I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17888

Identity Under Scrutiny: Media Attention and Rule Compliance

Sulin Sardoschau ®
Giorgio Gulino ®
Federico Masera ®

MAY 2025

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17888

Identity Under Scrutiny: Media Attention and Rule Compliance

Sulin Sardoschau
Humboldt Universität zu Berlin, RFBerlin and IZA

Giorgio Gulino
Tor Vergata University of Rome

Federico Masera
University of New South Wales

MAY 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Identity Under Scrutiny: Media Attention and Rule Compliance*

How does media coverage of minorities affect their rule compliance? Using data from 800,000 random audits at supermarket self-checkouts in Italy, we show that heightened refugee media coverage reduces under-reporting of items among shoppers born in major refugee-source countries, but not other migrants or natives. The effect is concentrated in the seven days following media exposure and is strongest when coverage is negative or highlights criminality. Results are not driven by changes in customer composition or perceived audit risk. Instead, our findings suggest that public scrutiny prompts minorities to counter negative stereotypes by increasing their compliance.

JEL Classification: D74, J15, D83, Z10, D72
Keywords: immigration, refugees, identity, crime, media, threat

Corresponding author:
Giorgio Gulino
Department of Economics and Finance
Tor Vergata University of Rome
Via Columbia 2, 00133 Rome, Italy
E-mail: giorgio.gulino@uniroma2.it

* Certified random order of the authors (Ray and Robson, 2018). We thank Filipe Campante, Elena Cettolin, Ruben Enikolopov, Giovanni Mastrobuoni, Paola Giuliano, Evan Rose, Moses Shayo, Francesco Sobbrio, Alessandra Voena, and participants at the Berlin Applied Micro & Behavioral Economics Workshop, Rome-Bologna Applied Economics Seminar Series for insightful comments. Financial support from the CRC 190 Rationality and Competition is gratefully acknowledged. All errors are our own.

1 Introduction

Public attention toward migrants in Western countries has risen substantially over recent decades, often explicitly or implicitly linking immigration to crime and illegality (UNHCR, 2015; Guriev & Papaioannou, 2022). An extensive literature documents how heightened salience of immigration, driven by political rhetoric or media coverage, influences attitudes and behaviors among native populations (Brader et al., 2008; Hopkins, 2010; Bursztyn et al., 2017; Müller & Schwarz, 2023; Djourelova, 2023; Couttenier et al., 2024; Schneider-Strawczynski & Valette, 2025). Yet, considerably less is understood about how migrants themselves respond to this public attention.

In this paper, we investigate how increased media coverage of refugees, especially when it is negative or focused on crime, affects the rule compliance of individuals born in major refugee-source countries. Social psychology identifies two countervailing behavioral responses among groups facing increased public scrutiny. On the one hand, stigmatizing public attention can erode commitment to norms or lower the perceived reputational costs of deviance, thereby increasing rule violations (Becker, 1963; Lemert, 1951; Belmi et al., 2015). On the other hand, stigmatizing public attention may increase the perceived risk of detection and punishment or trigger identity-threat responses (Steele et al., 2002) that lead individuals to disconfirm negative stereotypes by demonstrating heightened rule compliance (Kray et al., 2001; Rudman & Fairchild, 2004). Thus, the effect of increased public attention on minorities' rule compliance is ex-ante ambiguous.

We examine this in the context of Italy between 2017 and 2019, a period characterized by heightened media coverage of refugees and a populist government in power, notably featuring Matteo Salvini as Minister of the Interior, known for his high-profile anti-refugee rhetoric. We measure media attention with the number of migration-related news articles published by Italy's leading national daily newspaper *Corriere della Sera* and retrieve information on their sentiment and content.[1] The peaks of this index primarily correspond to news coverage of border crossings, refugee boat arrivals, and search-and-rescue operations.

We measure rule compliance by tracking the under-reporting of items at supermarket self-service checkouts, using shopping trip level information on over 800,000 random audits conducted in 36 Italian supermarkets over three years.[2] We focus on under-reporting – instances where the total value of reported items is lower than the actual value of items in a customer's shopping basket – for two reasons. First, under-reporting is common (about 13% of audits find at least one under-reported item) and measured

[1] We use the near-universe of articles extracted from the Factiva news archives. We first filter these articles by migration-related keywords and subsequently employ a Large Language Model (ChatGPT-4o) to classify articles that focus explicitly on migration. We also use alternative measures of media attention by looking at coverage of other major and local newspapers and with migration-related Google search volumes.

[2] We use high-frequency data by the supermarket chain Coop Alleanza, using random audits in the provinces of Ferrara and Modena between January of 2017 and January of 2020. This dataset comprises detailed customer characteristics (including place of birth) and tracks customer behavior with self-service checkout technology. A subsample of these data was used previously in Gulino & Masera (2023), who analyzed the effect of political corruption scandals at the local level on under-reporting among all shoppers.

with sufficient granularity, allowing us to identify immediate behavioral responses to surges in media coverage. Second, the associated material stakes are relatively low: there are no fines (shoppers only need to pay the difference) and refugee-origin customers are typically well-integrated, long-term residents with no risk of deportation or serious legal consequences from under-reporting.[3] This allows us to focus on the role of social identity and limit the influence of material considerations in determining rule compliance.

Using a difference-in-differences design, we compare changes in under-reporting among migrants relative to locals following increased media coverage of immigration. We find that heightened media attention significantly reduces under-reporting among immigrants from top refugee-origin countries but does not affect any other migrant groups (international migrants or internal migrants). The effects are short-lived but large in size. A one standard deviation increase in media attention in the 7 days preceding the shopping trip reduces the incidence of under-reporting by 1.2 percentage points, about 10% relative to the mean. The effect is driven by the top quartile of news salience, suggesting that media attention must generate sufficient public awareness to trigger behavioral responses among minority groups.

Our identification strategy assumes that, absent heightened media salience, there are no differential changes in under-reporting between refugee-origin customers and locals. To address potential violations, such as migrants preemptively adjusting behavior in anticipation of rising anti-minority sentiment, we examine the dynamic treatment effect of media salience. Future surges in news coverage do not impact compliance in preceding weeks. Instead, only salience up to seven days prior to a shopping trip lowers under-reporting for refugee-nationality customers, assuaging concerns about anticipation effects and mitigating the concern that our estimates reflect persistent effects of serially correlated past shocks.

To test whether responses are stronger when coverage is migration-focused, negative, or linked to crime, we retrieve information on the sentiment and content of news articles. In a first step, we compare our baseline results derived from articles classified by the LLM as primarily migration-focused to the residual category of articles mentioning migration only tangentially. The effect of explicitly migration-focused news is three times larger and estimated with greater precision, suggesting that refugee-origin shoppers respond specifically to targeted media attention. Second, we use an LLM to categorize news article sentiment into negative, neutral, or positive. Surges in negative public attention towards migration produce much larger effects. Negative-sentiment coverage reduces under-reporting by more than 3 times the effect of non-negative coverage. Third, we examine content directly, identifying articles explicitly linking migration to crime, illegality, or deviance. Consistent with expectations, refugee-origin shoppers' compliance responds most strongly to crime-related salience. The effect of crime-related articles is almost 5 times larger than other forms of migration coverage.

[3]The average refugee shopper in our sample has been shopping at Coop supermarkets for 10 years on average. 98% of shoppers are employed or students and 36% are women. The average probability of under-reporting does not differ between refugee-origin, Southern-Italian and other foreign-born customers.

Next, we explore alternative response margins and potential selection effects. Refugee-origin shoppers may alter their shopping behavior in response to media salience to minimize social scrutiny or may even stop shopping entirely. Using detailed shopping trip and location information, we do not find that refugee-origin shoppers change other behaviors, such as shopping at less busy hours or traveling to supermarket branches farther away from their location of residence. We also examine whether the composition of shoppers changes during salience shocks, either because of differential auditing or because of a change in shopping habits. The results show no significant differences in shopper characteristics, such as place of birth, socioeconomic traits, or past under-reporting behavior. We also include worker shift fixed effects, controlling for potential influences from specific guards or cashiers on duty during audits and find no changes in our estimated effects.

In addition, we leverage information on over-reporting, which plausibly reflects inattention (potentially through psychological distress) rather than compliance and find no evidence of changes in over-reporting following high salience weeks. These findings suggest that refugees' increased rule compliance occurs spontaneously and specifically at the point of scrutiny, consistent with an immediate psychological response rather than a more premeditated behavioral shift.

Examining group-specific salience, we show that the effect is significantly stronger among male refugee-origin shoppers and among newer supermarket members, consistent with heightened sensitivity to reputational concerns. Male shoppers are more directly targeted by narratives linking refugees to crime, while recent supermarket members likely have weaker claims to local in-group identity and thus face greater pressure to demonstrate conformity.

Finally, we consider whether the decrease in under-reporting could be driven by changes in perceived audit risk. Although audits remain randomly assigned, refugee-origin shoppers might mistakenly believe their likelihood of audit increases during high-salience periods. We examine this possibility leveraging two sources of variation for inference: shoppers' prior experiences with audits during high-salience periods, and the ethnic composition of customers audited at the same store within the hour preceding a given shopping trip. In both scenarios, refugee-origin shoppers might infer a higher audit probability – either due to prior frequent audits or observing audits of other refugee-origin customers – and thus exhibit greater compliance. However, we find no evidence of heterogeneous effects, suggesting that changes in perceived audit probability does not drive our main results.

Our findings are robust across a range of additional specifications and checks, including placebo checks using news unrelated to migration, accounting for unobserved individual-level heterogeneity with shopper fixed effects, accounting for shopping trip characteristics, alternative treatment definitions (such as proxies for the demand side for news through Google searches, or local newspaper and center-left newspaper coverage).

Overall, our evidence underscores how minority groups adjust their behavior under intensified scrutiny, disconfirming negative stereotypes about their in-group. Notably,

the shoppers in our sample are unlikely to be recent refugees themselves; instead, they entered the country many years ago, are middle-income and shop at an Italian supermarket. This suggests that policies and rhetoric framing refugees as deviant or non-compliant may impose psychological pressure even on established, well-integrated minority groups.

Our paper contributes to several strands of the literature. First, we extend research on how media coverage and out-group salience shape majority-group preferences and behavior. Media can influence public attitudes toward stigmatized groups, often linking immigration to crime (Djourelova, 2023; Keita et al., 2024; Couttenier et al., 2024). Salience reactivates latent prejudices, prompting individuals to adopt more extreme positions (Brader et al., 2008; Hopkins, 2010; Cantoni et al., 2019; Schneider-Strawczynski & Valette, 2025); targeted propaganda against minorities can lead to intergroup violence (Yanagizawa-Drott, 2014; Adena et al., 2015). Such salience is often further amplified by populist politicians or local protests (Bursztyn et al., 2020; Müller & Schwarz, 2021; Grosjean et al., 2023; Müller & Schwarz, 2023; Fages & Martínez, 2023; Sardoschau & Casanueva-Artís, 2025). We contribute by showing how media-driven attention not only shapes majority views but also influences minority behavior.

Second, we contribute to the literature on identity threat, assimilation, and minority responses to scrutiny. Prior studies show that discrimination and threat can drive assimilationist behaviors (Bisin & Tura, 2019; Fouka, 2019; Saavedra, 2021; Jaschke et al., 2022) but it can also produce backlash (Fouka, 2020; Glover, 2019; Dahl et al., 2022). We focus on one threat dimension – stigmatizing public attention – to shed light on the dynamics underlying identity threat responses.

Third, our paper also contributes to the economic literature on stereotype threat and identity salience. Negative stereotypes can undermine minorities' performance by increasing cognitive load, anxiety, or by inducing disengagement (Steele, 1997; Bertrand & Duflo, 2017; Hoff & Pandey, 2006; Fryer et al., 2008; Spencer et al., 2016). Economists have documented identity-driven gaps across diverse contexts, demonstrating how salience of ethnic, caste, gender, sexual, religious or racial identities can alter minority behavior and outcomes (Benjamin et al., 2010; Charness & Chen, 2020; Oh, 2023; Badgett et al., 2024). Our study extends this literature by providing empirical evidence that media-driven identity salience affects rule compliance among established minority populations.

Lastly, our research intersects with studies on immigration and crime, particularly how migrants are perceived regarding criminality. While extensive empirical evidence demonstrates that migrants are no more prone to criminal behavior than natives and may even reduce local crime rates (Mastrobuoni & Pinotti, 2015; Pinotti, 2017; Ajzenman et al., 2023; Abramitzky et al., 2024)—public discourse persistently associates immigration with crime, shaping native policy preferences (Dustmann & Preston, 2007; Citrin et al., 1997). Our paper highlights an overlooked dimension: migrants, when subjected to generalized suspicion of dishonesty, significantly increase rule compliance, underscoring a behavioral feedback effect of anti-immigrant portrayals.

2 Background and Data

2.1 Background

Between 2015 and 2017, Italy was a central entry point for refugees and asylum seekers traveling through the Central Mediterranean route, primarily originating from North Africa (UNHCR, 2021). Approximately 119,000 migrants reached Italy's shores in 2017 alone. This migration influx was widely portrayed as overwhelming Italy's infrastructure and resources. Immigration became a primary public concern by 2017, with roughly one-third of Italians listing it as a top national issue, up significantly from earlier years (European Commission, 2018). Italians considerably overestimated the immigrant population size and associated refugees with increased risks of crime and terrorism (Ipsos MORI, 2018) with prominent political figures amplifying these narratives (Camilli, 2019).

Subsequent EU-supported efforts to curtail Mediterranean crossings, including Italy's controversial 2017 Libya Agreement, contributed to a sharp decline in migrant arrivals (UNHCR, 2021). Further restrictive measures, such as Interior Minister Marco Minniti's 2017 NGO "code of conduct" intensified media attention by framing humanitarian rescue efforts as complicit in illegal immigration. The 2018 elections brought a populist coalition of the Five Star Movement and the right-wing League party to power, with Salvini swiftly implementing strict anti-refugee policies. These included closing Italian ports to NGO-operated rescue boats, leading to subsequent dramatic standoffs with rescue boats and arrests of rescue boats' crew members.[4] Salvini also enacted legislation curtailing humanitarian protection and significantly tightening refugee regulations (Camilli, 2019; Geddes et al., 2020). Although Salvini's tenure was brief, these high-profile actions and policy shifts marked a significant and enduring securitized migration stance in Italy.

2.2 Data on migration-related news articles

To measure public attention towards migrants and refugees, we construct an index of media salience using newspaper articles (or Google Searches) related to migration. We use these as proxies for general public attention rather than as direct evidence that customers altered their behavior specifically after reading these articles.

We use data from the FACTIVA archive, which covers the near-universe of articles in *Corriere della Sera*, Italy's most widely read national newspaper. We begin by selecting articles that contain at least one of the following migration-related keywords: *Rifugiato* (refugee), *Migrante/i* (migrant), *Sbarchi* (refugee boat arrivals), *Immigrato/i* (immigrant). This yields 11,311 articles for the 3-years period of our study. To isolate articles primarily concerned with migration, as opposed to those in which migration is mentioned only tangentially, we employ a large language model (ChatGPT-4o)

[4]The most notable cases are the Acquarius boat that was denied entry to Italian ports, leaving over 600 rescued migrants stranded at sea and the impounding of the Open Arms and Sea-Watch rescue boats and the arrests of it crew.

to classify article content. Approximately one-third of the articles are identified as migration-focused. Appendix B provides additional details on the classification procedure. Summary statistics of these articles can be found in Table A1. We complement this dataset with newspaper articles from other national and local newspapers. We also use daily Google Searches for the same migration-related keywords to build a standardized index from the first principal component of these search terms, capturing the overall interest in migration-related news.

Figure A1 plots the media salience index for newspaper articles and Google Searches over time, illustrating that news stories broadly reflect the public interest in these stories through Google Trends.[5] The peaks in both indices coincide with events surrounding the arrival of refugee boats and search-and-rescue missions in the Mediterranean outlined in section 2.1. During this period public debates focused on refugees rather than immigration in general.

2.3 Data from random supermarket audits

We use individual-level data from Coop Alleanza 3.0 supermarkets in the Italian provinces of Modena and Ferrara, for customers who used the "time-saver" self-service checkout system.[6] We leverage information on approximately 800,000 random audits conducted from January 2017 to January 2020 to verify the accuracy of customer-declared values against the true value of their shopping carts.

The audit process follows a randomized structure, where a fixed proportion of customers is selected for audits at checkout. During each audit, customers are required to remove items from their shopping cart, which are then rescanned by a cashier. The value of the rescanned items is compared to the value initially declared by the customer. Discrepancies are categorized as either under-reporting (declared value lower than the actual value) or over-reporting (declared value higher than the actual value). If under-reporting is detected, customers must pay the difference between the declared and actual basket values but face no further penalties, such as supermarket bans or police involvement. Thus, the direct material consequences of under-reporting are low.

Table A2 presents summary statistics for our sample of customers. On average, 13% of audits reveal under-reporting, while 5% involve over-reporting. We consider over-reporting as unintentional, providing a benchmark for assessing accidental errors versus deliberate dishonest behavior. The value of under-reported items is comparatively small, amounting to 4€ on average, which corresponds to about 8% of the basket value. While most audits were randomly assigned, exceptions included increased audit

[5] These are based on Google Search at the national level. When focusing on the region of Emilia-Romagna, a near identical picture emerges. Because of the lower volume of searches, the volatility of the regional time series is larger, and the time series often reaches the bottom coding that Google imposes.

[6] Coop Alleanza 3.0, Europe's largest consumer cooperative with over 2.7 million members, operates approximately 350 stores across eight Italian regions, accounting for 14.5% of the national market share and about 35% of the market share in Emilia-Romagna (home to the provinces of Modena and Ferrara). The cooperative is committed to promoting local identity and supporting Italian produce, ensuring that its offerings reflect the culinary traditions and preferences of their Italian communities.

probabilities for customers with a history of significant under-reporting and rare discretionary audits initiated by security agents. However, these exceptions are minimal, ensuring that the dataset predominantly reflects random sampling.[7]

The dataset encompasses 140,000 unique customers, which we identify through their membership card number, enabling us to track individual shopping behavior over time and across Coop shop locations. Additionally, the dataset includes detailed demographic information, such as age, gender, postal code of residence, and occupation. Table A3 summarizes the demographics of audited customers, benchmarking them against the overall population in the same province. Women and employed individuals are slightly overrepresented in the audited sample compared to the general population.

Most importantly, we have information on the country of birth of the customer and their province of birth for Italian nationals. We distinguish between four groups: locals (those born and residing in Northern Italy), Southern Italians (the rest of the Italian customers), refugee-origin shoppers and other migrants. We define refugee-origin shoppers as individuals born in non-European countries that belong to the top 30 asylum-seeker origin countries in Italy between 2017 and 2019 (see Table A5 for the full list of countries). This focus ensures that we capture the group most directly targeted by media and political rhetoric on refugee inflows (UNHCR, 2015). In robustness checks, our estimates remain unchanged when expanding or narrowing this country list or using the list of countries from refugee boat arrival records.

Overall, about 1.5% of our sample is comprised of refugee-origin shoppers (about 2,200 clients in total and 10,000 audits), other foreigners make up 4%, with locals (80%) and South-Italians (15%) representing the large majority in our sample. Table A4 shows that refugee-origin shoppers have likely been in the country for a long time. The average duration of their Coop membership is 10 years, and 98% of refugee-origin shoppers are either employed or studying. While the data does not specify whether the audited shoppers are asylum seekers themselves, it is unlikely given their employment status and duration of stay. However, media coverage of incoming refugees may remain salient for their compatriots. Notably, there is no significant difference in the average probability of under-reporting between refugee-origin, South-Italian, or other foreign-born customers.

3 Research Design and Results

3.1 Empirical Framework

To examine the effect of out-group salience on the rule conformity of minority groups, we exploit variation at the customer and day level in our sample of random audits to estimate the following linear probability model:

[7] We verify that these discretionary audits do not drive our results by controlling for guards-shifts fixed effects and showing that there is no over-sampling of minority groups by auditors during treated periods. In another analysis, we exclude individuals who have under-reported during their previous audit to alleviate concerns that known under-reporters are audited with a higher probability.

$$Y_{i,t} = \alpha_t + \sum_{j=2,3,4} \gamma_j MigGroup_i + \sum_{j=2,3,4} \beta_j Salience_t \times MigGroup_i + \epsilon_{i,t} \quad (1)$$

where $Y_{i,t}$ is a binary variable equal to 1 if customer i, audited on day t, under-reports, and 0 otherwise. $Salience_t$ measures the salience of the topic of migration. For our baseline specification, we use the count of migration-related newspaper articles over the preceding seven days.[8] For ease of interpretation, we standardize this variable to have a mean of 0 and a standard deviation of 1 (1.92 articles). We differentiate between three migrant groups: shoppers with a refugee nationality (as defined in section 2), shoppers with other foreign nationalities, and shoppers born in Southern Italy. The reference group comprises locals, defined as individuals born and residing in Northern Italy. The model includes day (α_t) and migrant group (γ_j) fixed effects, controlling for systematic differences in under-reporting across groups and temporal change in under-reporting common to all groups, such as national news cycles, elections, or economic fluctuations. Our coefficients of interest, β_j, estimate the differential probability of under-reporting for each group j relative to Northern Italians when out-group salience is 1 standard deviation higher. Standard errors are clustered at the migrant group and day level, with robustness checks confirming the results under alternative clustering approaches.[9]

Our identification strategy rests on the assumption that, absent the salience shock, there are no differential changes in under-reporting between migrant groups and Northern Italians. A potential violation would occur if migrants preemptively adjusted their behavior in anticipation of rising anti-minority sentiment, which coincides with a subsequent increase in refugee news coverage. As illustrated in Figure A1, news coverage of migration is primarily driven by unexpected events such as the arrival of refugee boats or search-and-rescue operations. These events are unlikely to be anticipated by migrant groups, precluding behavioral adjustments prior to the salience shock.

To verify this, we estimate dynamic treatment effects in Figure 1, examining the impact of media salience over multiple time periods before and after the shopping trip.[10] Results show that future salience does not influence present-day under-reporting. Furthermore, only salience in the preceding seven days predicts under-reporting, thus mitigating the concern that our estimates reflect persistent effects of serially correlated past shocks.

[8] The choice of the window is informed by our analysis in Figure 1, where we estimate the dynamics of the effect and show that they are all concentrated in the 7 days after the increase in salience. Nevertheless, we show in Section 3.4 that our results are robust to alternative time windows.

[9] Table A5 describes all variables in detail.

[10] Specifically, our estimating equation writes: $Y_{i,t} = \alpha_t + \sum_{j=2}^{4} \gamma_j \, MigGroup_{i,j} + \sum_{k=-L}^{K} \sum_{j=2}^{4} \beta_{j,k} \big(Salience_{t+k} \times MigGroup_{i,j}\big) + \epsilon_{i,t}.$, where we set L as 28 days before the shipping trip and K 28 days after the shopping trip, using 7-day increments.

3.2 Media attention and compliance

Table 1 presents our main results. Column 1 reports the results for our baseline specification which includes day and group fixed effects. Column 2 adds controls for 36 shop fixed effects and time-of-day fixed effects, accounting for systematic differences in under-reporting behavior across stores and potential temporal variations throughout the day. In column 3, we include detailed customer characteristics, such as country or province of birth, gender, occupation, postcode of residence, and the year the customer joined Coop. This specification aims to capture heterogeneity in behavior linked to demographic and socioeconomic factors.

Column 4 introduces customer fixed effects, controlling for any time-invariant, individual-specific attributes that may influence under-reporting. This also ensures that the observed effects are driven by changes in behavior rather than the selection of customers shopping during periods of high salience. Column 5 further refines the analysis by incorporating worker shift fixed effects, controlling for potential influences from specific guards or cashiers on duty during audits.[11] Column 6 includes all aforementioned controls, providing the most comprehensive specification. Finally, column 7 extends column 3 by adding interaction terms between gender and occupation fixed effects with the salience index, ruling out that our results are capturing gendered or occupation-based responses to salience shocks.

Across all specifications, the interaction term between refugee nationality and media salience is negative and highly significant, indicating a consistent decrease in under-reporting for this group during periods of high salience. Notably, the magnitude of the effect remains stable even in the most exhaustive specifications.[12] Our estimates suggest that a one standard deviation increase in the number of migration-related newspaper articles decreases under-reporting between 1.1 and 1.3 percentage points, or about 10% relative to the mean. In contrast, other immigrant groups and natives from Southern Italy do not exhibit a similar pattern. As expected, their coefficients are small in magnitude and not significant.

In addition, Figure A2 shows that the effect is primarily driven by the top quartile of news salience. This suggests that media attention must surpass a certain perception threshold to trigger behavioral responses among minority groups either because heightened salience directly reaches minority individuals, or because minorities anticipate that the majority group is sufficiently aware of this coverage. As we have shown in Figure 1, the effect of media salience is rather short-lived, potentially because of information or salience decay, whereby news cycles move on quickly (Figure A1) such that majority and minority awareness of group-level scrutiny fall back within days.

[11] We define a worker shift as the morning (before 1PM) or the afternoon (after 1PM) in a particular day-shop. The assumption is that the set of guards and cashiers is fixed during a shift.

[12] Observations vary slightly between columns due to missing customer characteristic data and the inclusion of fixed effects.

3.3 Drivers of compliance

News Topic and Sentiment. Table 2 examines how the sentiment and content of news coverage affects under-reporting. Column 1 replicates our baseline estimates, using all migration-related articles classified as primarily focusing on migration by the LLM. To make estimates comparable across all columns, we always normalize salience based on the mean and standard deviation of this baseline measure of salience. Column 2 shows the effect for articles mentioning migration only tangentially. The effect is substantially smaller and significantly different from our baseline estimates (0.34 percentage points, about a third of the baseline effect), suggesting that refugee-origin shoppers primarily respond to focused migration coverage rather than more generalized news cycles.

Columns 3 and 4 split migration-related articles into negative and non-negative sentiment, using our LLM sentiment classification procedure described in Appendix B. Column 3 shows that negatively toned articles, result in a large and significant decrease in under-reporting (4.39 percentage points) among refugee-origin shoppers. Column 4 finds a significantly smaller and precisely estimated effect (1.35 percentage points) for articles without an explicit negative framing. This difference suggests greater sensitivity among refugee-origin customers to negatively framed migration news, consistent with stereotype disconfirmation behavior.

Columns 5 and 6 distinguish between articles that explicitly link migration to crime or illegality and those that do not. The results reveal that refugee-origin shoppers reduce under-reporting by nearly 6 percentage points in response to crime-framed coverage - a fivefold increase relative to the baseline. This sharp divergence supports the stereotype-disconfirmation mechanism: when public narratives associate one's group with deviance, individuals may over-compensate through heightened rule compliance to counteract perceived suspicion. In contrast, migration coverage lacking criminal associations elicits only modest behavioral changes, underscoring that it is not generic salience but targeted moral framing that prompts this effect.

Other Response Margins. Figure 2 examines other changes in shopping behavior following increased out-group salience. It is possible that refugees alter their shopping behavior, such as choosing less busy times or visiting shops farther from their residence to avoid potential interactions or judgment. These behavioral shifts could reflect increased sensitivity to social exclusion or efforts to mitigate perceived risks associated with salience shocks. In addition, we examine over-reporting, which serves as a test for inattention, as it reflects unintentional errors rather than deliberate behavior.

We observe no effects on over-reporting and broader shopping behaviors, like distance traveled to the shop, shop size, shopping during busy hours, and shopping in the morning, are small and statistically insignificant across groups. The only large and statistically significant effects in Figure 2 panel (a) confirm our main results. Individuals from refugee nations are not only less likely to underreport but also reduce the number and value of under-reported items following a salience shock. These results support

the interpretation that the reduction in under-reporting is driven by increased rule conformity rather than broader shifts in shopping habits or inattentiveness. We also explore the changes in the behavior of Southern Italians with respect to locals. This group exhibits the same baseline probability of under-reporting as refugees, have also migrated but are not subjected to the same degree of stigmatizing media attention. As with under-reporting, Southern Italians do not change any of their shopping behaviors.

Shopper Composition. Figure 2 panel (b) explores whether changes in the composition of shoppers following salience shocks might drive the results. For example, refugee-origin customers might avoid shopping altogether due to fear of social scrutiny or discrimination. Alternatively, salience shocks might attract shoppers more inclined to conform to social norms, altering the average likelihood of rule violations. Although our results hold when including customer fixed effects in Table 1 column 4, we examine this issue more explicitly by analyzing whether the characteristics of customers shopping during salience shocks differ significantly from those shopping at other times. We present estimates of a one standard deviation increase in salience on various shopper characteristics for the entire sample (red) and the sample of customers of refugee nationality (green). These include previous under-reporting and shopping habits, socioeconomic factors, demographic traits, and their place of birth. Most coefficients are small in magnitude and statistically insignificant, indicating no systematic changes in shopper composition following salience shocks. Importantly, these results also confirm that the audits themselves do not disproportionately sample certain groups over others during salience shocks, reinforcing the validity of the random audit design.

Group-Specific Salience. To further investigate the mechanisms behind the observed compliance response, we examine how the effect varies across individual characteristics that may affect the salience of group identity. In Table 3 column 1, we find that the effect is significantly stronger among male refugee-origin shoppers, in line with public narratives that associate this group with criminality. This gendered response is consistent with a stereotype-disconfirmation mechanism, where individuals who are more likely to be the target of negative stereotypes respond to heightened scrutiny by increasing rule compliance. Second, in column 2 of Table 3 we find that the compliance response is concentrated among newer members of the Coop supermarket system.[13] In contrast, those with longer tenure, who are more likely to have acquired Italian nationality or are more fully integrated, do not respond to salience shocks. This pattern supports an identity-based explanation: individuals with a weaker claim to in-group status are more sensitive to reputational risks and therefore more likely to conform under public scrutiny.

[13]We define new members as those with fewer than 15 years of membership, corresponding to the bottom three quartiles of tenure among refugee-origin individuals.

Perceived Audit Probability. A natural alternative explanation is that refugee-origin shoppers reduce under-reporting during high-salience periods because they anticipate increased monitoring. As shown previously in Figure 2, audits remain randomly assigned even during these periods. Nevertheless, refugee-origin shoppers might still believe there will be an increase in audit probability and adjust their rule compliance accordingly.

We investigate this possibility in two ways. First, we exploit quasi-random variation in shoppers' past audit experiences. The idea is that shoppers who previously experienced audits during high-salience periods might (wrongly) associate media salience with higher audit likelihood. Thus, we test whether salience effects vary with the average salience level during a shopper's past audits. Column 3 of Table 3 shows no significant heterogeneity along this dimension. Second, we leverage variation in the composition of recently audited shoppers at the same store around the time of an individual's shopping trip. The idea is that a refugee-origin shopper may infer that the auditing probability is higher if they observe that many recently audited shoppers are also refugee-origin individuals. To study this, in Column 4 of Table 3, we interact salience with the share of refugee-origin audits in the preceding 60 minutes and find no statistically significant heterogeneity. Both results support the interpretation that our results are not driven by changes in perceived audit risk.

3.4 Robustness

To assess whether news coverage accurately captures the salience of migration, we replicate our baseline analysis using the Google Search Index described in Section 2.2. Table A6 shows consistent results. We also show in Figure A3 that the dynamic effects of Google searches display a similar pattern as our baseline result. To ensure that this Google-based measure of public attention does not merely reflect periods of increased general online activity, we perform placebo tests using unrelated search terms (travel, shopping, movies, restaurants, and football) in Table A7 and only find small, statistically insignificant effects.

Next, we address the possibility of unobserved confounders correlated with salience that could influence differences in under-reporting. For example, if refugee-related issues receive more attention during periods of political uncertainty, the observed behavioral changes might reflect differences in responses to the overall political environment. However, excluding the four weeks leading up to and following national elections does not alter our results (right panel of Figure A2).

In a series of additional robustness checks, we show that our results do not change when we (i) use alternative clustering of standard errors in Table A8; (ii) use articles from different newspapers in Table A9 (iii) extend or shorten the treatment period between one and twenty-one days in Table A10; (iv) define refugee-origin shoppers based on nationalities from data on refugee boat arrivals, as well as the top 20 and top 40 asylum seeker, or extend to refugees from European countries in Table A11; (v) exclude shoppers that have under-reported items on their previous audited shopping

trip in Table A12, alleviating concerns about selective auditing of known, recent under-reporters.

4 Discussion and Conclusion

This paper examines how heightened public attention to migration shapes the behavior of minority groups, focusing on rule compliance among refugee-origin individuals. Increased migration-related coverage, especially when negative or crime-related, increases compliance among refugee-origin shoppers but not other groups.

Our results point to a mechanism where media-driven salience generates identity-specific compliance incentives, not through direct enforcement but via social visibility and reputational concerns for the in-group. While this behavioral adjustment may be interpreted as a positive form of integration, it also raises concerns about asymmetric psychological burdens placed on minority groups.

These findings suggest that the effects of media narratives extend beyond shaping majority attitudes and can directly influence the behavior of minority groups. As policymakers and media actors consider the tone and framing of migration debates, they should weigh not only the attitudinal impacts on the majority but also the behavioral responses—and potential stressors—imposed on minority populations. Future research could explore whether such responses are transient or persistent, and whether they extend beyond low-stakes settings into domains like civic engagement or institutional trust.

References

Abramitzky, R., Boustan, L., Jácome, E., Pérez, S., & Torres, J. D. (2024). Law-abiding immigrants: The incarceration gap between immigrants and the us-born, 1870–2020. *American Economic Review: Insights*, 6(4), 453–71.

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., & Zhuravskaya, E. (2015). Radio and the rise of the nazis in prewar germany. *The Quarterly Journal of Economics*, 130(4), 1885–1939.

Ajzenman, N., Dominguez, P., & Undurraga, R. (2023). Immigration, crime, and crime (mis) perceptions. *American Economic Journal: Applied Economics*, 15(4), 142–176.

Badgett, M. L., Carpenter, C. S., Lee, M. J., & Sansone, D. (2024). A review of the economics of sexual orientation and gender identity. *Journal of Economic Literature*, 62(3), 948–994.

Becker, H. S. (1963). *Outsiders: Studies in the Sociology of Deviance*. Free Press.

Belmi, P., Barragan, R. C., Neale, M. A., & Cohen, G. L. (2015). Threats to social identity can trigger social deviance. *Personality and social psychology bulletin*, 41(4), 467–484.

Benjamin, D. J., Choi, J. J., & Strickland, A. J. (2010). Social identity and preferences. *American Economic Review*, 100(4), 1913–1928.

Bertrand, M. & Duflo, E. (2017). Field experiments on discrimination. *Handbook of economic field experiments*, 1, 309–393.

Bisin, A. & Tura, G. (2019). Marriage, Fertility, and Cultural Integration in Italy. NBER Working Paper 26303.

Brader, T., Valentino, N. A., & Suhay, E. (2008). What triggers public opposition to immigration? anxiety, group cues, and immigration threat. *American Journal of Political Science*, 52(4), 959–978.

Bursztyn, L., Egorov, G., & Fiorin, S. (2017). *From extreme to mainstream: How social norms unravel*. Technical report, National Bureau of Economic Research.

Bursztyn, L., Egorov, G., & Fiorin, S. (2020). From extreme to mainstream: The erosion of social norms. *American economic review*, 110(11), 3522–3548.

Camilli, A. (2019). How italy's interior minister salvini is weaponizing migration. *International Migration*, 57(4), 165–180.

Cantoni, D., Hagemeister, F., & Westcott, M. (2019). Persistence and Activation of Right-wing Political Ideology. MWB Discussion Paper No.143.

Charness, G. & Chen, Y. (2020). Social identity, group behavior, and teams. *Annual Review of Economics*, 12(1), 691–713.

Citrin, J., Green, D. P., Muste, C., & Wong, C. (1997). Public opinion toward immigration reform: The role of economic motivations. *The Journal of Politics*, 59(3), 858–881.

Couttenier, M., Hatte, S., Thoenig, M., & Vlachos, S. (2024). Anti-muslim voting and media coverage of immigrant crimes. *The Review of Economics and Statistics*, 106(2), 576–585.

Dahl, G. B., Felfe, C., Frijters, P., & Rainer, H. (2022). Caught Between Cultures: Unintended Consequences of Improving Opportunity for Immigrant Girls. *The Review of Economic Studies*, 89(5), 2491–2528.

Djourelova, M. (2023). Persuasion through slanted language: Evidence from the media coverage of immigration. *American economic review*, 113(3), 800–835.

Dustmann, C. & Preston, I. (2007). Racial and economic factors in attitudes to immigration. *The B.E. Journal of Economic Analysis Policy*, 7(1), Article 62.

European Commission (2018). *Standard Eurobarometer 90: Public Opinion in the European Union*. Technical report, European Commission. Accessed: 2025-04-01.

Fages, D. M. & Martínez, A. A. (2023). Hate speech and social media: evidence from bolsonaro's election in brazil.

Fouka, V. (2019). How do Immigrants Respond to Discrimination? The Case of Germans in the US During World War I. *American Political Science Review*, 113(2), 405–422.

Fouka, V. (2020). Backlash: The Unintended Effects of Language Prohibition in US Schools After World War I. *The Review of Economic Studies*, 87(1), 204–239.

Fryer, R. G., Levitt, S. D., & List, J. A. (2008). Exploring the impact of financial incentives on stereotype threat: Evidence from a pilot study. *American Economic Review*, 98(2), 370–375.

Geddes, A., Hadj-Abdou, L., & Brumat, L. (2020). *Migration and Mobility in the European Union*. Macmillan International Higher Education.

Glover, D. (2019). Job Search and Intermediation Under Discrimination: Evidence from Terrorist Attacks in France. Chaire Securisation des Parcours Professionels, Working Paper 2019-02.

Grosjean, P., Masera, F., & Yousaf, H. (2023). Inflammatory political campaigns and racial bias in policing. *The Quarterly Journal of Economics*, 138(1), 413–463.

Gulino, G. & Masera, F. (2023). Contagious dishonesty: Corruption scandals and supermarket theft. *American Economic Journal: Applied Economics*, 15(4), 218–251.

Guriev, S. & Papaioannou, E. (2022). The political economy of populism. *Journal of Economic literature*, 60(3), 753–832.

Hoff, K. & Pandey, P. (2006). Discrimination, social identity, and durable inequalities. *American Economic Review*, 96(2), 206–211.

Hopkins, D. J. (2010). Politicized places: Explaining where and when immigrants provoke local opposition. *American Political Science Review*, 104(1), 40–60.

Ipsos MORI (2018). *Perceptions are not reality: What the world gets wrong*. Technical report, Ipsos MORI. Accessed: 2025-04-01.

Jaschke, P., Sardoschau, S., & Tabellini, M. (2022). *Scared straight? Threat and assimilation of refugees in Germany*. Technical report, National Bureau of Economic Research.

Keita, S., Renault, T., & Valette, J. (2024). The usual suspects: Offender origin, media reporting and natives' attitudes towards immigration. *The Economic Journal*, 134(657), 322–362.

Kray, L. J., Thompson, L., & Galinsky, A. D. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology*, 80(6), 942–958.

Lemert, E. M. (1951). *Social Pathology: A Systematic Approach to the Theory of Sociopathic Behavior*. McGraw-Hill.

Mastrobuoni, G. & Pinotti, P. (2015). Legal status and the criminal activity of immigrants. *American Economic Journal: Applied Economics*, 7(2), 175–206.

Müller, K. & Schwarz, C. (2021). Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*, 19(4), 2131–2167.

Müller, K. & Schwarz, C. (2023). From hashtag to hate crime: Twitter and antiminority sentiment. *American Economic Journal: Applied Economics*, 15(3), 270–312.

Oh, S. (2023). Does identity affect labor supply? *American Economic Review*, 113(8), 2055–2083.

Pinotti, P. (2017). Clicking on heaven's door: The effect of immigrant legalization on crime. *American Economic Review*, 107(1), 138–168.

Rudman, L. A. & Fairchild, K. (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology*, 87(2), 157–176.

Saavedra, M. (2021). Kenji or Kenneth? Pearl Harbor and Japanese-American Assimilation. *Journal of Economic Behavior & Organization*, 185, 602–624.

Sardoschau, S. & Casanueva-Artís, A. (2025). *The Cost of Tolerating Intolerance: Right-Wing Protest and Hate Crimes*. IZA Discussion Paper 17763, Institute of Labor Economics (IZA).

Schneider-Strawczynski, S. & Valette, J. (2025). Media coverage of immigration and the polarization of attitudes. *American Economic Journal: Applied Economics*, 17(1), 337–68.

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, 67(1), 415–437.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613–629.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In *Advances in experimental social psychology*, volume 34 (pp. 379–440). Elsevier.

UNHCR (2015). Press Coverage of the Refugee and Migrant Crisis in the EU: A Content Analysis of Five European Countries. Accessed: 2025-04-01.

UNHCR (2021). *Operational Data Portal: Mediterranean Situation*. Technical report, United Nations High Commissioner for Refugees. Accessed: 2025-04-01.

Yanagizawa-Drott, D. (2014). Propaganda and conflict: Evidence from the rwandan genocide. *The Quarterly Journal of Economics*, 129(4), 1947–1994.

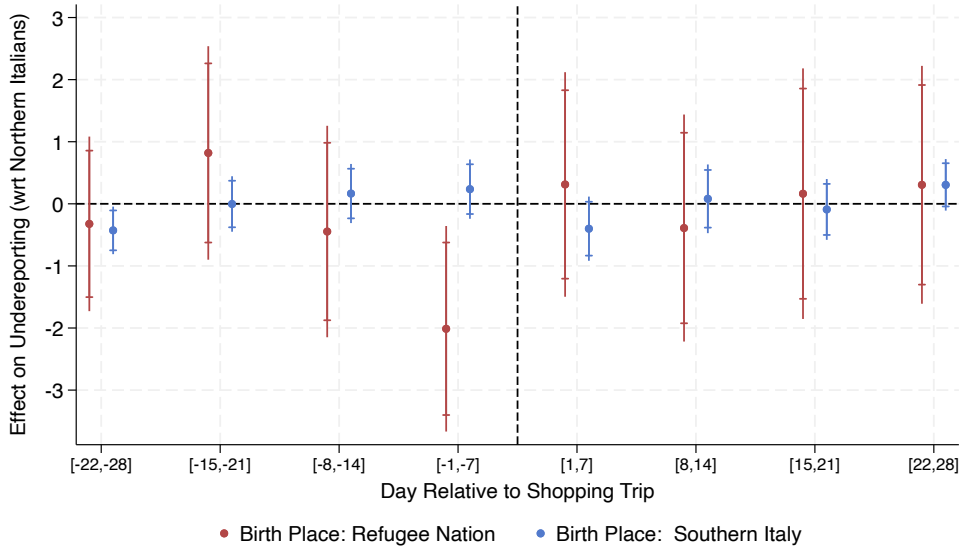
5 Main Tables and Figures

Table 1. Out-group Salience and the Likelihood of Under-reporting

	100 × 1(Under-Reporting)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Salience ×							
Refugee Nationality	-1.221***	-1.255***	-1.208***	-0.926***	-1.381***	-1.019***	-1.283***
	(0.308)	(0.308)	(0.306)	(0.352)	(0.316)	(0.379)	(0.308)
Other Foreigners	-0.086	-0.098	-0.107	0.003	-0.126	-0.042	-0.172
	(0.185)	(0.183)	(0.185)	(0.203)	(0.200)	(0.224)	(0.188)
Southern Italian	-0.019	-0.032	-0.014	0.093	-0.024	0.068	-0.062
	(0.105)	(0.104)	(0.106)	(0.107)	(0.107)	(0.107)	(0.114)
Controls	Baseline	Shopping Trip	Client Info	Client FE	Guard FE	Full Controls	Saturated
Observations	807,381	807,381	803,756	771,330	803,087	766,627	803,756

Note: The table reports the point estimates and standard deviations obtained from the estimation of equation 1. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. The unit of observation is the customer, with a given nationality, and audited on a given day. Salience measures the standardized number of migration-related newspaper articles published in *Corriere della Sera* over the previous seven days. Refugee Nationality is a dummy taking a value of 1 if the customer belongs to a refugee nationality (as defined in section 2), Southern Italian is a dummy taking a value of 1 if the customer was born in Southern Italy, Other Foreigners is a dummy taking a value of 1 if the customer is born in another non-refugee country. The omitted group is Northern Italians. In column (1), $X_{i,t}$ includes day and group fixed effects. In column (2), we also control for shop, hour, and day-of-the-week fixed effects. Column (3) instead controls for country/province of birth fixed effects, gender, occupation, postcode of residence, and year the shopper became a member of Coop. Column (4) instead adds customer fixed effects. Column (5) instead controls for guards' shifts by controlling for shop times shift fixed effects. Columns (6) include all controls mentioned above. Column (7) adds to column (3) gender fixed effects times Salience, and occupation fixed effects times Salience. Standard errors are clustered at the group times day level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Figure 1. Dynamic Treatment Effects: Under-reporting by Migrant Group



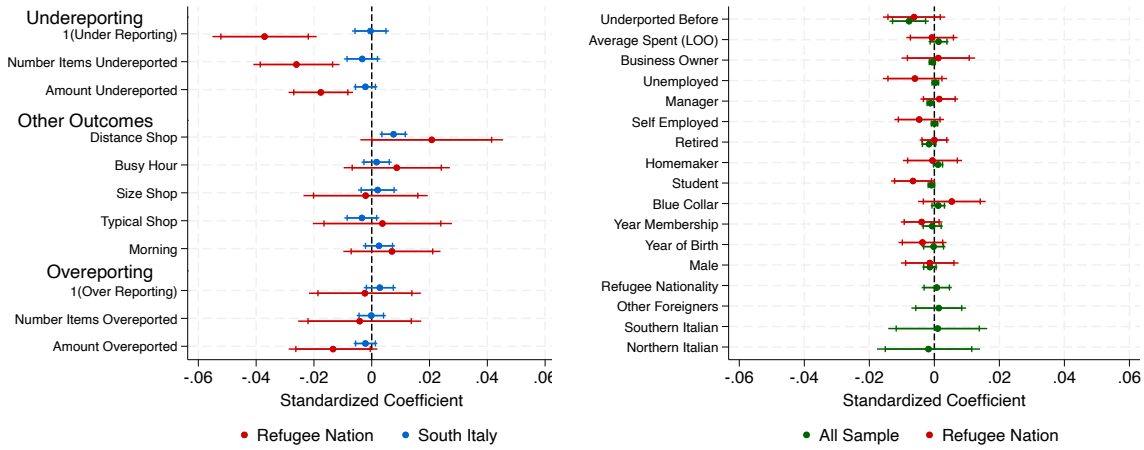
Note: The figures display the point estimate and confidence interval coming from the estimation of the following equation $y_{i,t} = \alpha_t + \sum_{g=2}^{4} \sum_{j=-4}^{4} \beta_{j,g} Salience_t^j \times 1(Group_i = g) + \gamma X_{i,t} + \epsilon_{i,t}$. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. $Salience_t^j$ measures the standardized count of migration-related newspaper articles published in *Corriere della Sera* for j weeks before (j negative) or j weeks after (j positive) the shopping trip. We display the $\beta_{j,g}$ for refugee-origin customers (as defined in section 2) and customers from Southern Italy. The omitted group is Northern Italians. $X_{i,t}$ includes country of birth and day fixed effects. Standard errors are clustered at the group times day level.

Table 2. News Type and Sentiment: Effects on Under-reporting

	100 × 1(Under-Reporting)					
	Main Topic		Tone		Crime	
	(1)	(2)	(3)	(4)	(5)	(6)
Salience ×						
Refugee Nationality	-1.221***	-0.339**	-4.690***	-1.397***	-6.341***	-1.318***
	(0.308)	(0.165)	(1.644)	(0.359)	(1.730)	(0.352)
Other Foreigners	-0.086	-0.028	0.530	-0.143	-1.072	-0.065
	(0.185)	(0.095)	(1.051)	(0.210)	(1.021)	(0.212)
Southern Italian	-0.019	0.021	-0.318	-0.010	-0.405	-0.007
	(0.105)	(0.057)	(0.563)	(0.120)	(0.536)	(0.122)
Category	Baseline	Tangential Migration	Negative	Not Negative	Crime Related	Not Crime Related
N. Articles (Mean)	3.45	6.89	0.32	3.12	0.48	2.96
P-val Diff.	0.01		0.05		0.00	
Observations	807,381	807,381	807,381	807,381	807,381	807,381
Controls	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline

Note: The table reports the point estimates and standard deviations obtained from the estimation of equation 1. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. The unit of observation is the customer, with a given nationality, and audited on a given day. Salience indicates the standardized number of migration-related articles published in Corriere della Sera over the previous seven days in column (1). In Column (2), Salience instead includes articles that only tangentially mention migration. In column (3), Salience refers to migration-related articles with negative sentiment, while column (4) captures those with non-negative sentiment. In column (5), Salience refers to migration-related articles containing crime-related stereotypes, while column (6) includes those without such stereotypes. Refugee Nationality is a dummy taking a value of 1 if the customer belongs to a refugee nationality (as defined in section 2), Southern Italian is a dummy taking a value of 1 if the customer was born in Southern Italy, Other Foreigners is a dummy taking a value of 1 if the customer is born in another non-refugee country. The omitted group is Northern Italians. $X_{i,t}$ includes, in all columns, day and group fixed effects. *P-val Diff.* reports the p-value for the test of equality between the coefficients on Refugee Nationality. Specifically, it compares columns (1) and (2), reported in column (2); columns (3) and (4), reported in column (4); and columns (5) and (6), reported in column (6). Standard errors are clustered at the group times day level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Figure 2. Alternative Outcomes and Selection



(a) Other Shopping Outcomes

(b) Shopper Characteristics

Note: Panel A shows the effect of Salience on various outcomes. All outcomes are standardized to have a mean of zero and a standard deviation of 1. The red coefficient is the effect on individuals with a refugee nationality (as defined in section 2). The blue coefficient is the effect on individuals born in Southern Italy. Both coefficients are compared to individuals born in Northern Italy. Panel B shows the correlation between individual shopper characteristics and Salience. All individual shopper characteristics are standardized to have a mean of zero and a standard deviation of 1. The coefficients using the whole sample are in green. The coefficients using only individuals born in a refugee nation are in red.

Table 3. Drivers of Compliance: Heterogeneity Analysis

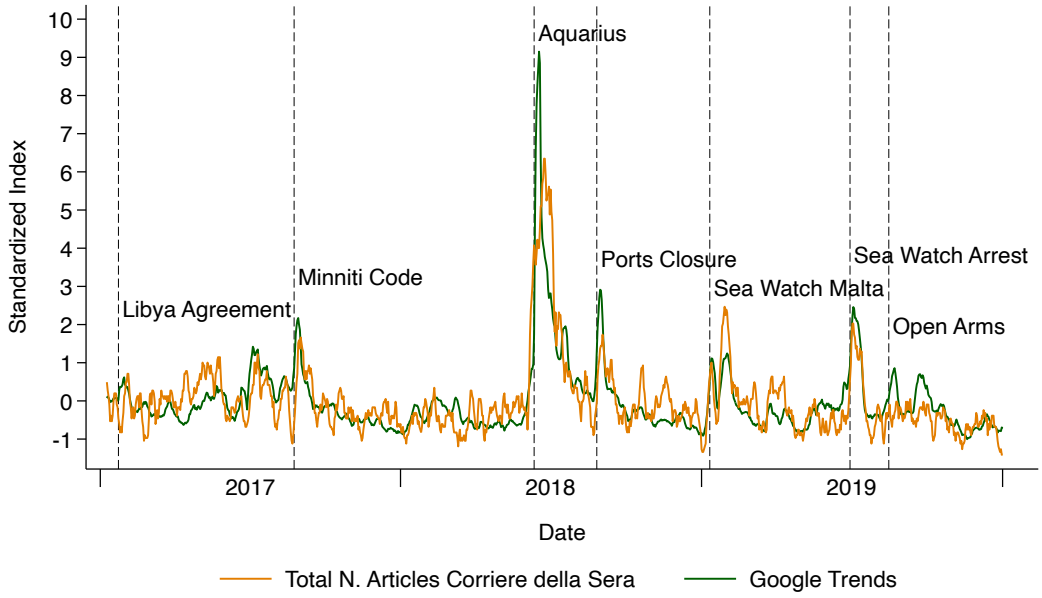
	Group Salience		Perceived Audit Probability	
	(1)	(2)	(3)	(4)
Salience ×				
Refugee Nationality	-0.385	-0.033	-1.047***	-1.131***
	(0.570)	(0.520)	(0.342)	(0.303)
Refugee Nationality × Heterogeneity	-1.375**	-1.707**	0.223	-0.058
	(0.671)	(0.694)	(0.549)	(0.069)
Heterogeneity Variable	Male	Newer Member	Own Experience	Co-Shopper Experience
Observations	807,381	802,542	659,665	695,723

Note: The table reports the point estimates and standard deviations obtained from the estimation of equation 1. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. The unit of observation is the customer, with a given nationality, and audited on a given day. Salience indicates the standardized number of migration-related articles published in Corriere della Sera over the previous seven days. Refugee Nationality is a dummy taking a value of 1 if the customer belongs to a refugee nationality (as defined in section 2). We also control for Refugee Nationality × Heterogeneity and Salience × Refugee Nationality × Heterogeneity, where the Heterogeneity variable varies across columns. In column (1), Heterogeneity is a dummy equal to 1 if the client is male, and 0 otherwise. In column (2), Heterogeneity is a dummy equal to 1 if the client has been a member of the supermarket for less than 15 years. In column (3), Heterogeneity is the average of all past values of the variable Salience recorded while the client was audited. In column (4), Heterogeneity is the share of treated clients present during the previous one-hour window. The omitted group is Northern Italians. $X_{i,t}$ includes, in all columns, also day and group fixed effects. Standard errors are clustered at the group times day level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Online Appendix

Appendix A: Appendix Tables and Figures

Figure A1. Google Trend and Factiva Articles Series

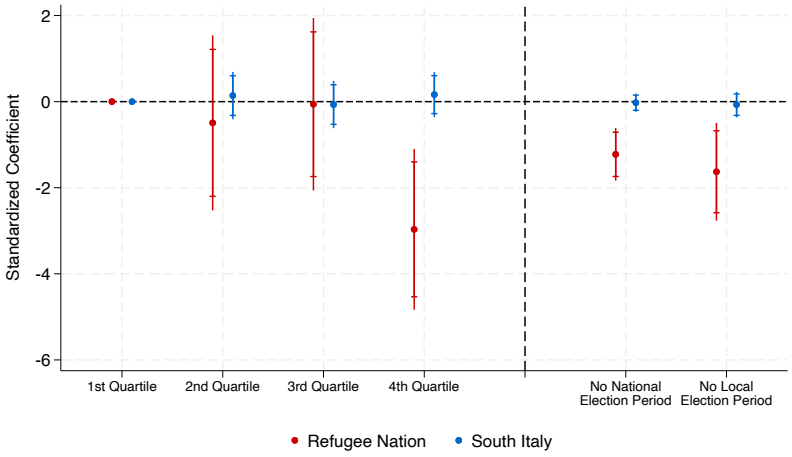


Note: The figure displays, in orange, the standardized number of newspaper articles published in *Corriere della Sera* that the LLM categorized as migration-related over the previous seven days, and, in green, the standardized search volume for migration-related keywords in Italy over the same period, as defined in Section 2.2.

Table A1. Summary Statistics Articles from *Corriere della Sera*

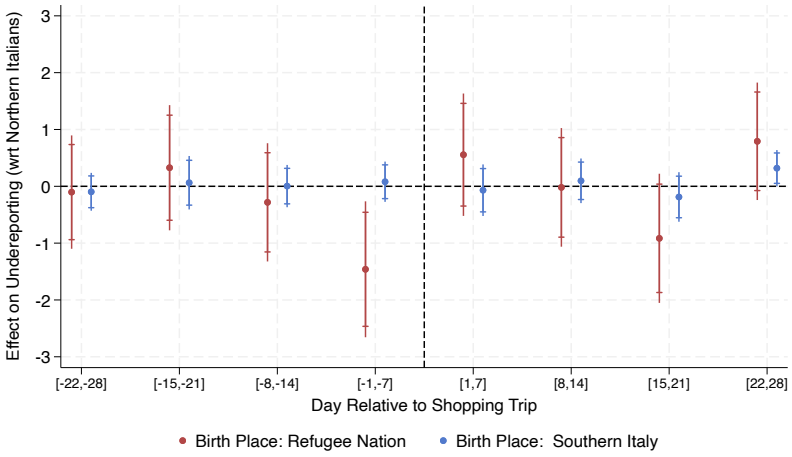
	Total	**Avg per Day**	**Sd**	**Min.**	**Max.**
Total No. of Articles (Factiva)	11311	10	5.5	2.6	44
Total No. of Articles (Identified by ChatGPT):	.	.	.	.	.
Articles About Migration	3771	3.5	2	.71	16
Articles Tangentially About Migration	7540	6.9	3.7	1.9	31
Articles with Negative Sentiment (Migration)	358	.33	.38	0	3
Articles with Non-Negative Sentiment (Migration)	3413	3.1	1.7	.71	15
Articles with Stereotypes	526	.48	.36	0	2.7
Articles without Stereotypes	3245	3	1.7	.71	14

Figure A2. Out-group Salience and the Likelihood of Under-reporting: Robustness



Note: The figure displays the point estimate and confidence interval coming from the estimation of equation 1. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. In the left panel of the graph, we created four dummy variables taking the value 1, one for each of the quartiles of the $Salience_t^{[1,7]}$ distribution. The omitted quartile is the first one. In the right panel of the graph, we replicate the baseline analysis, excluding the 30 days before and after national and local elections. We display the estimates for refugee-origin customers (as defined in section 2) and customers born in Southern Italy. The omitted group is Northern Italians. $X_{i,t}$ includes country of birth and day fixed effects. Standard errors are clustered at the group times day level.

Figure A3. Dynamic Treatment Effects: Under-reporting by Migrant Group: Google Trends Index



Note: The figure displays the point estimate and confidence interval coming from the estimation of the following equation $y_{i,t} = \alpha_t + \sum_{g=2}^{4} \sum_{j=-5}^{5} \beta_{j,g} Salience_t^j \times 1(Group_i = g) + \gamma X_{i,t} + \epsilon_{i,t}$. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. $Salience_t^j$ measures the standardized Google search volume for migration-related keywords in Italy for j weeks before (j negative) or j weeks after (j positive) the shopping trip. We display the $\beta_{j,g}$ for refugee-origin customers (as defined in section 2) and customers born in Southern Italy. The omitted group is Non-Souther Italians. $X_{i,t}$ includes country of birth and day fixed effects.Standard errors are clustered at the group times day level.

Table A2. Summary Statistics

	N	Mean	Std. Dev.	Min.	Max.
Total Obs	813518				
Clients	148181				
Shops	36				
Municipalities	1871				
Customer Characteristics:					
Age	.	51	15	18	119
Male	.	42	49	0	100
Years Since Membership	.	16	11	0	74
White collar employee	.	32	46	0	100
Blue collar employee	.	20	40	0	100
Retired	.	10	30	0	100
Homemaker	.	6.5	25	0	100
Self employed	.	4.3	20	0	100
Business owner	.	.53	7.3	0	100
Teacher	.	3.5	18	0	100
Student	.	4.9	22	0	100
Unemployed	.	1.3	11	0	100
Other employment	.	10	30	0	100
Born in Italy	.	95	23	0	100
Refugee Nationality	.	1.4	12	0	100
Other Foreingers	.	4	20	0	100
Southern Italian	.	15	35	0	100
Province of Birth	116				
Audits Record:					
Under-Reporting	.	13	34	0	100
Over-reporting	.	5.4	23	0	100
Total Value	.	53	48	0	1055
Share of Value of Under-Reported	.	7.9	16	.022	99
Share of Value of Over-Reported	.	11	60	.027	1979

Table A3. Summary statistics (Whole Population)

	Mean Population	Mean Our Sample
Customers with Residence in Modena:		
Age [1]	43.4	53.8
Male [1]	48.9	40.6
Retired Men [2]	21.6	16.5
Retired Women [2]	22.3	11.9
All types of self-employed [3] *	25.8	4.5
All types of employee [3] **	74.2	55.3
Italian nationality [4]	90.3	94.2
Customers with Residence in Ferrara:		
Age [1]	47.4	56.3
Male [1]	48.1	43.0
Retired Men [2]	26.4	19.6
Retired Women [2]	26.1	10.2
All types of self-employed [3] *	21.6	3.7
All types of employee [3] **	78.4	55.5
Italian nationality [3]	93.6	97.3
Customers with Residence Elsewhere:		
Age [1]		47.5
Male [1]		47.9
Retired Men [2]		8.3
Retired Women [2]		5.5
All types of self-employed [3] *		3.2
All types of employee [3] **		51.0
Italian nationality [4]		95.3

Source (1): year 2019
Source (2): year 2019
Source (3): year 2016
Source (4): year 2016
$*$ Categories included: Self employed and Business owner
$**$ Categories included: White collar employee, Blue collar employee, Teacher

Table A4. Summary Statistics - Only Refugees Nationality

	N	Mean	Std. Dev.	Min.	Max.
Total Obs	9468				
Clients	2131				
Shops	35				
Municipalities	156				
Customer Characteristics:					
Age	.	44	12	18	86
Male	.	65	48	0	100
Years Since Membership	.	10	8.1	0	49
White collar employee	.	9.8	30	0	100
Blue collar employee	.	48	50	0	100
Retired	.	2	14	0	100
Homemaker	.	7.5	26	0	100
Self employed	.	4.7	21	0	100
Business owner	.	.66	8.1	0	100
Teacher	.	.99	9.9	0	100
Student	.	4.5	21	0	100
Unemployed	.	2.1	14	0	100
Other employment	.	14	35	0	100
Born in Italy	.	0	0	0	0
Refugee Nationality	.	100	0	100	100
Other Foreingers	.	0	0	0	0
Southern Italian	.	0	0	0	0
Audits Record:					
Under-Reporting	.	14	35	0	100
Over-reporting	.	5	22	0	100
Total Value	.	42	46	0	890
Share of Value of Under-Reported	.	9.1	13	.069	98
Share of Value of Over-Reported	.	18	70	.087	900

Table A5. Variables description and data sources: main variables

Main treatment:

N. of Articles indicates the standardized number of migration-related articles published in *Corriere della Sera* over the previous seven days.

Google Trends Index. measures the standardized search volume for migration-related keywords in Italy over the preceding seven days.

Main Nationality Group:

Center-Northern Italians. It is a dummy taking the value 1 if the customer was born in the Center-North of Italy (List of Region: Abruzzo, Emilia-Romagna, Friuli-Venezia Giulia, Lazio, Liguria, Lombardia, Marche, Piemonte, Sardegna, Toscana, Trentino-Alto Adige Umbria, Valle d'Aosta, and Veneto).

Center-Southern Italians. It is a dummy taking the value 1 if the customer was born in the Center-South of Italy (List of Region: Basilicata, Calabria, Campania, Molise, Puglia, and Sicilia).

Refugee Nationality. It is a dummy taking the value 1 if the customer was born in one of the following foreing countries: Afghanistan, Bangladesh, Burkina Faso, Camerun, Colombia, Cote d'Ivorie, Egypt, El Salvador, Eritrea, Ghana, Guinea, India, Iraq, Libia, Mali, Morocco, Nigeria, Pakistan, Peru, China, Senegal, Sierra Leone, Syria, Somalia, Sri Lanka, Sudan, Togo, Tunisia, Turkey, and Venezuela. These are the top 30 non-European countries from the following list of countries compiled by Eurostat.

Other Foreigners. It is a dummy taking the value 1 if the customer was born in a foreign country that is not in the list Refugee Nationality above.

Main outcomes:

Under-reporting. It is a dummy taking value 100 if customer is found to under-report at least a product and 0 otherwise.

Over-reporting. It is a dummy taking value 100 if customer is found to over-report at least a product and 0 otherwise.

Customer characteristics:

Age. The age of the customer.

Male. It is a dummy taking value 1 if customer is male and 0 otherwise.

Years since Membership It is the number of years since the customer has been a member of the supermarket chain.

White collar employee. It is a dummy taking value 1 if customer is employee and 0 otherwise.

Blue collar employee. It is a dummy taking value 1 if customer is a process worker and 0 otherwise.

Retired. It is a dummy taking value 1 if customer is retired and 0 otherwise.

Housewife. It is a dummy taking value 1 if customer is a housewife and 0 otherwise.

Self-employed. It is a dummy taking value 1 if customer is self-employed and 0 otherwise.

Business owner. It is a dummy taking value 1 if customer is a business owner and 0 otherwise.

Teacher. It is a dummy taking value 1 if customer is a teacher and 0 otherwise.

Student. It is a dummy taking value 1 if customer is a student and 0 otherwise.

Unemployed. It is a dummy taking value 1 if customer is unemployed and 0 otherwise.

Other employment. It is a dummy taking value 1 if customer employment is not classified under any of the previous categories and 0 otherwise.

Italian nationality. It is a dummy taking value 1 if customer has Italian nationality and 0 otherwise.

Province of birth. The province of birth of the customer.

Table A6. Out-group Salience and the Likelihood of Under-reporting: Google Trends Index

	100 × Under-Reporting						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Google Search Index ×							
Refugee Nationality	-1.091***	-1.128***	-1.076***	-1.021***	-1.275***	-1.171***	-1.187***
	(0.345)	(0.345)	(0.344)	(0.367)	(0.355)	(0.392)	(0.347)
Other Foreigners	0.003	-0.017	-0.028	0.214	-0.040	0.190	-0.111
	(0.168)	(0.166)	(0.166)	(0.177)	(0.201)	(0.207)	(0.168)
Southern Italian	0.070	0.057	0.076	0.128	0.061	0.115	-0.017
	(0.096)	(0.095)	(0.096)	(0.094)	(0.100)	(0.096)	(0.106)
Controls	Baseline	Shopping Trip	Client Info	Client FE	Guard FE	Full Controls	Saturated
Observations	807,939	807,939	804,310	771,878	803,639	767,170	804,310

Note: The table reports the point estimates and standard deviations obtained from the estimation of equation 1. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. The unit of observation is the customer, with a given nationality, and audited on a given day. Google Trends Index measures the standardized search volume for migration-related keywords in Italy over the preceding seven days. Refugee Nationality is a dummy taking a value of 1 if the customer belongs to a refugee nationality (as defined in section 2), Southern Italian is a dummy taking a value of 1 if the customer was born in Southern Italy, Other Foreigners is a dummy taking a value of 1 if the customer was born in another non-refugee country. The omitted group is Northern Italians. In column (1), $X_{i,t}$ includes day and group fixed effects. In column (2), we also control for shop, hour, and day-of-the-week fixed effects. Column (3) instead controls for country/province of birth fixed effects, gender, occupation, postcode of residence, and year the shopper became a member of Coop. Column (4) instead adds customer fixed effects. Column (5) instead controls for guards' shifts by controlling for shop times shift fixed effects. Columns (6) include all controls mentioned above. Column (7) adds to column (3) controls gender fixed effects times Google Trends Index, and occupation fixed effects times Google Trends Index. *P-val Diff.* reports the p-value for the test of equality between the coefficients on Refugee Nationality and Southern Italian. Standard errors are clustered at the group times day level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A7. Out-group Salience and the Likelihood of Under-reporting: Placebo Google Index

	100 × Undereporting					
	(1)	(2)	(3)	(4)	(5)	(6)
Google Search Index ×						
Refugee Nation	-0.301	-0.167	0.080	0.465	-0.176	0.051
	(0.391)	(0.361)	(0.328)	(0.383)	(0.353)	(0.382)
Southern Italy	0.121	-0.041	0.344***	0.297**	-0.032	0.286***
	(0.105)	(0.097)	(0.103)	(0.115)	(0.097)	(0.106)
Other Foreigners	-0.192	0.121	0.121	0.328	-0.009	0.130
	(0.196)	(0.197)	(0.193)	(0.204)	(0.196)	(0.191)
Observations	807939	807939	807939	807939	807939	807939
Search Index	Travel	Shopping	Movies	Restaurant	Football	Index

Note: The table reports the point estimates and standard deviations obtained from the estimation of equation 1. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. The unit of observation is the customer, with a given nationality, and audited on a given day. Google Trends Index measures the standardized search volume for different keywords in Italy over the preceding seven days: in column (1), keywords relate travel, in column (2) keywords related to shopping, in column (3) keywords related to movies, in column (4) keywords related to restaurant, in column (5) keywords related to movies, and in column (6) is the principal component of the previous variables. Refugee Nationality is a dummy taking a value of 1 if the customer belongs to a refugee nationality (as defined in section 2), Southern Italian is a dummy taking a value of 1 if the customer was born in Southern Italy, Other Foreigners is a dummy taking a value of 1 if the customer belongs to an other foreign nationalities. The omitted group is Northern Italians. In column (1), $X_{i,t}$ includes day and group fixed effects. In column (2), we also control for shop, hour, and day-of-the-week fixed effects. Column (3) instead controls for country/province of birth fixed effects, gender, occupation, postcode of residence, and year the shopper became a member of Coop. Column (4) instead adds customer fixed effects. Column (5) instead controls for guards' shifts by controlling for shop times shift fixed effects. Columns (6) include all controls mentioned above. Column (7) adds to column (3) controls gender fixed effects times Google Trends Index, and occupation fixed effects times Google Trends Index. Standard errors are clustered at the group times day level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A8. Out-group Salience and the Likelihood of Under-reporting: Alternative Cluster

	100 × Under-Reporting						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Clusters:							
Country of Birth	-1.221***	-1.255***	-1.208***	-0.926***	-1.381***	-1.019***	-1.283***
	(0.306)	(0.301)	(0.307)	(0.322)	(0.301)	(0.345)	(0.317)
Country of Birth × Day Level	-1.221***	-1.255***	-1.208***	-0.926**	-1.381***	-1.019***	-1.283***
	(0.310)	(0.308)	(0.311)	(0.360)	(0.323)	(0.381)	(0.315)
Shop Level	-1.221***	-1.255***	-1.208***	-0.926**	-1.381***	-1.019**	-1.283***
	(0.268)	(0.271)	(0.275)	(0.359)	(0.309)	(0.391)	(0.269)
Shop × Day Level	-1.221***	-1.255***	-1.208***	-0.926***	-1.381***	-1.019***	-1.283***
	(0.311)	(0.309)	(0.312)	(0.355)	(0.330)	(0.373)	(0.314)
Municipality Level	-1.221***	-1.255***	-1.208***	-0.926***	-1.381***	-1.019***	-1.283***
	(0.259)	(0.261)	(0.267)	(0.295)	(0.278)	(0.287)	(0.267)
Municipality × Day Level	-1.221***	-1.255***	-1.208***	-0.926***	-1.381***	-1.019***	-1.283***
	(0.307)	(0.306)	(0.308)	(0.349)	(0.336)	(0.386)	(0.312)
Observations	807,381	807,381	803,756	771,330	803,087	766,627	803,756
Controls	Baseline	Shopping Trip	Client Info	Client FE	Guard FE	Full Controls	Saturated

Note: The table reports the point estimates and standard deviations obtained from the estimation of equation 1. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. In column (1), $X_{i,t}$ includes day and group fixed effects. In column (2), we also control for shop, hour, and day-of-the-week fixed effects. Column (3) instead controls for country/province of birth fixed effects, gender, occupation, postcode of residence, and year the shopper became a member of Coop. Column (4) instead adds customer fixed effects. Column (5) instead controls for guards' shifts by controlling for shop times shift fixed effects. Columns (6) include all controls mentioned above. Column (7) adds to column (3) controls gender fixed effects times Salience, and occupation fixed effects times Salience. Each row shows the effect on refugee nationality shoppers based on different standard errors clustering. These, in order, are country of birth, country of birth by day, shop, shop by day, municipality, and municipality by day level in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A9. Out-group Salience and the Likelihood of Under-reporting: Alternative Journals

	100 × Under-Reporting						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Salience per Journal:							
Total	-1.344***	-1.381***	-1.318***	-1.049***	-1.461***	-1.112***	-1.397***
	(0.312)	(0.311)	(0.309)	(0.363)	(0.323)	(0.393)	(0.311)
La Repubblica	-1.162***	-1.195***	-1.129***	-1.018***	-1.184***	-0.996**	-1.201***
	(0.322)	(0.322)	(0.321)	(0.378)	(0.343)	(0.409)	(0.321)
Gazzetta di Modena	-1.144***	-1.178***	-1.102***	-0.641*	-1.274***	-0.719*	-1.162***
	(0.332)	(0.331)	(0.328)	(0.364)	(0.346)	(0.392)	(0.329)
La Nuova Ferrara	-1.352***	-1.381***	-1.337***	-1.200***	-1.453***	-1.292***	-1.390***
	(0.336)	(0.335)	(0.334)	(0.381)	(0.352)	(0.405)	(0.335)
Observations	807,381	807,381	803,756	771,330	803,087	766,627	803,756
Controls	Baseline	Shopping Trip	Client Info	Client FE	Guard FE	Full Controls	Saturated

Note: The table reports the point estimates and standard deviations obtained from the estimation of equation 1. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. Salience measures the standardized number of migration-related newspaper articles published in all Journals (row 1), *La Repubblica* (row 2), *La Gazzetta di Modena* (row 3), and *La Nuova Ferrara* (row 4) over the previous seven days. The reported coefficient is the one related to shoppers with refugee nationality. In column (1), $X_{i,t}$ includes day and group fixed effects. In column (2), we also control for shop, hour, and day-of-the-week fixed effects. Column (3) instead controls for country/province of birth fixed effects, gender, occupation, postcode of residence, and year the shopper became a member of Coop. Column (4) instead adds customer fixed effects. Column (5) instead controls for guards' shifts by controlling for shop times shift fixed effects. Columns (6) include all controls mentioned above. Column (7) adds to column (3) controls gender fixed effects times Salience, and occupation fixed effects times Salience. Standard errors are clustered at the group times day level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A10. Out-group Salience and the Likelihood of Under-reporting: Alternative Time Windows

	100 × Under-Reporting						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Salience:							
Window Salience [-1,-1]	-0.868***	-0.891***	-0.835**	-0.945***	-1.075***	-1.049***	-0.881***
	(0.327)	(0.327)	(0.325)	(0.345)	(0.331)	(0.362)	(0.327)
Window Salience [-1,-7]	-1.221***	-1.255***	-1.208***	-0.926***	-1.381***	-1.019***	-1.283***
	(0.308)	(0.308)	(0.306)	(0.352)	(0.316)	(0.379)	(0.308)
Window Salience [-1,-14]	-1.157***	-1.199***	-1.137***	-0.889**	-1.285***	-0.991***	-1.210***
	(0.292)	(0.292)	(0.289)	(0.353)	(0.306)	(0.382)	(0.290)
Window Salience [-1,-21]	-0.984***	-1.022***	-0.959***	-0.785**	-1.073***	-0.858**	-1.018***
	(0.287)	(0.286)	(0.281)	(0.360)	(0.310)	(0.388)	(0.282)
Observations	795,866	795,866	792,403	759,839	791,622	755,183	792,403
Controls	Baseline	Shopping Trip	Client Info	Client FE	Guard FE	Full Controls	Saturated

Note: The table reports the point estimates and standard deviations obtained from the estimation of equation 1. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. Salience measures the standardized number of migration-related newspaper articles published in Corriere della Sera. The time window over which articles are aggregated varies across rows: in row 1, Salience is computed based on articles published on the previous day only; in row 2, over the previous seven days; in row 3, over the previous fourteen days; and in row 4, over the previous twenty-one days. We report coefficients for refugee-origin customers. The omitted group is Northern Italians. In column (1), $X_{i,t}$ includes day and group fixed effects. In column (2), we also control for shop, hour, and day-of-the-week fixed effects. Column (3) instead controls for country/province of birth fixed effects, gender, occupation, postcode of residence, and year the shopper became a member of Coop. Column (4) instead adds customer fixed effects. Column (5) instead controls for guards' shifts by controlling for shop times shift fixed effects. Columns (6) include all controls mentioned above. Column (7) adds to column (3) controls gender fixed effects times Salience, and occupation fixed effects times Salience. Standard errors are clustered at the group times day level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A11. Out-group Salience and the Likelihood of Under-reporting: Alternative Definitions of Country Lists

	100 × 1(Under-Reporting)					
	(1)	(2)	(3)	(4)	(5)	(6)
Salience ×						
Refugee Nationality	-1.221***	-0.616**	-0.922***	-1.220***	-1.148***	-1.063***
	(0.308)	(0.265)	(0.299)	(0.351)	(0.272)	(0.265)
Other Foreigners	-0.086	-0.179	-0.199	-0.160	-0.033	-0.038
	(0.185)	(0.204)	(0.190)	(0.181)	(0.194)	(0.199)
Southern Italian	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019
	(0.105)	(0.105)	(0.105)	(0.105)	(0.105)	(0.105)
Category	Baseline	Baseline Plus Europe	List Arrivals	Top 20	Top 35	top 40
Observations	807,381	807,381	807,381	807,381	807,381	807,381

Note: The table reports the point estimates and standard deviations obtained from the estimation of equation 1. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. Salience indicates the standardized number of migration-related articles published in Corriere della Sera over the previous seven days. The definition of Refugee Nationality varies across columns: in column (1), we use our baseline definition, as described in Section 2; column (2) uses the same baseline definition plus European countries; column (3) uses data on the nationalities of migrants arriving by boat; column (4) restricts the baseline list to the top 20 nationalities; column (5) extends the list to the top 35; and column (6) to the top 40. Southern Italian is a dummy taking a value of 1 if the customer was born in Southern Italy, Other Foreigners is a dummy taking a value of 1 if the customer belongs to another foreign nationality. The omitted group is Northern Italians. $X_{i,t}$ includes, in all columns, day and group fixed effects. Standard errors are clustered at the group times day level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A12. Out-group Salience and the Likelihood of Under-reporting: Excluding Known Under-reporters

	100 × 1(Under-Reporting)				
	(1)	(2)	(3)	(4)	(5)
Salience ×					
Refugee Nationality	-1.045***	-1.118***	-1.054***	-1.066***	-1.048***
	(0.337)	(0.336)	(0.339)	(0.334)	(0.339)
Other Foreigners	0.097	0.101	0.097	0.166	0.245
	(0.187)	(0.184)	(0.183)	(0.185)	(0.187)
Southern Italian	0.030	0.023	0.020	0.045	0.015
	(0.105)	(0.103)	(0.104)	(0.106)	(0.104)
Kwon Stealer Above	10 Euros	7 Euros	5 Euros	3 Euros	2 Euros
Observations	658,194	653,146	646,771	632,999	619,975

Note: The table reports the point estimates and standard deviations obtained from the estimation of equation 1. The *dependent variable* is a dummy variable equal to 100 if the customer is found to under-report at least one product while shopping and zero otherwise. The unit of observation is the customer, with a given nationality, and audited on a given day. We exclude customers that have under-reported items of varying total value in their previous audit. We vary the value threshold across columns: more than 10 euros in column (1), 7 euros in column (2), 5 euros in column (3), 3 euros in column (4), and 2 euros in column (5). Salience measures the standardized number of migration-related newspaper articles published in *Corriere della Sera* over the previous seven days. Refugee Nationality is a dummy taking a value of 1 if the customer belongs to a refugee nationality (as defined in section 2), Southern Italian is a dummy taking a value of 1 if the customer was born in Southern Italy, Other Foreigners is a dummy taking a value of 1 if the customer was born in another non-refugee country. The omitted group is Northern Italians. Standard errors are clustered at the group times day level. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Appendix B: Text Classification Procedure via ChatGPT

To analyze the content of newspaper articles, we employed the ChatGPT API (GPT-4o model by OpenAI) to automate text classification. Specifically, we submitted the full text of each article along with two distinct prompts designed to extract structured information regarding both the thematic content and the language used. The first prompt focused on categorizing each article along thematic, narrative, and evaluative dimensions, while the second aimed to identify political statements and assess their sentiment. The API responses were returned in JSON format to facilitate automated processing in both cases. The two prompts used are reported below.

First Prompt:

Hello, you will be provided with a series of articles and questions about them. The articles were published in the four prominent Italian newspapers, and the general theme is immigration. Kindly act as a university professor and classify each article into one of the following categories based on its content. Specifically, we are interested in understanding the real content of the article. Provide your answer in JSON format.

A) Could you provide a concise summary of the article? Thank you.

B) Could you provide an extremely brief summary of the article in a few words? Thank you.

C) Does the article genuinely address the issue of migration or refugee issues, or does it focus on something else? Please respond ONLY with one of the following options: 'Migration' or 'Else.' Thank you.

D) What is the article's stance toward immigration (not its overall tone)? Please respond ONLY with the following fields: 'Negative', 'Positive,' and 'Neutral.' Use 'Negative' for statements criticizing or opposing immigration or immigrants, 'Positive' for statements supporting immigration or immigrants, and 'Neutral' Neutral if it is factual or does not express a clear opinion about immigration. Thank you.

E) Could you categorize the primary topic of the article into one of the following categories? Please respond ONLY with one of these options: 'Immigrants,' 'Crime,' 'Politics,' 'Economy,' 'Infrastructures,' or 'International'. Use 'Immigrants' if it primarily discusses immigrants or refugees themselves, 'Crime' if it focuses on criminal acts or law enforcement, 'Politics' if it emphasizes political events, policies, or figures that do not relate to immigrants or refugees, 'Economy' if it centers on economic issues like jobs, markets, or financial aspects, 'Infrastructures' if it

mainly discusses roads, buildings, public works, etc, 'International' (if it chiefly covers international affairs, conflicts, or foreign relations. Thank you.

F) Could you sub-classify the topic of the article that you have classified as Immigrants? Please respond ONLY with one of these options: 'Migrations,' 'Refugees,' or 'Arrivals.' Use 'Migrations' if it primarily discusses migrants from OECD or European countries, 'Refugees' if it primarily discusses refugees, 'Arrivals' if it primarily discusses refugee boat arrivals, illegal crossings, or search and rescue missions. Thank you.

G) Does the article explicitly discuss border security or illegal immigration involving maritime arrivals? Please respond ONLY with the following fields: 'YES', 'NO'. Thank you.

Second Prompt:

Hello, you will be provided with a series of articles and questions about them. The articles were published in the four prominent Italian newspapers, and the general theme is immigration. Kindly act as a university professor and classify each article into one of the following categories based on its content. Specifically, we are interested in politicians' statements, what they are talking about or commenting on, and their sentiments about it. Provide your answer in JSON format.

A) Does the article report a statement made by a politician? Please respond ONLY with the following fields: 'YES', 'NO'. Thank you.

B) If a statement is reported, what is the name of the politician? If multiple politicians are cited, provide a list of names separated by the following symbol '@'. Thank you.

C) Please summarize the previous answer in just a couple of words. Thank you.

D) If a statement is reported, what is the sentiment of the statement? Please respond ONLY with the following fields: 'Negative,' 'Positive,' and 'Neutral.' Use 'Negative' for statements criticizing or opposing immigration, 'Positive' for statements supporting immigration, and 'Neutral' for factual or non-committal statements. Thank you.

E) Can you tell if the article discusses the topic of immigration in general or specifically refers to a specific event? Please respond ONLY with the following fields: 'General', 'Specific'. Thank you.