, structured codebooks are effective (Gilardi, Alizadeh and Kubli 2023). As a robustness check, we benchmarked our model (GPT-4o) against other LLMs, including GPT 3.5, Gemini 1.5 Pro, and Gemini 2 Flash. While GPT-4o outperformed all of these alternative models (Appendix A.1.6.1), its performance was comparable to other more modern Gemini models, demonstrating that our results are not dependent on the idiosyncrasies of a single model. GPT-3.5 underperforms these newer models, which is consistent with research emphasizing its less sophisticated instruction-following capabilities (Zheng et al. 2023; Achiam et al. 2023).

Figure 2: Classification Performance



Notes: Classification metrics in the plot are based on the human-annotated press releases' weighted averages. To align the evaluation sample with our empirical analysis, we exclude from both variables releases describing multiple incidents, ensuring that each observation corresponds to a single crime event. For the *Non-German* variable, we additionally omit traffic-related incidents, which fall outside the scope of our theoretical interest and are rarely coded for nationality or ethnicity. You can find the tables that include a breakdown of the performance for each sub-category in appendix A.1.6.

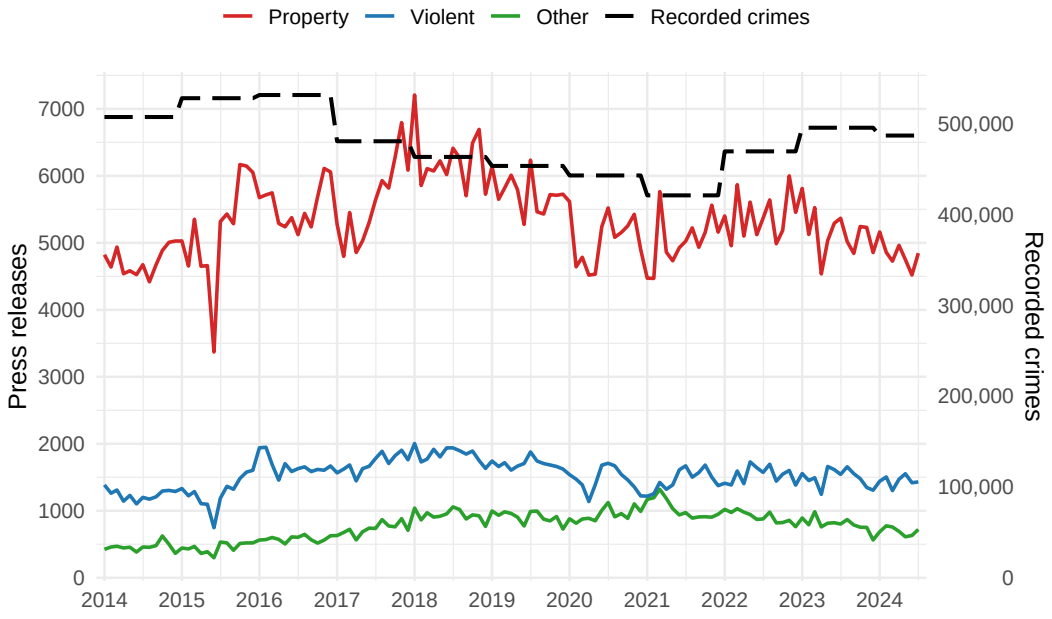
For further verification, we used a sample labeled by a second human coder to calculate intercoder reliability (ICR) both among human coders and between human coders and GPT (see Appendix A.1.7). This dual comparison provides a crucial validation step: ICR among coders validates coding consistency, while ICR with GPT assesses if the LLM performs comparably. Across all key variables, agreement among human coders is high, and GPT achieves broadly comparable reliability levels. These results underline that our approach not only achieves high classification performance but also effectively replicates the judgment patterns of human coders.

One might be tempted to benchmark the prevalence of out-group cues in our data against official crime statistics (*Polizeiliche Kriminalstatistik, PKS*). However, such comparison is methodologically problematic and risks producing misleading inferences. This is because the

press release data presents an intractable baseline problem. The vast majority of releases (over 85% in most years of our sample) omit any information on suspect nationality, making it impossible to construct a non-arbitrary baseline of in-group suspects. If we assume every "no mention" to be indicative of a German suspect, our baseline number of German suspects would be very large. If we do not make that assumption and instead rely on explicit mentions of German nationality only, the number would be very small. We thus refrain from any comparison. Instead, our analysis centers on the act of out-group cue disclosure itself, treating its variation over time as the key dependent variable to identify what political factors drive this strategic choice.

Figure 3 visualizes the number of press releases over time by crime type. Approximately 70% of all press releases concern property-related offenses, followed by around 20% related to violent crime. The overall distribution remains relatively stable over the ten-year period,

Figure 3: Police Press Releases vs. Crime Statistics, 2014–2024



Notes: The figure shows the monthly number of police press releases from our corpus by crime type (January 2014–July 2024), plotted on the left axis. The dashed black line (right axis) shows the average total number of recorded crimes by month, calculated by dividing the annual data from the *Polizeiliche Kriminalstatistik* (Statista 2025*b*) by twelve.

with no major shifts in the types of crimes reported.

5 Empirical Strategy

Our empirical strategy tests whether the disclosure of out-group cues in police press releases changes discontinuously around events that heighten the political salience of immigration. We leverage two distinct settings: the 2015–16 New Year's Eve attacks in Cologne (Study I) and state election cycles (Study II). For both studies, we employ a regression discontinuity in time (RDiT) design, where the unit of observation is the individual police press release.

The running variable, X_i, measures the time (in days) from the publication date of a press release i to the relevant cutoff date:

$$X_i \;=\; \Big(\text{publication date of release } i\Big) \;-\; \Big(\text{cut-off date}\Big).$$

We exclude observations where $X_i = 0$, as the treatment status on the cutoff day itself is ambiguous. Our main outcome variable, Y_i, is a binary indicator that equals 1 if a press release contains an out-group cue (e.g., non-German ethnicity or foreign nationality) and 0 otherwise.

For Study I, the cutoff is January 1, 2016, the day after the Cologne attacks. We hypothesize that the reputational shock and heightened salience led police to increase disclosure *after* this date. We define the treatment indicator accordingly:

$$\text{Post-NYE}_i \;=\; \mathbf{1}\{X_i > 0\}.$$

This indicator partitions the sample into releases published before ($X_i < 0$) and after ($X_i > 0$) the event. We estimate the following OLS model within bandwidth h:

$$Y_i \;=\; \alpha + \beta\,\text{Post-NYE}_i \;+\; \varepsilon_i, \tag{1}$$

where the coefficient β captures the average change in the probability of an out-group cue disclosure immediately following the New Year's Eve attacks.

For Study II, the cutoff is the date of the most proximate state election. We hypothesize that, in anticipation of heightened public attention, police increase disclosure *before* an election. The cutoff date for each press release is the date of the nearest state election. The treatment indicator is defined as:

$$\text{Pre-Election}_i \;=\; \mathbf{1}\{X_i < 0\}.$$

We estimate the effect using an OLS model that includes state $\times$ year fixed effects, $\alpha_{s,t}$, to ensure we are comparing reporting behavior around the same election:

$$Y_i \;=\; \alpha_{s,t} + \beta\,\text{Pre-Election}_i \;+\; \varepsilon_i. \tag{2}$$

Here, β measures the discontinuous change in out-group cue disclosure just *before* an election. Standard errors are clustered by election (state $\times$ year). Our main specifications use a bandwidth of $h = 2$ days, and we show results for alternative bandwidths in Appendix Figure A.6. Our results are substantively unchanged when using police-station $\times$ year fixed effects, i.e., when we zoom in on changes within the same police station before and after state elections (see Table 4).

A key identifying assumption for our empirical strategy is that the underlying *composition* of offenders remains stable across the cutoff. A direct test of this assumption is not possible, as official crime data lacks the necessary temporal granularity and is released only annually.[9] Nevertheless, we view this assumption as plausible, as it is unlikely that the true share of crimes committed by foreigners would discontinuously change within short time spans. Importantly, our design is robust to shifts in the overall number of crimes, provided that the

[9]The foreign suspect share in the overall crime statistics gradually rose from 24.3% in 2014 to 35.4% in 2024, tracking the increase in the foreign share of the general population (Statista 2025*a*).

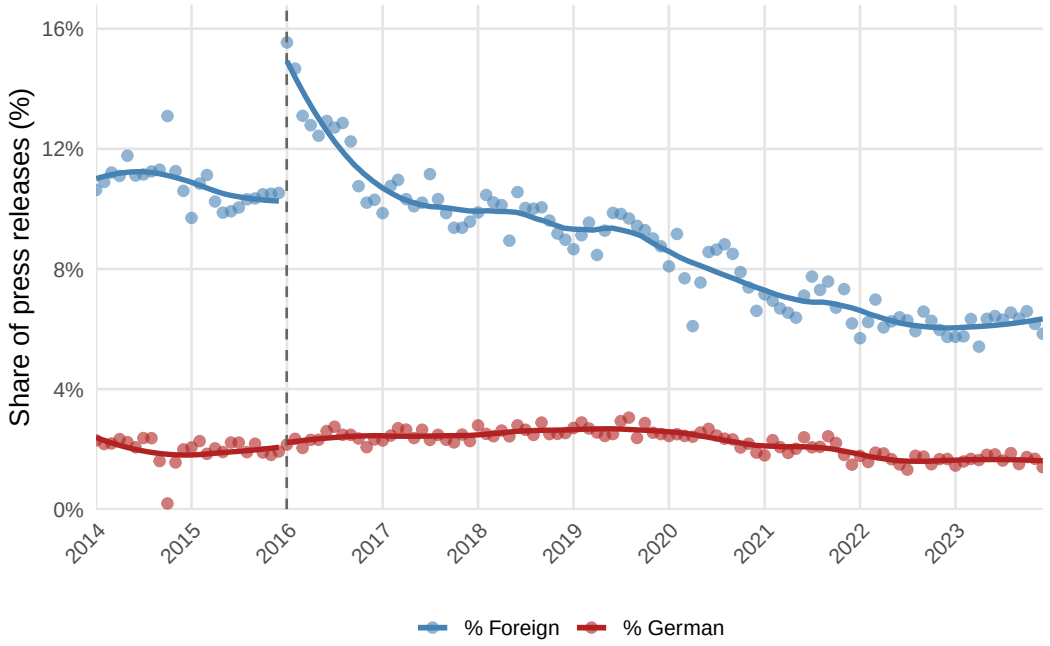
relative share of offenses by different groups remains smooth around the cutoff.

To validate our empirical strategy, we conduct two placebo treatment tests following Eggers, Tuñón and Dafoe (2024). First, for Study I (NYE), we re-estimate our model using every day between January and November 2015 as a placebo cutoff. The resulting distribution of placebo coefficients is tightly centered at zero, while the actual estimate from the true cutoff (January 1, 2016) is an extreme outlier, substantially larger than any placebo estimate (Figure A.5). This suggests the discontinuous increase in out-group disclosure is not driven by random temporal fluctuations. Second, for Study II (elections), we randomly reassign each state election to 1,000 different placebo Sundays and re-estimate the main specification each time. The randomization distribution of t-statistics is centered at zero, and our observed statistic lies in the extreme tail ($p = 0.003$; Figure A.7), indicating the effect is unlikely to arise from chance or simple day-of-the-week effects (see SI Figure A.2 for the distribution of press releases by weekday). Complementing this, we also show that the overall volume of press releases does not shift discontinuously at elections (SI Table A.17).

6 Study I: The 2015–16 New Year's Eve Attacks

On New Year's Eve 2015, hundreds of women in Cologne reported being sexually harassed, assaulted, and robbed in the central square near the main train station. Subsequent investigations revealed that attacks were carried out by groups of men described as being of North African and Middle Eastern origin (Baumgärtner et al. 2016). The events in Cologne quickly became a defining moment in Germany's 'refugee crisis'. Within days, media outlets and political actors framed the assaults as emblematic of the perceived dangers of Chancellor Angela Merkel's open-door asylum policy (Connolly 2016). What began as a local incident escalated into a national scandal that dominated headlines for weeks, fueling heated debates on immigration, integration, and public safety, severely straining the legitimacy of

Figure 4: Ethnicity cues in police press releases before and after 2015–16 New Year's Eve



Notes: The figure plots the monthly share of police press releases that include an explicit nationality/ethnicity cue for the suspect, shown separately for foreign/out-group and German references. Points denote monthly means. The vertical dashed line marks December 31, 2015 (Cologne New Year's Eve), and separate LOESS curves are fitted to observations on either side of this cutoff. Shares are expressed in percent.

the German police.[10]

A central element of the controversy was the police's initial communication strategy. Early reports, both in Cologne and nationally, downplayed the severity of events and omitted any reference to the suspects' ethnic backgrounds. When these omissions became public, police were accused of incompetence, engaging in a cover-up, and bowing to political pressure to avoid inflaming anti-immigrant sentiment (Bielicki 2019). This episode created a legitimacy crisis for law enforcement, putting them on the defensive.

[10]The change in federal hate-crime recording between 2015 and 2016 noted by Schwitter and Liebe (2023) does not confound our analysis. We study press releases for all crime types, not anti-refugee violence specifically. Hate crimes are a vanishingly small subset of all crime; in 2016, 8983 xenophobic offenses (PMK) constituted only $\approx 0.14\%$ of over 6 million total offenses. Reported right-wing attacks on asylum accommodations numbered less than 1,000. Our analysis is thus not affected by this reporting change. (BMI 2017).

Table 3: Differential reporting around NYE 2015

	Out-group cue (0/1)			
	(1)	(2)	(3)	(4)
Post-NYE	0.054***	0.052***	0.043***	0.053***
	(0.005)	(0.005)	(0.008)	(0.005)
R^2	0.006	0.083	0.004	0.006
Observations	15,837	15,837	8,272	15,399
Police station fixed effects		✓		
Excluding NRW			✓	
Excluding Cologne				✓

Notes: Results from OLS regressions where the outcome is a binary indicator for the presence of out-group cues in a police press release. The unit of observation is the individual press release. *Post-NYE* is a binary indicator equal to one for releases issued after NYE 2015. The analysis uses a bandwidth of ±30 days around NYE. Column (1) includes all data. Column (2) includes additionally station fixed effects. Column (3) removes the state of North Rhine-Westphalia (NRW) from the sample. Column (4) removes only Cologne from the sample. Significance codes: *** p<0.01, ** p<0.05, * p<0.1.

The New Year's Eve assaults function as a sharp, nationwide shock to the salience of immigration. Unlike gradual shifts in the political agenda, the events of a single night triggered a rapid, exogenous surge in public concern about immigration and crime. This provides a critical setting to test our theory: we can observe whether and how police strategically adjusted their communication practices in response to heightened accountability pressures. Our data reveal a marked discontinuity in the disclosure of out-group cues in press releases immediately following the assaults. As we show in Figure 4, local police stations across Germany—including those far from Cologne—increased the frequency with which they highlighted foreign nationality or ethnicity in crime reports. This change in disclosure is specific to out-group markers; we find no corresponding shift in the reporting of in-group (German) cues (see also SI Table A.14).

This nationwide shift cannot be explained by reporting on the Cologne perpetrators alone. The surge in disclosure is sudden, appears across diverse crime types, and is observed

in regions throughout Germany. These patterns are consistent with a broad recalibration of police communication strategy in response to heightened salience, as police sought to signal transparency and regain public trust. To quantify the magnitude of this shift, we estimate the specification described in section 5. We estimate the effect of the event on the likelihood that a press release contains an out-group cue. The results, presented in Table 3, show a substantive and statistically significant increase in such disclosures.

Column (1) shows that the probability of an out-group cue appearing in a press release was approximately 5 percentage points lower in the 30 days before New Year's Eve compared to the 30 days after. This finding is robust to the inclusion of police station fixed effects (Column 2), which accounts for time-invariant differences in reporting practices across stations. To demonstrate that this effect is not driven solely by reporting on the Cologne events, we exclude the state of North Rhine-Westphalia (where Cologne is located) in Column (3). The coefficient estimate is only slightly reduced, remaining large and statistically significant, which confirms that the shift in reporting was a nationwide phenomenon. Column (4) further corroborates this by showing the effect persists even when only the city of Cologne is excluded. These results indicate that the New Year's Eve attacks triggered a widespread, strategic shift in how police communicate information about crime and suspect nationality.

7 Study II: Electoral cycles

Our first analysis demonstrated that a major, exogenous shock to immigration crime salience—the 2015/16 New Year's Eve assaults—triggered a sharp increase in police reporting of out-group crime. This finding raises a crucial question: is such strategic behavior limited to salient 'focusing events' that attract national media attention, or does it represent a more systematic pattern? To answer this, we turn to a second source of over-time variation in immigration salience: regional (state) elections. We focus on state (*Landtag*) elections because political authority over the police is vested at the state level.

Elections are predictable moments of heightened political attention in which parties,

journalists, and voters converge on a narrow set of contested issues. While other issues, such as the war in Ukraine or the energy crisis, have temporarily captured public attention, no single topic has dominated the political agenda in Germany since 2014 as consistently and intensely as immigration (Forschungsgruppe Wahlen e.V. 2025). However, the degree to which immigration dominates the campaign agenda is not constant across all elections, offering an additional source of variation that we can leverage in our analysis. Elections thus provide a setting that allows us to test whether police reporting behavior covaries with peaks in public attention and immigration salience, when incentives to signal transparency are highest.

We estimate the specification described in section 5. Our main OLS results are presented in Table 4. We identify a substantively large and statistically significant increase in the probability of a police press release containing an out-group cue immediately before a state election. Specifically, our estimates reveal a sharp increase of approximately 3 percentage points in the reporting of out-group cues. This discontinuity at the election cutoff date corresponds to an increase of 32% relative to the mean of our outcome variable. One concern is that this finding may reflect changes in reported crime types, such as a shift from reporting property crimes to violent crimes, which may have a higher baseline rate of out-group suspects. We test for this alternative explanation in Table A.20, finding no discontinuous shift in the shares of reported crime types at the electoral cutoff.

We present evidence in SI Table A.15 that this increased rate of out-group cues in the pre-election period is accompanied by a corresponding statistically significant decrease in the share of press releases that provide no information regarding the ethnicity or nationality of the alleged offender. Analogous to our analysis of reporting behavior around the 2015/16 NYE, we do not find evidence of discontinuous shifts for the rate at which in-group (German) cues are reported in police press releases around state elections. Our effects are thus specific to out-groups and not driven by a general shift in the propensity to report ethnic or national

Table 4: Differential reporting around elections

	Out-group cue (0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-Election	0.029**	0.024**	0.029***	0.027***	0.023***	0.028***
	(0.012)	(0.010)	(0.009)	(0.010)	(0.007)	(0.008)
Pre-Election × Mig. salience				0.019**	0.018***	0.015*
				(0.008)	(0.006)	(0.007)
R^2	0.002	0.178	0.035	0.006	0.179	0.036
Observations	2,219	2,186	2,218	2,219	2,186	2,218
Police station x Election fixed effects		✓			✓	
State x Election fixed effects			✓			✓

Notes: Results from OLS regressions where the outcome is a binary indicator for the presence of out-group cues in a police press release. The unit of observation is the individual press release. *Pre-election* is a binary indicator equal to one for releases issued before a given state election. The analysis uses a bandwidth of ±2 days around state elections. The moderator is a survey-based measure of immigration salience for each state election, defined as the share of respondents naming integration, foreigners, or refugees as one of the two most important political issues facing the country. Immigration salience is calculated from the two most recent Politbarometer survey waves conducted in each state prior to its respective state election, incorporating survey weights, and standardized (mean = 0, SD = 1). Standard errors are clustered at the state-election level. Significance codes: *** p<0.01, ** p<0.05, * p<0.1.

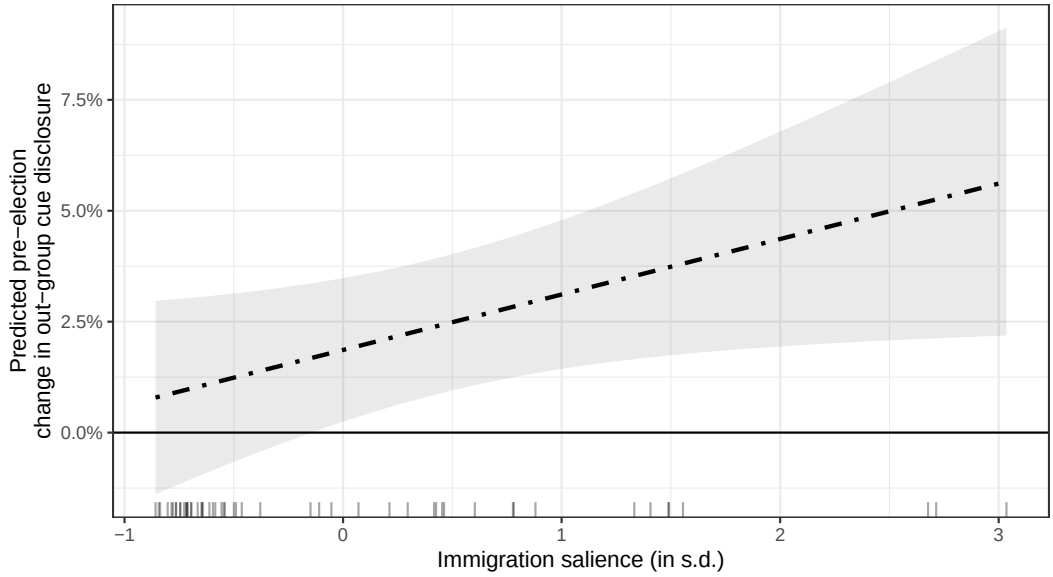
identifiers around elections.[11]

Consistent with our theoretical framework, we find that the differential increase in out-group cue disclosure prior to elections is particularly pronounced in elections where immigration is salient. Figure 5 demonstrates this by exploiting cross-election variation in immigration salience across state contests. We proxy salience with the share of respondents in the two *Politbarometer* waves preceding each election who name *integration/foreigners/refugees* as one of the two most important political issues (Forschungsgruppe Wahlen 2025). Similarly, we find that our pre-election effects are amplified in areas with high AfD support, where local demand for out-group crime information is likely highest (Table A.21).

An alternative mechanism is that police might engage in a form of strategic political partisanship, intentionally seeking to influence electoral outcomes. This account posits that

[11] As an additional placebo test, we analyze press releases issued by the federal police (*Bundespolizei*). This federal agency lacks the specific state-level reputational incentives we theorize for state police. We find a null result: federal police show no statistically significant, discontinuous change in their disclosure of out-group cues around state elections (Table A.22).

Figure 5: Marginal effects by immigration salience



Notes: The plot shows the estimated post-election change in the probability that a police press release contains an out-group cue as a function of immigration salience. The solid line is the estimated marginal effect with 95% confidence intervals. The rug marks the empirical distribution of the moderator across state elections. Immigration salience is measured as the share of respondents naming *integration/foreigners/refugees* as one of the two most important political issues in the two Politbarometer waves prior to each election (standardized). Estimates are based on the model presented in column 6 in Table 4.

police personnel, who tend to lean right-wing (Clasen et al. 2024; Reny et al. 2025), selectively emphasize out-group crime to promote a "law and order" narrative. By increasing the salience of immigration and crime, they could hope to benefit right-leaning parties, which not only align with their ideological preferences but are also more likely to increase police budgets and grant them greater operational leeway. While this account is plausible, we present two pieces of empirical evidence that speak against it. First, the data are inconsistent with an explanation of top-down political directives. If such mechanisms were at play, the effect should be stronger under an interior ministry with partisan incentives to elevate the issue. Yet, we find no statistically significant difference in the pre-election shift in out-group disclosure based on the partisan control of the state interior ministry (Table A.18). Although point estimates are somewhat larger under center-right (CDU/CSU) control, this difference is not statistically significant. Second, the behavior does not appear calibrated to

maximize electoral impact. If police were strategically attempting to sway elections, their efforts should be most pronounced in tight races where a small shift in public opinion could alter the outcome. Contrary to this, we find that the pre-election increase in out-group cue disclosure does not systematically vary with electoral competitiveness (Table A.19). The effect is no larger in close contests than in uncompetitive ones. Taken together, these results speak against strategic electoral manipulation.

8 Discussion

Across two empirical case studies, we document that local police stations across Germany strategically increase their reporting of out-group crime when the political salience of immigration is high. Our findings suggest that the political information environment is shaped not only by media outlets but also by upstream state actors who exercise discretion over what information becomes available for the news media to report.

The generalizability of our findings likely depends on two key scope conditions. First, the issues of crime and immigration must be closely linked in public discourse, such that the disclosure of ethnicity information about suspects is politicized and represents a meaningful signal of police transparency. This condition is likely met in many settings beyond Germany; similar debates over whether the police should disclose the ethnicity or immigration status of suspects are salient in many countries, including the United Kingdom (Reuters 2025), Sweden (The Independent 2016), Switzerland (SWI 2021), and Australia (The Guardian 2018). Second, the magnitude of the effects we document likely depends on the institutional context. Germany's decentralized policing structure and vague reporting guidelines create substantial local discretion for police press offices. In systems with more centralized control or stricter, legally mandated disclosure rules – either mandating or forbidding the release of ethnicity information in press releases – the scope for this type of strategic behavior may be smaller.

A question beyond the scope of this paper is whether the strategic disclosure of out-group

cues "works" – that is, whether increasing the disclosure of out-group cues in police communications boosts perceived accountability or public trust in the police. We descriptively examine this question in SI Section A.5, where we correlate local out-group reporting with individual-level survey data on trust in the police (Hirndorf et al. 2023). We do not find any correlational evidence that a higher share of out-group reporting is associated with increased trust in the police (Table A.23). This finding is, however, merely correlational. Future research (e.g., using vignette experiments) is needed to causally study how citizens interpret and respond to selective transparency in police communication.

In addition, we cannot fully rule out the alternative explanation that strategic disclosure is at least partially driven by the individual ideological preferences of police personnel. Our analysis speaks against the most direct forms of political capture. We find that disclosure patterns do not systematically vary with the partisanship of the state interior minister or electoral competitiveness. We cannot, however, rule out the influence of individual-level ideology. Future research should test this channel more directly by linking station-level disclosure behavior to measures of police officers' ideological preferences.[12]

Our findings have important policy implications. Across many advanced democracies, there is a striking mismatch between perceptions and reality with regard to crime: while aggregate crime rates have declined in many countries, public concerns about crime remain high or have even increased (Financial Times 2025; Pew Research Center 2024*c*). Our research identifies a potential supply-side driver of this phenomenon. The strategic amplification of out-group crime during politically salient periods can create a distorted information environment, leading the public to perceive crime as more prevalent and more closely linked to immigration than official statistics would suggest. One potential remedy could be to reform and potentially limit discretion in reporting. Prior research suggests that a policy of

[12]While such data exists (e.g., "MEGAVO – Motivation, Einstellung und Gewalt im Alltag von Polizeivollzugsbeamten"), it is not publicly accessible, and our requests for this data were unsuccessful, precluding a direct test of this mechanism.

consistent disclosure – where the ethnicity and nationality of *all* suspects, including members of the native in-group, are reported – may limit the negative political consequences of out-group crime reporting (Keita, Renault and Valette 2024). However, the effectiveness of such a policy likely depends on the representativeness of the initial pool of crimes selected for public communication. If bias is already present in the first selection step (i.e., which crimes are chosen for a press release), then a policy of full disclosure in the second step could inadvertently amplify threat narratives linking out-groups to crime. This suggests that effective reform is not simply about disclosing more or less information, but about ensuring the entire reporting process – from the initial selection of which crimes to publicize to the details disclosed within those reports – is systematic and consistent, rather than strategic and selective.

References

Achiam, Josh, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat et al. 2023. "Gpt-4 technical report." *arXiv preprint arXiv:2303.08774* .

Alizade, J. 2025. "The Electoral Politics of Immigration and Crime." https://doi.org/10.31219/osf.io/h967e_v2.

Alizade, Jeyhun, Rafaela Dancygier and Jonathan Homola. 2024. "Structures of Bias: How the State Systematically Downplays Right-Wing Extremism.".
URL: *https://www.jeyhunalizade.com/s/State_Bias_RWE_Germany.pdf*

Banks, Jeffrey S and Barry R Weingast. 1992. "The political control of bureaucracies under asymmetric information." *American Journal of Political Science* pp. 509–524.

Baumann, Joachim, Paul Röttger, Aleksandra Urman, Albert Wendsjö, Flor Miriam Plaza-del Arco, Johannes B Gruber and Dirk Hovy. 2025. "Large language model hacking: Quantifying the hidden risks of using LLMs for text annotation." *arXiv preprint arXiv:2509.08825* .

Baumgärtner, Maik, Markus Brauck, Jürgen Dahlkamp, Jörg Diehl, Ulrich Fichtner, Jan Friedman, Matthias Geyer, Hubert Gude, Horand Knaup, Alexander Kühn, Dialika Neufeld, Ralf Neukirch, Ann-Kathrin Nezik, Miriam Olbrisch, Maximilian Popp, Gordon Repinski, Sven Röbel, Barbara Schmid, Fidelius Schmid, Andreas Ulrich and Antje Windman. 2016. "Chaos and Violence: How New Year's Eve in Cologne Has Changed Germany." *Spiegel Online* . (last accessed: July 31st, 2021).
URL: *http://www.spiegel.de/international/germany/cologne-attacks-trigger-raw-debate-on-immigration-in-germany-.a-1071175.html*

Berk, Nicolai. 2025. "The impact of media framing in complex information environments." *Political Communication* pp. 1–17.

Bielicki, Jan. 2019. Cologne's New Year's Eve Sexual Assaults: The Turning Point in German Media Coverage. In *Refugee News, Refugee Politics: Journalism, Public Opinion and Policy Making in Europe*, ed. Giovanna Dell' Orto and Irmgard Wetzstein. New York: Routledge.

BMI. 2017. Politisch motivierte Kriminalität im Jahr 2016. Technical report Bundesministerium des Innern Berlin: .

URL: *https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/2017/pmk-2016.pdf?__blob=publicationFile&v=1*

Carpenter, Daniel. 2001. *The forging of bureaucratic autonomy: Reputations, networks, and policy innovation in executive agencies, 1862-1928*. Princeton University Press.

Carpenter, Daniel. 2010. *Reputation and power: Organizational image and pharmaceutical regulation at the FDA*. Princeton University Press.

Carpenter, Daniel P and George A Krause. 2012. "Reputation and public administration." *Public administration review* 72(1):26–32.

Clasen, Julia, Eva Groß, Joachim Häfele, Ilka Kammigan, Stefanie Kemme, Amelie Nickel, Anabel Taefi, Ulrike Zähringer, Iniobong Essien and Marleen Stelter. 2024. "Demokratiebezogene Einstellungen und Werthaltungen in der Polizei Hamburg (DeWePol): Erste Befunde aus Befragungswelle 1.".

URL: *https://akademie-der-polizei.hamburg.de/resource/blob/1029176/3acbf3a748308047 2eb5e2ed1a6bb611/projekt-dewepol-bericht-2024-do-data.pdf*

Connolly, Kate. 2016. "Cologne Inquiry into 'coordinated' New Year's Eve Sex Attacks." *The Guardian* . (last accessed: July 31st, 2021).

URL: *http://www.theguardian.com/world/2016/jan/05/germany-crisis-cologne-new-years-eve-sex-attacks*

Cook, Scott J and David Fortunato. 2023. "The politics of police data: State legislative capacity and the transparency of state and substate agencies." *American political science*

review 117(1):280–295.

Couttenier, Mathieu, Sophie Hatte, Mathias Thoenig and Stephanos Vlachos. 2021. "Anti-muslim voting and media coverage of immigrant crimes." *Review of Economics and Statistics* pp. 1–33.

Eggers, Andrew C, Guadalupe Tuñón and Allan Dafoe. 2024. "Placebo tests for causal inference." *American Journal of Political Science* 68(3):1106–1121.

Evangelische Zeitung. 2024. "Nicht alle Bundesländer nennen Nationalität von Tatverdächtigen." https://evangelische-zeitung.de/nicht-alle-bundeslaender-nennen-die-nationalitaet-von-tatverdaechtigen. Accessed 2025-10-02.

Financial Times. 2025. "The great crime paradox.".
URL: *https://www.ft.com/content/7488fe4c-5e1d-4b2b-adab-f42ad5273fc9*

Fitzgerald, Jennifer, K Amber Curtis and Catherine L Corliss. 2012. "Anxious publics: Worries about crime and immigration." *Comparative Political Studies* 45(4):477–506.

Forschungsgruppe Wahlen e.V. 2025. "Politik II — Langzeitentwicklung: Themen im Überblick (Politbarometer)." https://www.forschungsgruppe.de/Umfragen/Politbarometer/Langzeitentwicklung_-_Themen_im_Ueberblick/Politik_II/. Last updated 2025-08-14; accessed 2025-08-15.

Forschungsgruppe Wahlen, Mannheim. 2025. "Politbarometer 1977-2023 (Partial Cumulation)." (ZA2391; Version 16.1.0) [Data set]. GESIS, Cologne. https://doi.org/10.4232/1.14527.

Frey, Arun. 2020. "'Cologne Changed Everything'—the effect of threatening events on the frequency and distribution of intergroup conflict in Germany." *European Sociological Review* 36(5):684–699.

Giavazzi, Francesco, Felix Iglhaut, Giacomo Lemoli and Gaia Rubera. 2024. "Terrorist attacks, cultural incidents, and the vote for radical parties: Analyzing text from Twitter."

American journal of political science 68(3):1002–1021.

Gilardi, Fabrizio, Meysam Alizadeh and Maël Kubli. 2023. "ChatGPT outperforms crowd workers for text-annotation tasks." *Proceedings of the National Academy of Sciences* 120(30):e2305016120.

Grunwald, Ben, Julian Nyarko and John Rappaport. 2022. "Police agencies on Facebook overreport on Black suspects." *Proceedings of the National Academy of Sciences* 119(45):e2203089119.

Helbling, Marc and Daniel Meierrieks. 2022. "Terrorism and migration: An overview." *British Journal of Political Science* 52(2):977–996.

Hestermann, Thomas. 2025. "Medienprofis in Uniform: Wie die Polizei Delikte darstellt." Presentation at the 30th Deutscher Präventionstag. 14:00–14:45 Uhr, Raum 7 (TC-2.24B). **URL:** *https://www.praeventionstag.de*

Hirndorf, Dominik, Viola Neu, Sabine Pokorny and Jochen Roose. 2023. "Perception of Crime in Germany 2020." (ZA7946; Version 1.0.0) [Data set]. GESIS, Cologne. https://doi.org/10.4232/1.14068.

Hoekstra, Mark and CarlyWill Sloan. 2022. "Does race matter for police use of force? Evidence from 911 calls." *American economic review* 112(3):827–860.

Hu, Xiaochen, Kourtnie Rodgers and Nicholas P. Lovrich. 2018. ""We are more than crime fighters": social media images of police departments." *Police Quarterly* 21(4):544–572.

Keita, Sekou, Thomas Renault and Jérôme Valette. 2024. "The Usual Suspects: Offender Origin, Media Reporting and Natives' Attitudes Towards Immigration." *The Economic Journal* 134(657):322–362.

Lipsky, Michael. 1980. *Street-level bureaucracy: Dilemmas of the individual in public service*. Russell sage foundation.

Marie, Olivier and Paolo Pinotti. 2024. "Immigration and crime: An international perspec-

tive." *Journal of Economic Perspectives* 38(1):181–200.

Ministerium des Inneren, für Digitalisierung und Kommunen Baden-Württemberg. 2022. "Medienkodex Polizei Baden-Württemberg: Richtlinien für die polizeiliche Berichterstattung – Publizistische Grundsätze." *Ministerium des Inneren, für Digitalisierung und Kommunen Baden-Württemberg.*

URL: *https://im.baden-wuerttemberg.de/de/sicherheit/polizei/medienkodex*

Ministerium für Inneres und Bau Mecklenburg-Vorpommern. 2020. "Benennung der Nationalität von Tatverdächtigen in der Kommunikationsarbeit der Polizei in Mecklenburg-Vorpommern.".

URL: *https://www.regierung-mv.de/Landesregierung/im/Aktuell/?id=157232&pager.items.offset=3830&pager.page.nr=383*

Moore-Berg, Samantha L, Boaz Hameiri and Emile G Bruneau. 2022. "Empathy, dehumanization, and misperceptions: A media intervention humanizes migrants and increases empathy for their plight but only if misinformation about migrants is also corrected." *Social Psychological and Personality Science* 13(2):645–655.

Nowacki, Tobias and Daniel M Thompson. 2021. "How much do elections increase police responsiveness? Evidence from elected police commissioners." *https://dthompson.scholar.ss.ucla.edu/wp-content/uploads/sites/19/2021/05/Nowacki_Thompson_Commissioners.pdf* .

O'Connor, Christopher D. 2017. "The police on Twitter: image management, community building, and implications for policing in Canada." *Policing and society* 27(8):899–912.

O'Connor, Christopher D. and Huda Zaidi. 2020. "Communicating with purpose: Image work, social media, and policing." *The Police Journal* pp. 1–20.

Pew Research Center. 2024*a*. "How Americans get local political news.". Accessed: 2025-06-18.

URL: *https://www.pewresearch.org/journalism/2024/07/24/how-americans-get-local-*

political-news/

Pew Research Center. 2024*b*. How Americans View the Situation at the U.S.-Mexico Border, Its Causes and Consequences. Technical report Pew Research Center.
URL: *https://www.pewresearch.org/politics/2024/02/15/how-americans-view-the-situation-at-the-u-s-mexico-border-its-causes-and-consequences/*

Pew Research Center. 2024*c*. "What the data says about crime in the U.S." https://www.pewresearch.org/short-reads/2024/04/24/what-the-data-says-about-crime-in-the-us/.

Reny, Tyler T, Marcel F Roman, Benjamin J Newman and David O Sears. 2025. "Late adolescents entering college intending a career as police officers hold more right-leaning views than their peers." *Proceedings of the National Academy of Sciences* 122(32):e2500220122.

Reuters. 2025. "UK police can disclose suspects' ethnicity to curb misinformation.".
URL: *https://www.reuters.com/world/uk/uk-police-can-disclose-suspects-ethnicity-curb-misinformation-2025-08-13/*

Rho, Eugenia H., Reid Pryzant, Meghan Healy, Yang Xu, Nicholas P. Camp, Dan Jurafsky and Jennifer L. Eberhardt. 2023. "Escalated police stops of Black men are linguistically and psychologically distinct in their earliest moments." *Proceedings of the National Academy of Sciences of the United States of America* 120(46):e2307977120.

Riaz, Sascha, Daniel Bischof and Markus Wagner. 2024. "Out-group threat and xenophobic hate crimes: evidence of local intergroup conflict dynamics between immigrants and natives." *The Journal of Politics* 86(4):1146–1161.

Riaz, Sascha and Maik Hamjediers. 2025. "Anti-Immigrant Bias in the Choice Between Punitive and Rehabilitative Justice.".
URL: *https://osf.io/preprints/socarxiv/yvfzx*

Riaz, Sascha and Tobias Roemer. 2025. "Racial Threat in Public Social Spaces." *The Journal of Politics* . Forthcoming.

URL: *https://www.journals.uchicago.edu/doi/10.1086/736112*

Rumbaut, Rubén G and Walter A Ewing. 2007. *The myth of immigrant criminality and the paradox of assimilation: Incarceration rates among native and foreign-born men.* Immigration Policy Center, American Immigration Law Foundation.

Sabet, Navid, Marius Liebald and Guido Friebel. 2025. "Terrorism and voting: The rise of right-wing populism in Germany." *American Economic Journal: Economic Policy* 17(3):407–440.

Schneider-Strawczynski, Sarah and Jérôme Valette. 2025. "Media coverage of immigration and the polarization of attitudes." *American Economic Journal: Applied Economics* 17(1):337–368.

Schwitter, Nicole and Ulf Liebe. 2023. "Not Cologne but the data collection (might have) changed everything: a cautionary tale on ignoring changes in data recording in sociological research." *European Sociological Review* 39(6):1005–1010.

Semyonov, Moshe, Rebeca Raijman and Anastasia Gorodzeisky. 2008. "Foreigners' impact on European societies: public views and perceptions in a cross-national comparative perspective." *International Journal of Comparative Sociology* 49(1):5–29.

Shayo, Moses and Asaf Zussman. 2011. "Judicial ingroup bias in the shadow of terrorism." *The Quarterly journal of economics* 126(3):1447–1484.

Skinner, Gideon and Glenn Gottfried. 2017. "Global views on immigration and the Refugee crisis." *URL: http://tinyurl. com/yynk9eqs* .

Stansfield, Richard and Brenna Stone. 2018. "Threat perceptions of migrants in Britain and support for policy." *Sociological Perspectives* 61(4):592–609.

Statista. 2025*a*. "Anteil der ausländischen Straftatverdächtigen in Deutschland von 2014 bis 2024.".

URL: *https://de.statista.com/statistik/daten/studie/2460/umfrage/anteile-*

nichtdeutscher-verdaechtiger-bei-straftaten-zeitreihe/

Statista. 2025*b*. "Anzahl der polizeilich erfassten Straftaten in Deutschland von 1987 bis 2024.". Quelle: Bundeskriminalamt (Polizeiliche Kriminalstatistik, PKS).
URL: *https://de.statista.com/statistik/daten/studie/197/umfrage/straftaten-in-deutschland-seit-1997/*

Statistisches Bundesamt. 2024. "Vollzeitäquivalent der Beschäftigten im Aufgabenbereich Polizei in den Kernhaushalten des Bundes und der Länder." Webseite, Stand 11. Oktober 2024.
URL: *https://www.destatis.de/DE/Themen/Staat/Oeffentlicher-Dienst/Tabellen/beschaeftigten-polizei.html*

Statistisches Bundesamt (Destatis). 2025. "Beschäftigte des öffentlichen Dienstes nach ausgewählten Aufgabenbereichen.". Zeitreihe, Stichtag jeweils 30. 06.; enthält den Aufgabenbereich *Polizei*.
URL: *https://www.destatis.de/DE/Themen/Staat/Oeffentlicher-Dienst/Tabellen/beschaeftigte-ausgew-aufgaben.html*

SWI. 2021. "Zurich police to disclose nationality of crime suspects.".
URL: *https://www.swissinfo.ch/eng/politics/zurich-police-to-disclose-nationality-of-crime-suspects/46413138*

The Guardian. 2018. "Is Melbourne in the grip of African crime gangs? The facts behind the lurid headlines.". Last modified 13 March 2018.
URL: *https://www.theguardian.com/australia-news/2018/jan/03/is-melbourne-in-the-grip-of-african-gangs-the-facts-behind-the-lurid-headlines*

The Independent. 2016. "Swedish police banned from describing criminals anymore in case they sound racist.".
URL: *https://www.independent.co.uk/news/world/europe/swedish-police-are-not-allowed-to-give-descriptions-of-alleged-criminals-so-as-not-to-sound-racist-a6810311.html*

Voigt, Rob, Nicholas P Camp, Vinodkumar Prabhakaran, William L Hamilton, Rebecca C Hetey, Camilla M Griffiths, David Jurgens, Dan Jurafsky and Jennifer L Eberhardt. 2017. "Language from police body camera footage shows racial disparities in officer respect." *Proceedings of the national Academy of sciences* 114(25):6521–6526.

Wirsching, Elisa. 2025. "Political power of bureaucratic agents: Evidence from policing in New York city.".

URL: *https://elisawirsching.github.io/research/policeresistance.pdf*

Xu, Wenfei, Michael Smart, Nebiyou Tilahun, Sajad Askari, Zachary Dennis, Houpu Li and David Levinson. 2024. "The racial composition of road users, traffic citations, and police stops." *Proceedings of the National Academy of Sciences* 121(24):e2402547121.

YouGov. 2019. "What's the impact of immigration, according to Europeans and Americans?".

URL: *https://today.yougov.com/politics/articles/23792-whats-impact-immigration-according-europeans-and-americans*

YouGov. 2025. "Western Europeans Say Immigration Is Too High and Poorly Managed, Survey Finds." The Guardian.

URL: *https://www.theguardian.com/uk-news/2025/feb/26/western-europeans-say-immigration-is-too-high-and-poorly-managed-survey-finds*

Zheng, Lianmin, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric Xing et al. 2023. "Judging llm-as-a-judge with mt-bench and chatbot arena." *Advances in neural information processing systems* 36:46595–46623.

A Supporting Information (Online Only)

Contents

A.1 Data collection & processing

A.1.1 Scraping

Presseportal

We scrape all police press releases published on *Presseportal* between 2020 and August 2024, collecting date, police station, headline, and full text. We collect releases from both state and federal police. We combine this data with data collected by Hestermann (2025), who scraped police press releases from the same platform, and collected the same variables, between 2014 and 2019.

Supplementary state police scraping

In five states, Berlin, Saxony, Brandenburg, Saxony-Anhalt, and Saarland, police press releases on Presseportal were initially available only for the federal police. To address this, we supplemented these records with state police press releases retrieved directly from the official websites of the Berlin, Brandenburg, and Saxony-Anhalt police. Saarland did not make state police press releases publicly available on its website. However, beginning in 2021, it began contributing regularly to Presseportal, and we include all reports from that year onwards. In the case of Saxony, the state police website restricts access to press releases published within the current week only. Efforts to recover earlier postings using archival tools (e.g., the Wayback Machine) proved unsuccessful due to limited archival coverage.

For the remaining three states – Berlin, Brandenburg, and Saxony-Anhalt – we were able to recover press releases for the full period from 2014 to 2024. In Berlin, however, a substantial number of older releases appear to have been removed from the official web archive. To address this, we rely on a separate dataset compiled by one of the authors for an earlier project, which includes all Berlin state police press releases from 2014 to 2019. For the years 2020 to 2024, we use the releases currently available in the online archive. As a result, coverage for Berlin may be less complete.

A.1.2 Identifying and dropping traffic-related press releases

The raw database contains several categories of releases that are not directly relevant to our analyses (e.g., traffic accidents) and press releases that do not concern a crime (e.g., invitations to events).

In a first step, we remove press releases on traffic accidents as a category of releases that is (i) relatively numerous and (ii) relatively easy to identify. In particular, we perform string matching in the headline on the term "[U]nfall" (accident). This is possible because press releases that mention "accident" in the headline almost certainly imply that the actual press release is about a traffic accident. Moreover, while police stations sometimes combine

press releases, accidents are rarely combined with more serious crimes, but rather with other accidents in the area. Manual validation of a random sample confirms that this first-step procedure is accurate, though conservative. We identify other categories of press releases that are not relevant to our analyses as well as traffic accidents that we missed here in a separate step using GPT-4 classification (see below sections, i.e A.1.3 and A.1.4).

A.1.3 Labeling

Our research assistant labeled a random sample of **5364** press releases using the codebook below. To calculate measures of Intercoder Reliability, we had another research assistant label a random sample of **1094** rows.

Codebook for human coders

We classify the following variables:

Type of Crime:

- Violent Crime
  - Homicide: intentional killing of one person by another
  - Assault: physical attack or threat of violence against another person
  - Robbery: taking or attempting to take something of value from someone by force or threat of force
  - Sexual assault: non-consensual sexual contact or behavior
  - Kidnapping: taking a person against their will and holding them captive
- Property Crime
  - Theft: taking someone else's property without their permission
  - Burglary: breaking and entering into a building or vehicle with the intent to steal
  - Arson: deliberately setting fire to property
  - Vandalism: damaging or destroying someone else's property
- Traffic crimes:
  - Driving without a license
  - Illegal street racing

  - Hit and run
  - Accidents
  - Driving under the influence of alcohol

- Other: Includes all other crimes that do not fall into any of the above categories, such as drug-related crime, white-collar crime, and any other type of crime that does not fall under the above categories.

- None: use this category for press releases that do not describe criminal activity. For ambiguous cases that do not neatly fit into the above categories, use your best judgment based on the context given in the press release. For example, if a crime involves both property damage and violence, but the violence is the focus of the report, classify it as 'violent'

Ethnicity:

- No Information: perpetrator/suspect is unknown, no information is available

- No Ethnicity or Nationality: perpetrator/suspect is known, but no information on ethnic background is given

- German: if perpetrator is described as German national or hints of German ethnicity are given or stereotypical German physical characteristics or native command of the German language (Deutsch, blonde, blue eyes, sprach akzentfrei deutsch, ...)

- Foreign: if perpetrator is described as a foreign national or of non-German ethnicity.

Multiple Crime Events Classification
Does the report include multiple crimes reported together?

- Yes

- No

Crime Conclusion
Has the crime investigation been concluded?

- Ongoing

- Concluded

- Not Applicable (NA)

We designed our codebook to focus on the most relevant outcome variables. For the *Type of Crime*, we initially considered separate categories for white-collar and drug-related crimes, but later merged these categories into "other" to simplify the task and improve performance. In the *Non German* category, we differentiated between cases with no available information about the perpetrator (coded as *No Information*) and cases where the police have some information but do not disclose ethnicity/nationality cues (coded as *No Ethnicity or Nationality*), which might indicate a deliberate decision not to reveal such details. Additionally, we added a variable for *Multiple Crime* events because police stations sometimes combine several events, often less significant press releases, such as traffic incidents, into one report.

A.1.4 Prompt Design and Learning Choice

To achieve consistent and reliable model outputs, we adopted an iterative, evaluation-driven prompt design approach. Our goal was to minimize hallucinations and empty outputs while maximizing agreement with human-coded classifications across a large and noisy dataset of German police press releases.

We use a zero-shot prompt that we developed over several months through extensive testing. At each stage, we evaluated prompt variants against a subset of the gold standard dataset labeled by trained research assistants (RAs), focusing on classification performance and response completeness. We tracked improvements by computing agreement rates with this subset and manually inspecting edge cases where the model diverged from human coders. This process yielded a *stable* prompt that produced low-variance results across random samples and minimized common LLM failures, such as skipping entries or overgeneralizing. To prevent overfitting, we set aside a hold-out evaluation set (from the labeled sample provided by our research assistants). We report results (figure 2 and table A.1) on the total sample; results remain similar when restricted to other data subsets. This evaluation technique helps ensure that our evaluation highlights the model's ability to generalize beyond the examples it has seen during prompt engineering, as we wanted to avoid the model simply reproducing patterns in the text it has already seen.

Separately, we noticed a persistent issue with the classification of German ethnicity/-national identity. Specifically, earlier versions of the prompt frequently labeled individuals as "German" based solely on their place of residence in Germany, even when no nationality was explicitly stated. To correct for this, we experimented with multiple phrasings and conditional rules in the prompt to instruct the model not to infer nationality from residence alone. We added a "German recoding" step that uses an additional prompt to address this issue and ensure the German category in the *Ethnicity* variable is correctly coded. That step

required repeated testing to eliminate systematic over-classification of German identity and ensure closer alignment with our coding standards.

A.1.5 OpenAI API Usage

We use the OpenAI Batch API with gpt4-o (gpt-4o-2024-08-06) model. We break down our data in batches of 5/10/or 20k rows. Batches were broken down according to the rate limits of the Batching API, allowing 50000 requests at a time and batch input files up to 200 MB. Batch processing was not done randomly, but in a sequential order by date.

A.1.6 Performance Metrics

In this section, we report for each category the performance metrics (precision, recall, and F1 score) of our main model (GPT-4o). Mainly, precision tells you how often the model's positive predictions are correct, and recall shows how many of the true positive cases the LLM manages to retrieve. The F1-score combines precision and recall by using their harmonic mean to give a single, well-balanced measure that is high only when both metrics are high. Finally, support gives the number of actual examples for each class in the evaluation set.

We report classification performance metrics based on the full set of human-annotated press releases used for evaluation. For the variables *Multiple Events* and *Non-German*, as also in our analysis, we exclude (1) traffic incidents and (2) releases that do not describe a crime, as identified through the *Type of Crime* classification. For both the *Type of Crime* variable and the *Non-German* variable, we also exclude multiple events. This is done to ensure that each observation corresponds to one single crime event.

Overall, the models achieve high weighted-average precision, recall, and F1-scores across all categories. This indicates a strong and well-balanced classification performance. The relatively lower precision for *Multiple Events* category is likely the case because it underlines the differences in the understanding of the label *Multiple Events* between the human coder and GPT. The large language model tends to use a broader definition of what "multiple events" are, namely, multiple crimes described in the same press release. However, human coders only labeled cases where, in the press release, it was clear that multiple releases had been copied and pasted into a single entry. For our task, using this variable to filter out reports with multiple events, *recall* would be the more important metric to consider. In our case, recall is high; we catch nearly all true multiple events, even if slightly overinclusive.

Table A.1: Performance Metrics, Main Model (GPT 4-o)

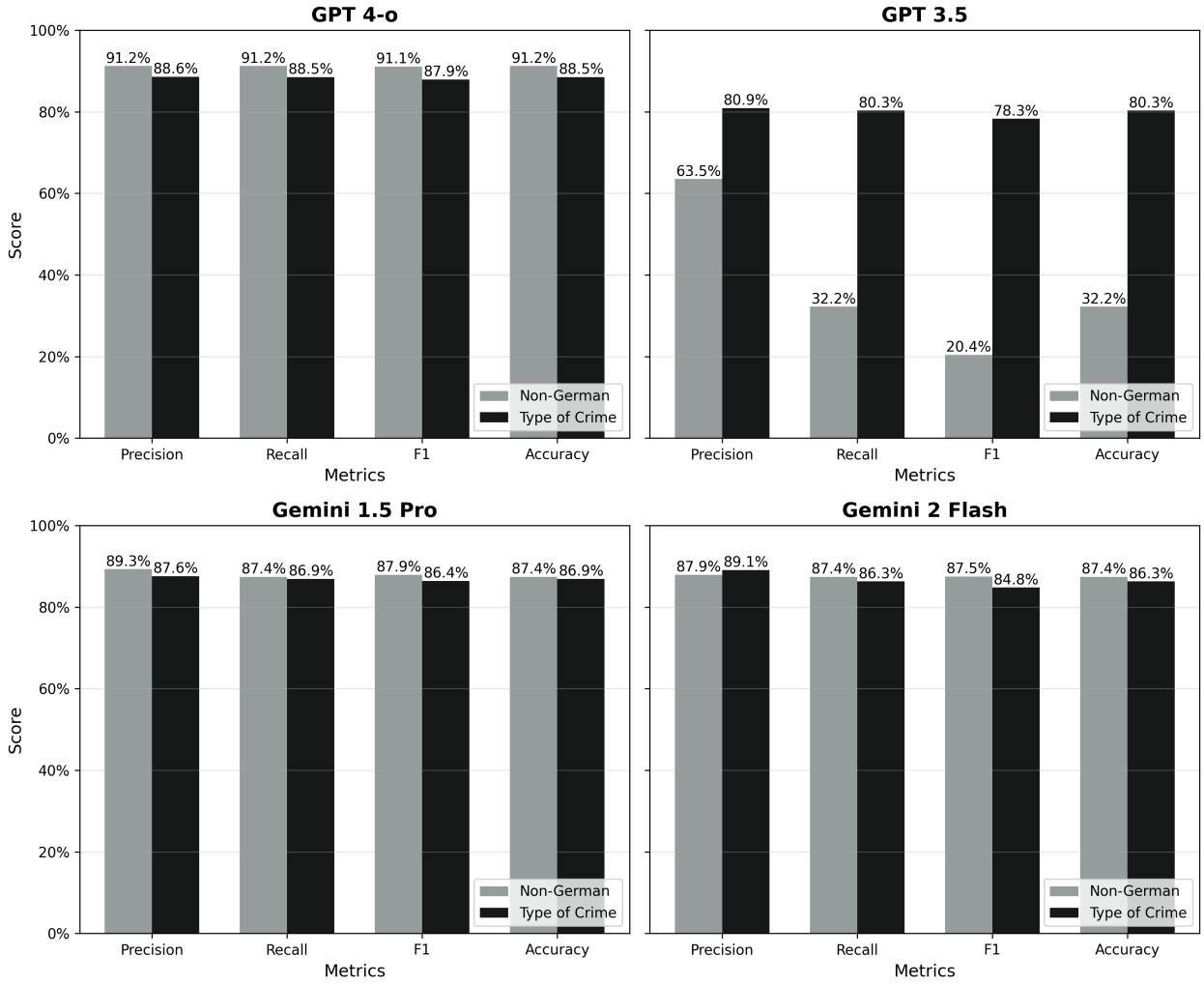
	precision	recall	f1-score	support
Type of Crime				
Violent	0.907	0.916	0.912	546.000
Property	0.879	0.974	0.924	1835.000
Other	0.726	0.768	0.747	280.000
None	0.898	0.564	0.693	624.000
Traffic	0.921	0.933	0.927	1074.000
macro avg	***0.866***	***0.831***	***0.840***	***4359.000***
weighted avg	***0.886***	***0.885***	***0.879***	***4359.000***
Multiple Events				
No	0.994	0.866	0.925	3042.000
Yes	0.465	0.954	0.626	373.000
macro avg	***0.729***	***0.910***	***0.775***	***3415.000***
weighted avg	***0.936***	***0.875***	***0.892***	***3415.000***
Non-German				
No Information	0.941	0.947	0.944	1443.000
No Ethnicity or Nationality	0.882	0.834	0.857	752.000
German	0.859	0.807	0.832	83.000
Foreign	0.866	0.964	0.912	334.000
macro avg	***0.887***	***0.888***	***0.886***	***2612.000***
weighted avg	***0.912***	***0.912***	***0.911***	***2612.000***

A.1.6.1 Benchmarking against other LLMs

We have tested the performance of our prompt with several large language models: GPT 3.5, Gemini 2 Flash, and Gemini 1.5 Pro. GPT 4-o outperforms these models in terms of classification performance, with a more comparable performance to newer Gemini models. GPT 3.5 underperforms in comparison to these models, which is, as mentioned, consistent with work that underlines GPT 3.5's limitations, especially the less sophisticated instruction-following capabilities (Achiam et al. 2023; Zheng et al. 2023).

In figure A.1 below, we report classification performance metrics based on the full set of human-annotated press releases used for our evaluation. For the variables *Multiple Events* and *Non-German*, as also in our analysis, we exclude (1) traffic incidents and (2) releases that do not describe a crime, as identified through the *Type of Crime* classification. Both for the *Type of Crime* variable and *Non German* variable, we exclude multiple events. This is done to ensure that each observation corresponds to one single crime event.

Figure A.1: Classification Metrics, Comparison Across LLMs



In the following tables, we also report the breakdown of performance for each category of the models classified.

Type of Crime:

Table A.2: Classification Report (Type of Crime, GPT 3.5)

	precision	recall	f1-score	support
Violent	0.846	0.923	0.883	573.000
Property	0.775	0.982	0.866	2040.000
Other	0.521	0.197	0.286	315.000
None	0.883	0.406	0.556	633.000
Traffic	0.895	0.812	0.851	1089.000
macro avg	***0.653***	***0.553***	***0.574***	***4650.000***
weighted avg	***0.809***	***0.803***	***0.783***	***4650.000***

Table A.3: Classification Report (Type of Crime, Gemini 1.5 Pro)

	precision	recall	f1-score	support
Violent	0.938	0.857	0.895	544.000
Property	0.847	0.984	0.910	1823.000
Other	0.587	0.629	0.607	280.000
None	0.928	0.585	0.718	620.000
Traffic	0.938	0.906	0.922	1075.000
macro avg	***0.848***	***0.792***	***0.811***	***4342.000***
weighted avg	***0.876***	***0.869***	***0.864***	***4342.000***

Table A.4: Classification Report (Type of Crime, Gemini 2.0 Flash)

	precision	recall	f1-score	support
Violent	0.937	0.904	0.920	408.000
Property	0.867	0.980	0.920	1416.000
Other	0.436	0.840	0.574	162.000
None	0.960	0.301	0.458	475.000
Traffic	0.955	0.960	0.957	902.000
macro avg	***0.831***	***0.797***	***0.766***	***3363.000***
weighted avg	***0.891***	***0.863***	***0.848***	***3363.000***

Multiple Events:

Table A.5: Classification Report (Multiple Events, GPT 3.5)

	precision	recall	f1-score	support
No	0.970	0.954	0.962	3030.000
Yes	0.665	0.756	0.708	365.000
macro avg	***0.818***	***0.855***	***0.835***	***3395.000***
weighted avg	***0.937***	***0.933***	***0.935***	***3395.000***

Table A.6: Classification Report (Multiple Events, Gemini 1.5 Pro)

	precision	recall	f1-score	support
No	0.993	0.861	0.922	3049.000
Yes	0.456	0.949	0.616	374.000
macro avg	***0.725***	***0.905***	***0.769***	***3423.000***
weighted avg	***0.934***	***0.871***	***0.889***	***3423.000***

Table A.7: Classification Report (Multiple Events, Gemini 2.0 Flash)

	precision	recall	f1-score	support
No	0.997	0.650	0.787	3038.000
Yes	0.244	0.986	0.391	348.000
macro avg	***0.621***	***0.818***	***0.589***	***3386.000***
weighted avg	***0.920***	***0.685***	***0.747***	***3386.000***

Non German:

Table A.8: Classification Report (Non German, GPT 3.5)

	precision	recall	f1-score	support
No Information	0.857	0.004	0.007	1614.000
No Ethnicity or Nationality	0.255	0.667	0.370	806.000
German	0.585	0.348	0.437	89.000
Foreign	0.499	0.978	0.661	357.000
macro avg	***0.439***	***0.399***	***0.295***	***2866.000***
weighted avg	***0.635***	***0.322***	***0.204***	***2866.000***

Table A.9: Classification Report (Non-German, Gemini 1.5 Pro)

	precision	recall	f1-score	support
No Information	0.958	0.925	0.941	1459.000
No Ethnicity or Nationality	0.832	0.742	0.784	740.000
German	0.398	0.892	0.550	83.000
Foreign	0.867	0.943	0.904	318.000
macro avg	***0.764***	***0.875***	***0.795***	***2600.000***
weighted avg	***0.893***	***0.874***	***0.879***	***2600.000***

Table A.10: Classification Report (Non-German, Gemini 2.0 Flash)

	precision	recall	f1-score	support
No Information	0.939	0.893	0.915	1182.000
No Ethnicity or Nationality	0.740	0.813	0.775	497.000
German	0.821	0.719	0.767	64.000
Foreign	0.884	0.958	0.920	215.000
macro avg	***0.846***	***0.846***	***0.844***	***1958.000***
weighted avg	***0.879***	***0.874***	***0.875***	***1958.000***

A.1.7 Intercoder Reliability

We calculate reliability both among human coders and between human coders and GPT[13]. Table A.11 reports the agreement across human coders. Table A.12 compares human coders' agreement with the GPT classification; ideally, GPT's reliability should approach the level of agreement observed among human coders. We calculate two standard measures of intercoder reliability (*ICR*), Cohen's Kappa and Krippendorff's Alpha. We observe that agreement between human coders is high across all the variables. It is also the case that GPT 4-o achieves comparable reliability levels. The agreement between human coders and GPT is lower on the *Multiple Events* variable, consistent with the greater ambiguity in coding this category as underlined in section A.1.6[14].

Table A.11: Intercoder Reliability: Between Human Coders

Variable	Cohen's Kappa	Krippendorff's Alpha
Multiple Events	0.770	0.770
Type of Crime	0.842	0.842
Non German	0.800	0.7999

[13]Tables include all labeled data by research assistants as described in A.1.3.

[14]See discussion on performance of the *Multiple Events* category.

Table A.12: Intercoder Reliability: Human Coders Agreement Against GPT 4-o Coding

Variable	Cohen's Kappa	Krippendorff's Alpha
Multiple Events	0.613	0.607
Type of Crime	0.869	0.869
Non German	0.940	0.940

A.1.7.1 Addressing hallucination

We dropped hallucinations on the non-German variable, which occurred in about 0.25% of cases. We manually checked a subset of these cases. The vast majority related to irrelevant events, such as fires without a perpetrator. The remaining cases were also not crimes, including incidents like birds causing damage or traffic issues. Accordingly, dropping such cases from our data set does not affect our substantive conclusions.

A.1.8 Prompts

A.1.8.1 Main Prompt

Role: You are an expert annotator for Natural Language Processing tasks.
Task: Extract information from a German police press release to code the following variables. Focus on both explicit and implicit details in the text.

Variable 1 – Type of Crime Classification (‘crime_type‘, enum) Identify the type of the most severe crime reported in the press release and classify it using the following categories:
- Violent crime (‘violent‘): Includes homicide, assault, robbery, sexual assault, and kidnapping.
- Property crime (‘property‘): Includes theft, burglary, arson, and vandalism.
- Traffic-related incidents (‘traffic‘): Includes driving without a license, illegal street racing, hit and run, accidents, and drunk driving.
- Other (‘other‘): Includes all other crimes that do not fall into any of the above categories, such as drug-related crime, white-collar crime, and any other type of crime that does not fall under the above categories.
- Not a crime (‘none‘): use this category for press releases that do not describe criminal activity. For ambiguous cases that do not neatly fit into the above categories, use your best judgment based on the context given in the press release. For example, if a crime involves both property damage and violence, but the violence is the focus of the report, classify it as ‘violent‘.

Variable 2 - Ongoing Police Investigation (‘investigation_type‘, enum)
Investigation ongoing (‘ongoing‘): if the investigation is still ongoing or if the police are still looking for information about the suspects or perpetrators.
Investigation concluded (‘concluded‘): if the investigation has been concluded or if the police have apprehended the suspects or perpetrators.
Not applicable (‘na‘): does not apply, for example, because the press release does not describe any criminal activity.

Variable 3 – Suspects’ or Perpetrators’ Ethnicity and Nationality Classification (‘eth_or_nationality‘, enum):
- No information (‘no_information‘): if the police do not have any information about

the suspects or perpetrators and thus cannot describe their physical appearance. Do not use this category if the suspects or perpetrators have been arrested or punished.
- No Ethnicity or Nationality provided (‘no_eth_or_nationality‘): if the suspects or perpetrators are known or their physical appearance is described, but no information about their nationality or ethnic background is provided. Do not use this category if the description of the suspects’ or perpetrators’ physical appearance hints at German or foreign ethnic origin, as described below.
- German (‘german‘): If the suspects or perpetrators are described or identified as German nationals (Deutscher, Deutsche), if there are hints of German ethnicity, stereotypical German physical characteristics, or native command of the German language. Look for clear mentions of German nationality or indicators of German ethnicity relating to physical characteristics, speaking the German language without accent, and skin complexion. Important: Classify someone as German if stereotypical physical features such as fair skin, blond (including variations of blond) or light brown hair, or blue or green eyes are mentioned. Anyone who is described as looking Caucasian, Western, or European must be coded as German. I also want you to code suspects or perpetrators as German if any of the following stereotypical linguistic features are mentioned: speaking proper German (for example “Hochdeutsch”), German with German-regional accents (for example “süddeutscher Akzent”), or accent-free German. Here is a non-exhaustive list of keywords that indicate German origin and must always be coded as German without exception: "Deutsch," "Deutscher," "Deutsche Staatsangehörige," "Deutsches Erscheinungsbild," "akzentfrei," "Hochdeutsch," "blond," "blaue Augen," "weiße Hautfarbe","mittelblonde Haare," etc.
IMPORTANT NOTE: Residence in Germany alone does not confirm German ethnicity or nationality. UNDER ABSOLUTELY NO CIRCUMSTANCES infer German ethnicity or nationality solely because residence in a German city/Germany is mentioned. IGNORE place-based identifiers (such as "Bergneustädter," "Kölner," "Wiehler") unless there are other indicators of ethnicity or nationality as described above. If multiple suspects or perpetrators are mentioned and at least one is non-German, code Variable 3 as foreign.
- Foreign (‘foreign‘): if the suspects or perpetrators are described or identified as foreign nationals or of non-German ethnicity. Look for explicit mentions of non-German nationality or implicit indicators of foreign background. Relevant keywords include "unbekannter Akzent," "ausländisches Aussehen," "gebrochenes Deutsch," "ausländischer Akzent," "dunkle Hautfarbe," "osteuropäische Erscheinung,"

"Südländer," "Kopftuch," "afrikanischer Herkunft," etc. Consider these keywords as indicators, especially when explicit nationality information is missing. If multiple suspects or perpetrators are mentioned and at least one is non-German, code Variable 3 as foreign.

Additional Instructions for Variable 3 (‘eth_or_nationality‘): When determining the suspects’ or perpetrators’ ethnicity or nationality, pay close attention to specific keywords or phrases that could indicate German or non-German origins. Remember, these keywords are illustrative and not exhaustive. When direct information is not provided, they serve as indicators to help you infer the suspects’ or perpetrators’ ethnicity or nationality. Use these keywords in conjunction with the context provided in the press release to make informed coding decisions.
IMPORTANT REMINDER: DO NOT USE place-based identifiers (such as "Berliner", "Hamburger", "Bayer", etc.) to infer German ethnicity or nationality.
If multiple suspects or perpetrators are mentioned and at least one is non-German, code Variable 3 as ‘foreign‘.

Variable 4 – Multiple events (‘mult_events‘): If the press release describes multiple crime events, code 1, otherwise, code 0.

Variable 5 – Keyword Identification (‘eth_keywords‘, list[str]): If German or non-German status is inferred, list the main keywords from the text that influenced your decision. Avoid general crime-related terms. Focus solely on keywords or phrases that provide insights into the suspects’ or perpetrators’ ethnic background or nationality. This may include references to language, nationality, specific ethnic descriptors, or other culturally or geographically specific identifiers.

Variable 6 – Nationality (‘nationality‘, list[str]): If specified, record the suspects’ or perpetrators’ nationality. If multiple nationalities are mentioned, note all. If no information is given, give an empty list.

Data Output Format: Here is a press release. You MUST return the classifications for the press release in the VERY STRICT format ‘Press release 0: ["crime_type", "investigation_type", "eth_or_nationality", "mult_events", (eth_keyword1, eth_keyword2, ..) or (NA), (nationality1, nationality2, ...) or (NA)]’

STRICTLY ADHERE TO THIS FORMAT. Only return this output and NOTHING ELSE. **In cases where the press release describes multiple crime events, code the most severe crime.**

A.1.8.2 Prompt for Improving German Classification:

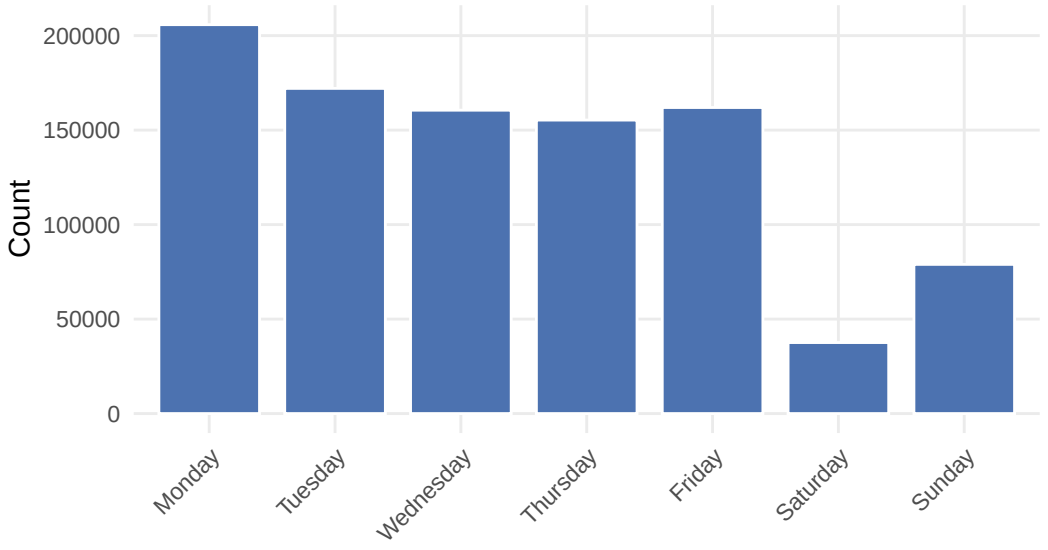
Task: Analyze the following German police press release to determine if the **only** marker of German ethnicity or nationality is a place-based identifier (such as "Berliner", "Hamburger", "Potsdamer", etc.).
Instructions: If the **only** marker of German ethnicity or nationality in the text is a place-based identifier, return 1. If there are other markers of German ethnicity or nationality (e.g., explicit mentions like "Deutscher", "Deutsche Staatsangehörige", "blond", "blaue Augen", "weiße Hautfarbe", "akzentfrei", "Hochdeutsch", etc.), return 0. Do **not** consider place-based identifiers as valid indicators of German ethnicity or nationality on their own. Do **not** infer German ethnicity or nationality solely based on residence in a German city or mentions of living in Germany. Your output should be **either** 1 **or** 0 **and nothing else**.

A.2 Descriptive statistics

Table A.13: Summary Statistics by Year

	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
N	77150	82850	93728	96479	108135	100261	89523	90176	94046	89382	52099
Type of Crime											
Property	0.74	0.75	0.71	0.70	0.69	0.69	0.68	0.67	0.69	0.69	0.69
Violent	0.19	0.18	0.21	0.21	0.20	0.20	0.20	0.19	0.20	0.20	0.21
Other	0.07	0.07	0.07	0.09	0.10	0.11	0.13	0.14	0.12	0.11	0.10
Ethnicity											
Foreign	0.11	0.10	0.13	0.10	0.10	0.09	0.08	0.07	0.06	0.06	0.06
German	0.02	0.02	0.02	0.02	0.03	0.03	0.02	0.02	0.02	0.02	0.02
No Info / Other	0.87	0.88	0.85	0.87	0.88	0.88	0.90	0.91	0.92	0.92	0.92
Concluded (bin.)	0.15	0.13	0.15	0.16	0.16	0.16	0.16	0.16	0.14	0.14	0.14
Multiple Events (bin.)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Federal Police (bin.)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Figure A.2: Number of press releases by weekday

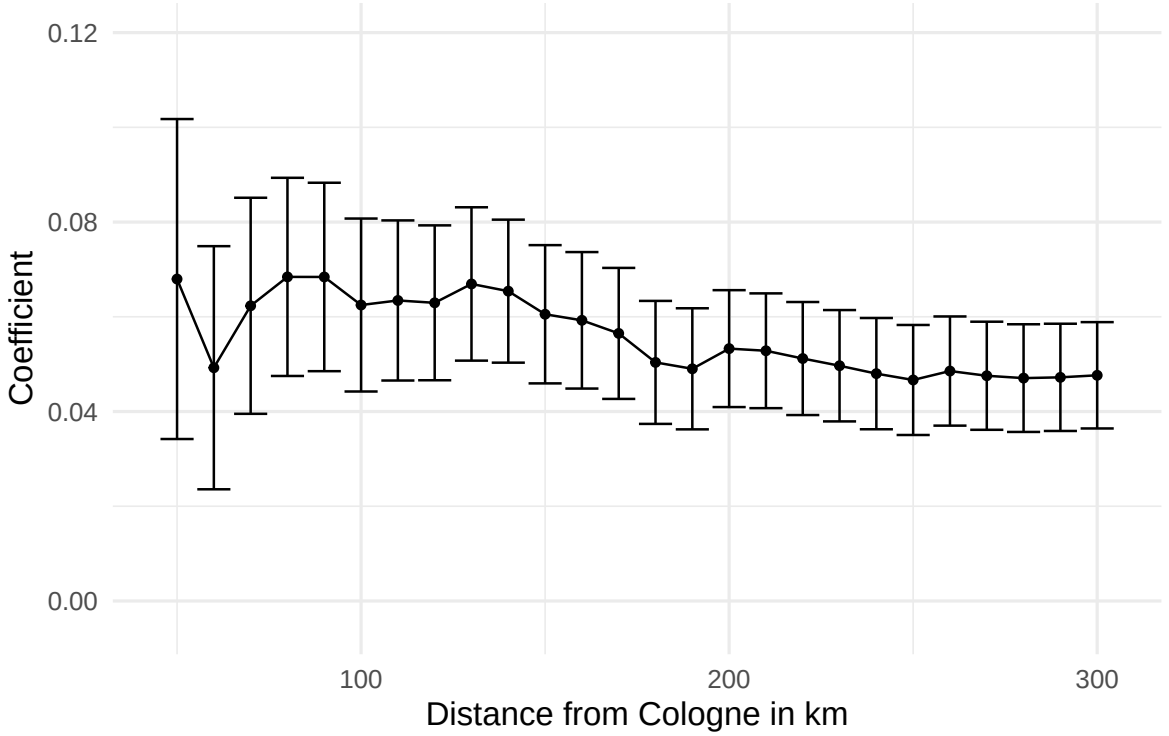


Notes: The figure shows the number of police press releases published by weekday.

Figure A.3: Press releases over time, across police stations and states. *Note:* Each row represents a given police station in our dataset, and a tile represents whether or not a given police station published at least one press release that month.

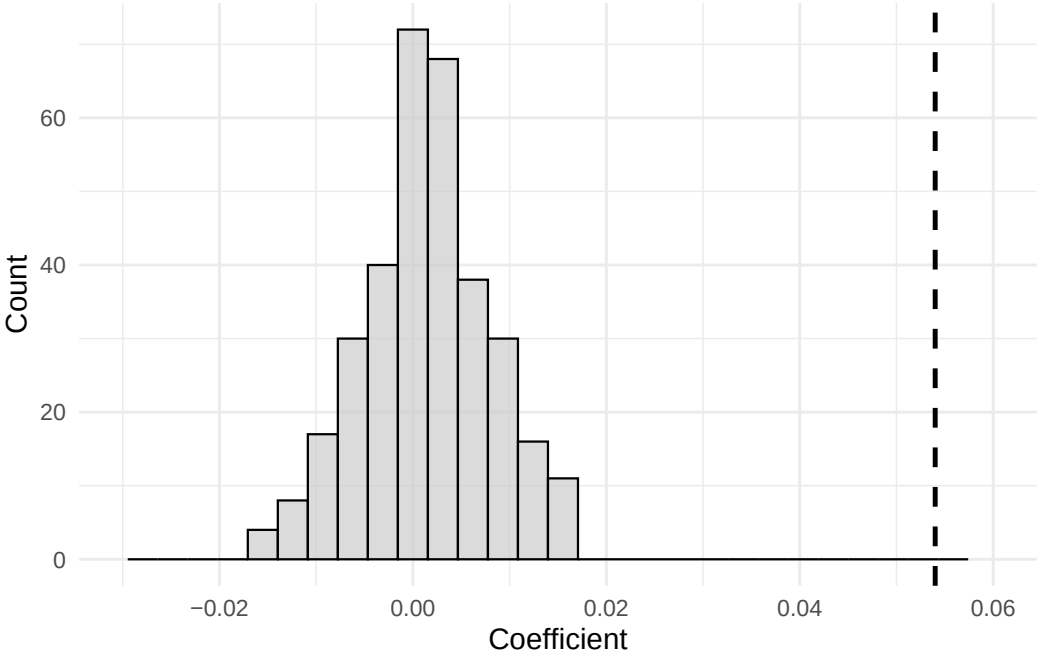
A.3 Additional results: Study I (2015/16 New Year's Eve)

Figure A.4: Distance from Cologne



Notes: We re-estimate the specification from column (1), subsetting the data to all observations within *dist* km from Cologne. The radius around Cologne is indicated on the x-axis. To calculate radii, we use the distance between the centroids for each municipality and the centroid of Cologne. The y-axis shows the respective coefficient estimate on that subset with corresponding 95% confidence intervals.

Figure A.5: Placebo treatments test



Notes: Distribution of coefficient estimates. We re-estimate the specification from column (1) using every day between January 1, 2015, and November 30, 2015, as placebo cut-offs.

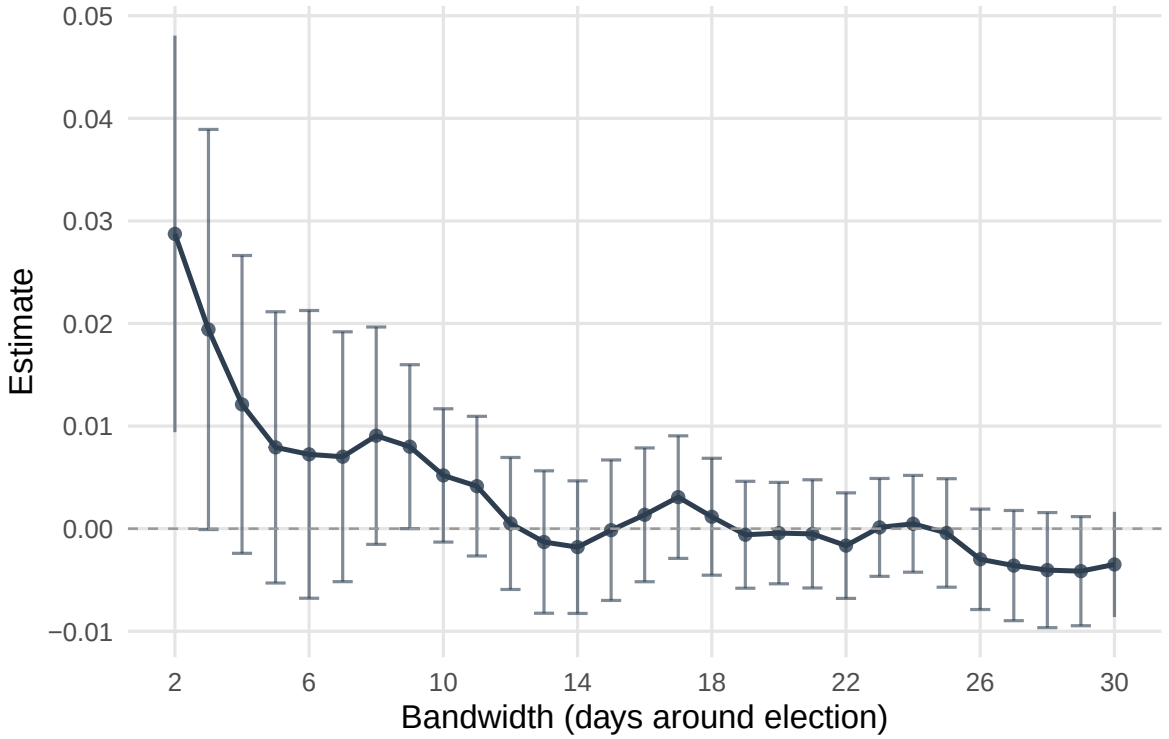
Table A.14: Other ethnicity cues, pre/post NYE 2015/16

	In-group (German) cue (1)	No ethnicity info. (2)
Post-NYE	0.002	-0.050***
	(0.002)	(0.006)
R^2	0.025	0.092
Observations	16,172	16,172
Police station fixed effects	✓	✓

Notes: Results from OLS regressions where the outcomes are binary indicators for (1) German in-group cues and (2) no cues about nationality/ethnicity in police press releases. Press releases are the unit of observation. *Post-NYE* is a binary indicator equal to one for press releases issued after NYE 2015. The analysis uses a bandwidth of ±30 days around NYE. Standard errors are clustered by police station. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

A.4 Additional results: Study II (Elections)

Figure A.6: Results for different bandwidths



Notes: The figure shows the results from the main specification for varying bandwidths between 2 and 30 days.

Table A.15: Other ethnicity cues

	In-group (German) cue (1)	No ethnicity info. (2)
Pre-Election	-0.0003	-0.029**
	(0.006)	(0.013)
R^2	0.018	0.037
Observations	2,218	2,218
State x Election fixed effects	✓	✓

Notes: Results from OLS regressions where the outcome variables are binary indicator for the presence of (1) German in-group cues or (2) no cues about nationality/ethnicity in a police press release. Police press releases are the unit of observation. Pre-election is a binary indicator that equals one for press-releases issued before a given state election. We use a bandwidth of 2 days around state elections. Standard errors are clustered at the state-election level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table A.16: Immigration salience estimated with additional survey rounds

	Out-group cue (0/1) (1)
Pre-Election	0.028***
	(0.008)
Pre-Election × Mig. salience	0.014*
	(0.008)
R^2	0.036
Observations	2,218
State x Election fixed effects	✓

Notes: Results from OLS regressions where the outcome variable is a binary indicator for the presence of out-group cues in a police press release. Police press releases are the unit of observation. Pre-election is a binary indicator that equals one for press-releases issued before a given state election. We use a bandwidth of 2 days around state elections. We estimate immigration salience using data from the four most recent waves of the Politbarometer survey conducted in each state prior to its respective state election, including survey weights. Both moderator variables are measured in standard deviations. Standard errors are clustered at the state-election level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table A.17: Difference in Daily Report Counts on Pre-Election Fridays/Saturdays

	Daily count of press releases		
	(1)	(2)	(3)
Pre-Election Fri/Sat (0/1)	0.767	1.50	1.50
	(3.01)	(2.11)	(1.79)
Observations	14,797	14,797	14,797
State fixed effects		✓	✓
Year fixed effects		✓	✓
Weekday fixed effects			✓

Notes: Results from an OLS regression where the outcome variable is the total count of police press releases on a given day. The unit of observation is a state-day (Friday or Saturday). The analysis compares the Fridays and Saturdays immediately before a state election to all other Fridays and Saturdays in the sample period. Pre-Election Fri/Sat is a binary indicator equal to one for the days immediately preceding an election in that specific state. The model includes state, year, and day-of-week fixed effects. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table A.18: Heterogeneity by party of state minister of the interior

Party of interior minister	Out-group cue (0/1) CDU/CSU (1)	SPD (2)
Pre-Election	0.037**	0.024*
	(0.013)	(0.013)
R^2	0.013	0.047
Observations	808	1,410
State x Election fixed effects	✓	✓

Notes: Results from OLS regressions where the outcome variable is a binary indicator for the presence of out-group cues in a police press release. Police press releases are the unit of observation. Pre-election is a binary indicator that equals one for press-releases issued before a given state election. We use a bandwidth of 2 days around state elections. The sample is split by party membership of the state interior minister at the time of a given state election. Standard errors are clustered at the state-election level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table A.19: Heterogeneity by competitiveness of a given election

	Out-group cue (0/1)			
	(1)	(2)	(3)	(4)
Pre-Election	0.029	0.033	0.141*	-0.070
	(0.017)	(0.218)	(0.079)	(0.067)
Pre-Election × Margin of victory (winner minus runner-up, in p.p.)	4.38×10^{-5}			
	(0.002)			
Pre-Election × Competitiveness index (1 – MoV)		-0.004		
		(0.232)		
Pre-Election × Herfindahl–Hirschman Index of vote-share concentration			-0.494	
			(0.329)	
Pre-Election × Effective Number of Parties (1 / HHI)				0.022
				(0.016)
R^2	0.035	0.035	0.036	0.036
Observations	2,218	2,218	2,218	2,218
State x Election fixed effects	✓	✓	✓	✓

Notes: Results from OLS regressions where the outcome is a binary indicator for the presence of out-group cues in a police press release. The unit of observation is the individual press release. *Pre-election* equals 1 for releases issued in the two days before the relevant state election; releases on election day are excluded. The analysis uses a bandwidth of ±2 days around state elections. Moderators are election-level measures of electoral competitiveness computed from official vote shares across seven categories (including a residual *Other* so shares sum to 100): (i) the winner–runner-up vote-share gap in percentage points; (ii) a normalized competitiveness index defined as one minus that margin divided by 100; (iii) the Herfindahl–Hirschman Index, the sum of squared proportional party vote shares; and (iv) the Effective Number of Parties, the inverse of the Herfindahl–Hirschman Index. All models include state×year fixed effects and standard errors clustered at the state–year level. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.20: Type of crime

	Other (0/1) (1)	Property crime (0/1) (2)	Violent crime (0/1) (3)
Pre-Election	0.014	-0.023	0.010
	(0.013)	(0.021)	(0.018)
R^2	0.018	0.021	0.025
Observations	2,218	2,218	2,218
State x Election fixed effects	✓	✓	✓

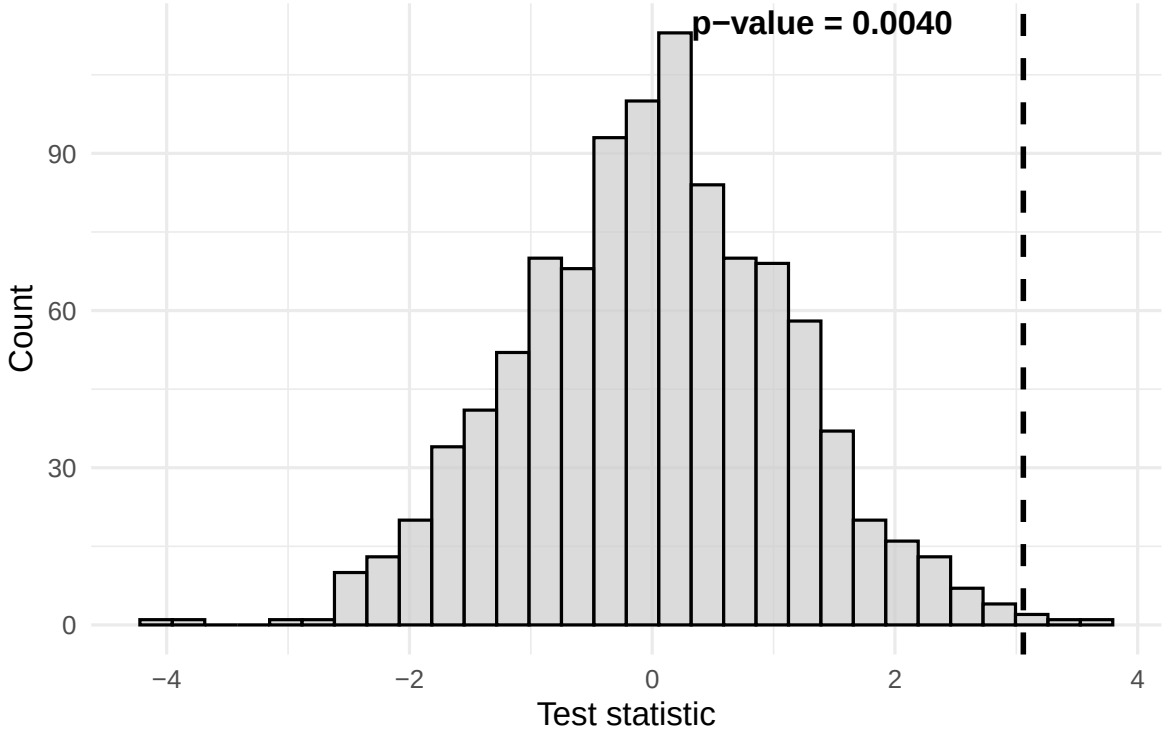
Notes: Results from OLS regressions where the outcome variables are binary indicators for different types of crime in a police press release: (1) other crimes, (2) property crimes, (3) violent crimes. Police press releases are the unit of observation. Pre-election is a binary indicator equal to one for press releases issued before a given state election. Standard errors are clustered at the state-election level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table A.21: Heterogeneity by local politics

	Out-group cue (0/1) (1)	(2)
Pre-Election	-0.016	0.144***
	(0.019)	(0.046)
Pre-Election × AfD vote share	0.004**	
	(0.002)	
Pre-Election × Left combined vote share (SPD + Greens + The Left)		-0.003**
		(0.001)
R^2	0.037	0.040
Observations	2,189	2,189
State x Election fixed effects	✓	✓

Notes: Results from OLS regressions where the outcome is a binary indicator for the presence of out-group cues in a police press release. The unit of observation is the individual press release. *Pre-election* equals 1 for releases issued in the days before a given state election; releases on election day are excluded. The analysis uses a bandwidth of ±2 days around state elections. Each column reports a separate interaction between *Pre-election* and either AfD vote share or the Left combined vote share (SPD + Greens + The Left), with all measures taken from the 2017 federal election at the municipality level (p.p.). Data on municipality-level vote shares was obtained from GERDA – German Election Database. All models include state×year fixed effects and standard errors clustered at the state–year level. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure A.7: Placebo treatments test



Notes: Distribution of placebo t-statistics. For each of 1,000 iterations, we replace the actual election date of every state with a *Sunday* drawn at random from the sample period. We then re-estimate the main specification in the ±2-day window around these placebo cutoffs and record the t-statistic on the `post` indicator. The histogram displays the resulting distribution; the dashed red vertical line marks the true-election t-statistic. The two-sided randomization p-value is the proportion of placebo statistics whose absolute value is at least as large as the true statistic.

Table A.22: Placebo outcome test – press releases issued by federal police

	Out-group cue (0/1)		
	(1)	(2)	(3)
Pre-Election	-0.059	-0.114	-0.124
	(0.082)	(0.110)	(0.096)
R^2	0.002	0.341	0.232
Observations	125	96	114
Police station x Election fixed effects		✓	
State x Election fixed effects			✓

Notes: Results from OLS regressions where the outcome is a binary indicator for the presence of out-group cues in a police press release. The unit of observation is the individual press release. *Pre-election* is a binary indicator equal to one for releases issued before a given state election. The analysis uses a bandwidth of ±2 days around state elections. This analysis examines press releases issues by the German federal police (extitBundespolizei). Standard errors are clustered at the state-election level. Significance codes: *** p<0.01, ** p<0.05, * p<0.1.

A.5 Police Communication and Trust in Police

To conduct an exploratory test for the association between police communication and public attitudes, we link our data on out-group reporting in police press releases to individual-level survey data from Hirndorf et al. (2023), collected in 2020. Table A.23 presents results from linear probability models predicting a binary outcome: whether the respondent reports high trust in the police.[15] We correlate this survey item with the share of press releases containing out-group cues at the county level in 2020. All models control for individual demographics, the true share of foreign crime offenders (*Polizeiliche Kriminalstatistik (PKS)*), and include state fixed effects. Standard errors are clustered at the county level. The analysis is restricted to respondents without a migration background.

Table A.23: Out-Group Crime Reporting and Public Perceptions

	Trust in Police (binary)	
	(1)	(2)
Share of out-group press releases	0.036	0.037
	(0.055)	(0.054)
True foreign crime share (PKS)	-0.005	0.0008
	(0.011)	(0.015)
Crime Clearance Rate (PKS)		-0.015
		(0.018)
Total Crimes in County (PKS)		-0.009
		(0.010)
R^2	0.025	0.025
Observations	2,063	2,063
State fixed effects	✓	✓

Notes: OLS linear probability models at the respondent level (sample restricted to respondents without migration background); the outcome is binary – High trust in police coded 1 for 'sehr großes/großes Vertrauen'; Share of out-group reports = 2020 county share of police press releases containing out-group cues; all columns include state fixed effects and controls for age, gender, and education; SEs clustered by county; Column (1) additionally controls for the overall county-level share of foreign crime suspects (PKS 2020); Column (2) further adds the county crime clearance rate (Aufklärungsquote) and total crimes (PKS 2020); all PKS variables are standardized; significance: *** p<0.01, ** p<0.05, * p<0.1.

[15] "Please tell me ... how much trust you have in [the Police]: very high trust, high trust, little trust, or no trust at all?"