#IamLGBT: Social Networks and Coming Out*

Jan Gromadzki†, Przemysław Siemaszko‡

October 13, 2025

Abstract

In recent decades, the number of people disclosing their LGBTQ identity has increased substantially. We use newly collected data from two waves of a spontaneous Twitter coming out campaign to study the role of peer effects in coming out. Importantly, we are able to distinguish actual public coming out decisions — costly, explicit disclosures — from mere engagement with the campaign. We combine data on users' pre-campaign networks with the information on the exact time of coming out actions to construct a time-varying measure of the exposure to peers coming out as LGBTQ. A one standard deviation increase in the exposure increases the hourly probability of coming out by almost 20 percent. We also exploit the non-overlapping network structure of users' peer groups as an exogenous source of variation, and we confirm the baseline results.

Keywords: LGBTQ; social networks; peer effects; social media; cultural change

JEL Codes: J15, D85, D74, Z13

*We are grateful to Luca Bagnato, Marianne Bitler, Paweł Bukowski, Kitt Carpenter, Scott Carrell, Shuai Chen, Stefano DellaVigna, Keith Marzilli Ericson, Joanna Franaszek, Jeff Frank, Daniel Hamermesh, Jorge García Hombrados, Philipp Kerler, Iga Magda, Nelson Sunuwar, Danisz Okulicz, Martha Olney, Tymon Słoczyński, Guo Xu, and Krzysztof Zaremba for very helpful discussions. We thank seminar and conferences participants at the UNU-MERIT, AEA/ASSA Annual Meeting, EEA-ESEM, (Ce)^2 Workshop, the Discrimination and Diversity Workshop, and the Workshop on life-changing transitions in the LGBTQ+ community for comments and suggestions. We gratefully acknowledge use of the Python/Stata template provided by von Gaudecker (2014). Siemaszko acknowledges funding from the National Science Centre through grant no. 2018/29/N/HS4/00890.

†Department of Economics, Vienna University of Economics and Business, Email: jan.gromadzki[at]wu[dot]ac[dot]at

‡SGH Warsaw School of Economics, Email: przemyslaw.siemaszko[at]gmail[dot]com

1 Introduction

The rapid increase in the number of openly LGBTQ individuals is one of the most spectacular examples of cultural change in recent decades. In the U.S., the percentage of survey respondents who said they know someone who is gay or lesbian rose from less than 25 percent in the early 1980s to nearly 90 percent in 2013 (Pew Research Center, 2013). This surge in visibility remains largely unexplained, likely due to methodological challenges in studying the disclosure of concealable identities, and the lack of data on such disclosures. What is the impact of social networks on individual decisions to disclose a concealable stigmatized identity? A controlled randomization of the exposure to peer disclosure is not feasible: we cannot force or incentivize LGBTQ individuals to publicly disclose their identity due to high potential costs of disclosure. Instead, we use high-frequency longitudinal data from a unique setting to examine the role of peer effects in LGBTQ individuals' decisions to disclose their identity.

In 2019, after violent attacks against a Pride march, thousands of Polish Twitter users posted public tweets with the hashtag *IamLGBT* (*#jestemLGBT*) to increase the visibility of LGBTQ people. We define coming out as the act of disclosing one's LGBTQ identity and classify as coming out only those posts that clearly communicate the author's LGBTQ identity. Coming out publicly on Twitter as part of the #IamLGBT campaign was a costly action. LGBTQ participants faced an increased risk of receiving hateful comments from anti-LGBTQ users and the risk of being identified by friends or family members. Although most users did not use their real names on Twitter, some could be identified because they posted photos of themselves before or during the campaign. Twitter data can also be used by law enforcement and employers. For example, social media posts have been used in U.S. courts as evidence against rioters in the January 6 Capitol attack.[1] Facial recognition companies sell data scraped from social media, including Twitter, to private companies that can use it in the hiring process.[2] Finally, Twitter usernames serve as primary identifiers in virtual relationships on the social media platform. Hence, users risked losing followers when coming

[1] `https://apnews.com/article/media-prisons-social-media-capitol-siege-sentencing-0a60a821ce19635b70681faf86e6526e`

[2] `https://www.buzzfeednews.com/article/ryanmac/clearview-ai-fbi-ice-global-law-enforcement`

out. Fears of being identified were mentioned in several coming out posts, suggesting that these costs were salient to participants.

We are able to determine the exact time of the costly coming out action for hundreds of LGBTQ users. Moreover, we use data on the users' activity patterns prior to the campaign to elicit their networks and generate additional variables. By linking the information about the timing of individual coming out tweets with the pre-campaign networks of LGBTQ users, we obtain a time-varying measure of the users' exposure to peers coming out as LGBTQ. We construct an hourly panel dataset and estimate a duration model to assess the impact of the exposure to peers coming out as LGBTQ on the decisions of individuals to disclose their LGBTQ identity. In other words, we test whether users who came out in a given hour had higher exposure to peers' coming out actions than users who chose to continue concealing their identity.

The exposure to peers coming out as LGBTQ had a strong impact on individuals' decisions to disclose their LGBTQ identity, as a one standard deviation increase in the exposure increased the hourly probability of coming out by 17 percent. The estimation of peer effects in observational data is very challenging due to unobserved heterogeneity (Angrist, 2014). Individuals sort into groups of peers with shared characteristics, and these groups of peers experience correlated shocks. We address these challenges by exploiting the timing of peer coming out for identification using a high-frequency panel data. The results of several additional tests and placebo exercises suggest that unobserved heterogeneity should not have led to an upward bias of our estimates. Instead, the peers-of-peers instrumental variables estimation results indicate that our baseline findings likely underestimate the peer effects.

Our results are not due to the spreading of information about the existence of the campaign through networks, as the effects were strongest after national media coverage of the campaign, and not at the beginning of the campaign. Moreover, individuals do not respond to the tweets posted by anti-LGBTQ peers despite potentially sharing the same information network. We provide suggestive evidence that a social learning mechanism largely explains our findings. First, we show that the magnitude of the peer effects is positively correlated with the positive reception of the peer coming out tweets, and negatively correlated with the intensity of anti-LGBTQ responses to peer tweets. Second, we find that the peer effects were stronger for peers of the same gender than for peers of different genders, which may indicate that observing peers of the same gender coming out provides better information about the level of net social cost an individual will face. Finally, it appears that the estimated results are entirely driven by mutual relationships, and not by fan-idol relationships which again points to the importance of the proximity of peers.

We finish with two additional observations. First, we analyze the impact of the exposure to peers coming out as LGBTQ on the participation of non-LGBTQ users. We find no effects on the participation of anti-LGBTQ users and small positive effects on the participation of straight allies. This suggests that peers' coming out actions do not radicalize anti-LGBTQ users. Second, we study the effects of the exposure to peers' coming out actions on LGBTQ users' post-campaign activity. We find no effects on tweets' sentiment or LGBTQ activism. If anything, the results indicate that coming out influenced by peer effects encourages LGBTQ individuals to speak more openly about their romantic relationships.

This paper makes three main contributions. First, we provide novel evidence on the role of peer effects in cultural change, complementing existing work on female labor supply (Nicoletti et al., 2018), paternity leave (Dahl et al., 2014), and student protest participation (Bursztyn et al., 2021; González, 2020). Thanks to our high-frequency data, our paper is the first to identify peer effects that occur immediately after peer actions. Most importantly, we are the first to study the determinants of disclosing a concealable identity using revealed preferences data. Our results are relevant for other concealable stigmatized identities, such as being a member of an ethnic or religious minority (Kudashvili and Lergetporer, 2022), being a victim of sexual abuse, having undergone an abortion in the past, or living with a chronic illness. Second, we contribute to the literature on LGBTQ people (Badgett et al., 2021; Badgett et al., 2024). Previous studies have found that LGBTQ people face discrimination in labor markets (Aksoy et al., 2018, 2019; Carpenter et al., 2022; Carpenter, 2005; Geijtenbeek and Plug, 2018), and tend to conceal their identity even in anonymous surveys (Coffman et al., 2017), suggesting a high cost of disclosure. We provide first rigorous evidence on the determinants of coming out decisions. Third, we contribute to the literature on social media and collective action, showing that networks matter even when campaign awareness is widespread (Boulianne and Theocharis, 2020; Bursztyn et al., 2019; Fergusson and Molina, 2019; García-Jimeno et al., 2022; Levy, 2021; Zhuravskaya et al., 2020).

2 Background

LGBTQ people in Poland

Our study uses data from Poland, where the levels of visibility of LGBTQ individuals and the levels of support for same-sex marriage in 2019 were similar to those that prevailed in the U.S. in the 1990s (Figure 1). According to a report by a European Union agency, Poland is one of the least LGBTQ-friendly countries in the EU (European Union Agency for Fundamental Rights, 2020). Poland has the highest percentage of LGBTQ people who report always avoiding holding their same-sex partner's hand in public for fear of being assaulted (58 percent compared to the EU average of 30 percent). These fears are not irrational: according to the same study, LGBTQ individuals in Poland experience physical attacks for being LGBTQ more often than their counterparts in any other EU country. According to the ILGA 2021 report, the status of LGBTQ rights is lower in Poland than it is in any other European Union country. Although same-sex sexual activity is legal in Poland, there is no legal recognition of same-sex partnerships, the formal gender recognition procedure is complicated and humiliating, and LGBTQ individuals are not protected from hate crimes and hate speech in Poland (ILGA-Europe, 2021).

The results of a large survey conducted by an European Union agency in 2019 shed light on the patterns of concealment. Figure 2 shows the concealment rates for different LGBTQ identities in various environments, including in public spaces, schools, workplaces, and the family. The concealment rates are universally higher in Poland than in Western Europe, which reflects the differences in the discrimination levels in these places. In all environments in both Poland and Western Europe, lesbian and bisexual women have the lowest concealment rates, while trans people have the highest concealment rates. The differences in the concealment behaviors of LGB women and men are largest in public settings: i.e., when holding hands in the street or when interacting with schoolmates.

After years of very slow progress, the human rights situation of LGBTQ individuals began to deteriorate after the far-right populist *Law and Justice* party won the parliamentary and presidential elections in 2015. Between 2014 and 2020, Poland fell from 23rd to last (27th) place in the ILGA ranking of EU countries. In March 2019, the government, the right-wing media, and Catholic bishops launched a campaign against LGBTQ people after the mayor of Warsaw signed a declaration in support of LGBTQ people. The *Law and Justice* leader described the Warsaw declaration as an attack on the family and children. In the following months, 94 municipalities declared that they are "LGBT-free zones" or are "free from LGBT ideology". Bishops called the LGBTQ movement a "rainbow plague", and right-wing newspapers distributed "LGBT-free zone" stickers. In response to accusations of persecuting LGBTQ people, the campaign leaders claimed they were against "LGBT ideology," not LGBTQ individuals.

At the same time, the mobilization of the LGBTQ community and their allies was as high as ever. A record number of 24 Pride marches were held in 2019, including marches in small towns. Local authorities attempted to ban Pride marches in several towns, but activists successfully challenged these bans in courts. On July 20, 2019, the first Equality March in Białystok was violently attacked by an angry mob inspired by local politicians and clergy: rainbow flags were burned and several people were injured and beaten up. The video reports from Białystok came as a shock to the LGBTQ community, because no similar incidents had occurred at at previous Pride marches, and the police had successfully protected the marches from counter-protests. A few days later, solidarity protests were organized in several towns against the brutal attacks in Białystok. Nevertheless, the *Law and Justice* party continued to use anti-LGBTQ rhetoric and again won parliamentary and presidential elections in 2019 and 2020.

#IamLGBT campaign

The mass coming out campaign started on Twitter on the afternoon of July 29, 2019, nine days after the violent attacks in Białystok. It was started by user *sebastian*, who tweeted:

> Let's f*** with right-wingers, make a hashtag #IamLGBT and post pictures from school and work to show that we are normal people who can be found everywhere in stores, on the streets, in offices, and not some ideology.

In the following hours, thousands of users joined the spontaneous campaign, and the hashtag *#IamLGBT* (*#jestemLGBT*) became the top trending phrase on Polish Twitter. At its peak hour, it was ranked 31st in the top trending phrases worldwide, with more than 34,000 tweets and retweets.[3] The tweets followed the pattern described in the initial tweet: in addition to making a coming out statement, users wrote about their jobs or career plans and that they were not ideology. Approximately 78 percent of users attached photos to their coming out posts. Some users expressed a sense of fear and helplessness, while other posts had humorous elements in them. While some users clearly came out in these posts, others did not come out explicitly (e.g., tweeting the hashtag only). Non-LGBTQ allies ("I'm not LGBT, but I support this campaign"), anti-LGBTQ users sending offensive replies, and large organizations and media outlets also joined in. Examples of tweets can be found in Appendix E.

The campaign was widely echoed in both Polish and foreign media. Major newspapers and TV channels tweeted about the campaign just a few hours after its launch. The European Commission expressed its support for the campaign by tweeting a statement with the hashtag *#IamLGBT*. After the two first days, the campaign began to die out.

The second wave of the campaign took place on May 27, 2020 after a court dismissed a defamation case against a right-wing activist who claimed on TV that "gays want to adopt children to rape them". The second wave was initiated by the same user as the first wave and used the same hashtag. Although the second wave was less successful than the first one in terms of media coverage, several thousand users participated, and the hashtag *#IamLGBT* was once again the top trending hashtags in Poland for several hours (around 20,000 tweets and retweets). Importantly for our study, hundreds of users who were concealing their

[3] The 34,000 figure corresponds to the highest number of tweets and retweets with the hashtag *#IamLGBT* recorded within a 24-hour interval. The number of tweets and retweets posted during the whole period of the campaign was higher but it is impossible to retrieve. See `https://getdaytrends.com/poland/2019-07-29/23/`

identity during the first wave of the campaign decided to come out in the second wave, almost one year later.

Twitter has been one of the most popular social media platforms in Poland. In 2019, Twitter was used by around six million Poles or 16 percent of the Poland's population.[4] For comparison, 22 percent of U.S. adults stated that they used Twitter in 2019.[5] The findings of a study of U.S. Twitter users shed light on patterns of their activity (Pew Research Center, 2019). Around 50 percent of Twitter users use or visit Twitter at least once a day. Nevertheless, the vast majority of Twitter users is largely passive: a median user posts only two tweets per month, and "likes" only one tweet per month. Hence, Twitter may be characterized as a platform with a large group of watchers and a very small group of very active content creators. While there is no comparable study on Polish Twitter users, their behavior is likely very similar to that of the U.S. users' behavior. Hence, we estimate that the coming out posts of LGBTQ users could have reached an audience of up to three million users.

Coming out publicly on Twitter as part of the #IamLGBT campaign was a costly action. LGBTQ participants faced an increased risk of receiving hateful comments from anti-LGBTQ users and the risk of being identified by friends or family members. Although most users did not use their real names on Twitter, some could be identified because they posted photos of themselves before or during the campaign. Twitter data can also be used by law enforcement and employers. For example, social media posts have been used in U.S. courts as evidence against rioters in the January 6 Capitol attack.[6] Facial recognition companies sell data scraped from social media, including Twitter, to private companies that can use it in the hiring process.[7] We know that these costs were salient to participants, as fears of being identified were mentioned in several coming out posts.

[4] https://www.gemius.pl/wszystkie-artykuly-aktualnosci/kobiety-i-mezczyzni-w-sieci.html

[5] https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018

[6] https://apnews.com/article/media-prisons-social-media-capitol-siege-sentencing-0a60a821ce19635b70681faf86e6526e

[7] https://www.buzzfeednews.com/article/ryanmac/clearview-ai-fbi-ice-global-law-enforcement

3 Data

We collected data from Twitter using the Twitter API and additional Python libraries. First, we downloaded the list of all tweets containing the hashtag *#IamLGBT* (*#IamLGBT*) posted during the two waves of the campaign. Second, we manually classified tweets into three groups: LGBTQ coming out tweets, tweets from straight allies, tweets from anti-LGBTQ users. Our classification rules were conservative which resulted in a large number of tweets not being classified into any of these categories (e.g., tweets that consisted of the hashtag only). For all campaign participants, we then retrieved the universe of their tweets posted between January 1 and October 30, 2019. We describe the data collection process in detail in Appendix C. In addition, we provide validity checks of our manual classification using the ChatGPT language model.

We define coming out as the user's first Twitter post that included the hashtag *#IamLGBT* and was classified as indicating their own LGBTQ identity. Individuals with concealable identities disclose their identity to different groups of people over time. We interpret the *#IamLGBT* post as a public disclosure because users, by participating in a viral hashtag campaign and clearly indicating their identity, made their identity visible to Twitter followers and the wider Twitter community. Even if users had come out to some or all of their peers before the campaign, they were aware that their posts could be seen by users previously unaware of their LGBTQ identity, as the information became public and easily accessible to every Twitter user through reposting and searches for the hashtag.

We restrict the sample based on the users' pre-campaign Twitter activity. First, to be included in the sample, users had to have posted at least 25 tweets and 10 replies in the period before the campaign (January 1, 2019 - July 28, 2019), and they had to have a minimum pre-campaign network size (the number of users to whom the user replied) of five. We restrict our analysis to users who posted at least one reply tweet between July 19 and July 28, 2019 (10 days prior to the campaign). Finally, we exclude LGBTQ activists from the sample, as their identity was publicly known before the campaign.

Figure 3 shows that the differences in the Twitter activity of users who came out during the first and the second waves of the campaign were stable before the campaign began. These differences only increased on the day of the Twitter campaign. This was due to an increase in the Twitter activity of the participants in the first wave, rather than to a decrease in the Twitter activity of those who did not join the campaign.

Table 1 presents summary statistics for users who came out during the first wave of the campaign, and for users who decided to conceal their identity during the first wave of the campaign (not-yet-out users). By the end of the first wave, users who came out witnessed a lot of peers coming out: on average, over 8 percent of their network participated in the campaign as LGBTQ users. Not-yet-out users observed many fewer coming out actions in their network. Participants in the first wave also had more peers joining the campaign as allies and as anti-LGBTQ users. There were also some significant differences in the pre-campaign characteristics of the two groups. Women made up the vast majority of the LGBTQ users during both the first and the second wave, and the share of women was higher during the second wave. These gender differences are in line with survey data, which report large differences in the concealment rates of women and men, particularly in public settings (see Figure 2). The average tweet length was significantly greater among the participants in the first wave, while the not-yet-out users were more likely to use emojis in their tweets. Users who decided to come out during the first wave were more likely to post LGBTQ-related tweets, and they had stronger ties with the media, politicians, and LGBTQ activists. These findings confirm the intuition that users with a history of engagement in LGBTQ-related discussions would come out sooner than users who were less politically active and were presumably younger (high emoji use). The averages of the remaining variables (tweet count, reply behavior, hashtag use, emotional words use, tweet sentiment, and network size) were similar for both groups. Compared to LGBTQ users, straight allies and anti-LGBTQ users had much stronger ties to politicians and the media. More than 80 percent of anti-LGBTQ users were men.

The exposure to peers coming out as LGBTQ at the end of the first wave combines peers' coming out actions that occurred before and after an individual posted a coming out tweet which raises concerns about reverse causality. Therefore, we construct an hourly panel dataset with outcome and network variables that vary over time. Our outcome variable, R_{it}, equals one if the user came out by hour t, and equals zero otherwise. Our main network variable is the fraction of the user i's full Twitter network who have disclosed their LGBTQ identity by t ($R_{j,t-1}$) weighted by the strength of pre-campaign ties between users i and j, $\mathbb{G}_{ji}$

$$NI_{it}^{LGBTQ} = \frac{\sum_{j \in \mathbb{I}} \mathbb{G}_{ji} R_{j,t-1}}{\sum_{j \in \mathbb{I}} \mathbb{G}_{ji}} \tag{1}$$

The connection strength is approximated by the number of replies sent from user i to user j in the pre-campaign period. Hence, our measure captures not just the fraction of users in the network who have come out by a given hour, but also the pre-campaign strength for ties to these users. As the exposure to peers coming out can only increase over time, the absolute levels of the exposure are not comparable for users who decided to come out at the beginning and at the end of the campaign. Hence, for each hour, we standardize the network variable with a mean of zero and a standard deviation of one. Standardization is based on the users not yet out at t to avoid simultaneity. We construct similar variables to control for the fraction of peers who joined the campaign as anti-LGBTQ users and straight allies. Table B.4 shows the correlates of the exposure to peers coming out as LGBTQ. There is a strong positive correlation between having LGBTQ activists in the network and the exposure to peers coming out as LGBTQ. Hashtag use, emoji use, and the exposure to media accounts in the network are negatively correlated with the exposure to peers' coming out actions. Women had a lower exposure to peers' coming out actions than men. Combining these findings with descriptives presented in Table 1 suggests that the direction of the potential bias is unclear. On the one hand, unobservables related to gender, emoji use, and the pre-campaign exposure to LGBTQ activists would bias our results upward. These unobserved factors may include the pre-campaign exposure to LGBTQ peers and age. On the other hand, users with a high exposure to media accounts observed few peers' coming out actions while having higher probability of coming out during the first wave of the campaign.

4 Theoretical framework

Why do peer actions matter for individual disclosure decisions? We start from the observation that individuals decide whether to irreversibly disclose their identity by weighing the benefits and costs of disclosure. Individuals have perfect information about certain personal costs and benefits (e.g., their shyness, the foregone implications of pretending to be someone else). At the same time, they face uncertainty about how society will receive their actions (e.g., employer discrimination, stigmatization, verbal abuse, support from others). Hence, this setting combines control over disclosure with uncertainty regarding social costs, which arises precisely because individuals have not yet decided to disclose their identity. We present the detailed model in Appendix A; in this section, we focus on describing the decision process and the theoretical predictions that guide our empirical analysis.

Concealing individuals do not observe their social cost of disclosure. The only way in which they can form their beliefs about the social cost they would face is by observing their peers who decide to come out. Fernández (2013) considers a related process, in which inactive women form their beliefs about disutility from work by observing their peers. The accuracy of the beliefs is particularly fundamental in the context of coming out, because, unlike labor force participation, the disclosure is irreversible. Hence, we can expect that the attention to signals from the network is very high.

We formalize the mechanism through which individuals observe their network and assess the expected gain from taking such action. Net social cost can assume high or low value, which is unknown to concealing individuals, but they have some prior beliefs about the probability of each state. As individuals observe others in their network come out, they immediately update these expected probabilities because they interpret peers' coming out decisions as evidence that their peers perceive the level of net social cost to be low (otherwise, those peers would not come out). Moreover, peer participation may directly influence beliefs about the level of social support an individual can expect from their network after coming out, increasing the perceived likelihood of the low-social-costs scenario. Once an individual believes that the high level of net social cost is unlikely, they take the risky action.

We show that for each individual, we can find a unique threshold of coming outs from the network, above which they will decide to come out. Peer similarity is crucial for the size of the peer effects. In the most extreme case, where all individuals have identical personal benefits of coming out, a peer's decision to come out would serve as a perfect signal of their beliefs about the level of net social cost. The welfare effects of actions, as described in the model, are ambiguous. Individuals make an irreversible decision based on their beliefs, and their final utility may differ from their ex-ante expectations. Due to idiosyncratic gains from coming out, individuals may mistake others' actions as more or less costly than they actually are.

We take the model to the data to test two major hypotheses. First, we test whether users with greater exposure to peers coming out as LGBTQ are more likely to disclose their own identity. Second, we test whether the strength of peer effects depends on peer proximity, measured along various dimensions of similarity. Additionally, we rule out a major alternative mechanism not included in our model. Since our empirical setting involves a viral campaign, peers may influence individual disclosure simply by spreading awareness of the campaign. To use a hashtag, one must first know that it exists. This information channel does not involve any change in perceived social costs or benefits.

Peer effects operating through the information channel should be strongest at the beginning of the campaign, when awareness of the campaign is not yet widespread. Moreover, if peer effects are purely about spreading information, the size of the effect should not depend on peers' characteristics or the type of relationship. Our second hypothesis contradicts this mechanism. Additionally, we test the relevance of the information channel by examining differences in the size of the effect depending on the stage of the campaign, the peer's type (LGBTQ vs. anti-LGBTQ users), and the relationship type (mutual followers vs. idol-fan relationships).

5 Empirical strategy

We employ two approaches to assess the role of social networks in individual coming out decisions during the first wave of the #IamLGBT campaign. Since coming out is an absorbing state, we cannot analyze changes in the probability of coming out over time for a single user using a panel fixed-effect estimator. Therefore, in our baseline approach, we estimate a simple discrete approximation of a duration model (Corno et al., 2020; Currie and Neidell, 2005) to test whether exposure to peer coming out as LGBTQ affects the hourly probability of public coming out.

In our baseline approach, we estimate the probability of coming out of user i at hour t as follows:

$$R_{i,t} = \kappa NI_{i,t-1}^{LGBTQ} + \beta X_i + \gamma_1 NI_{i,t-1}^{ally} + \gamma_2 NI_{i,t-1}^{anti-LGBTQ} + \alpha_t + \epsilon_{i,t} \tag{2}$$

where $R_{i,t}$ is a binary variable coded as one in the hour the user decides to come out, and as zero otherwise. $NI_{i,t-1}^{LGBTQ}$ is the weighted fraction of the network who have disclosed their LGBTQ identity before hour t. We control for time-invariant factors (gender, pre-campaign measures of Twitter activity), X_i, time-varying exposure to allies' and anti-LGBTQ users' tweets ($NI_{i,t-1}^{ally}$ and $NI_{i,t-1}^{anti-LGBTQ}$), and hour fixed effects (α_t). κ is the main coefficient of interest. Since $NI_{i,t-1}^{LGBTQ}$ is standardized for each hour, positive values of κ mean that, at the time of their coming out, the disclosing users had consistently higher exposure to peers' coming out actions than the concealing users. In our sample, we include users who came out in the first wave of the campaign, and users who came out during the second wave of the campaign. Hence, our sample includes users who did not choose to come out during the study period (first wave of the campaign).

The estimation is performed on user-hour panel data. Users are included in the dataset until they come out, after which they exit the data. We start our analysis from the fourth hour of the campaign, as the first anti-LGBTQ and ally users joined the campaign during the third hour of the campaign. We end our analysis after 54 hours of the campaign because up until that point, at least one participant joined every hour, except at night (see Figure 4). After that time, there were only a few users coming out with long gaps in between.

The major challenges we face in identifying the effects of peers coming out as LGBTQ is omitted variable bias. We address unobserved heterogeneity in several ways. It is likely that exposure to peer LGBTQ coming out actions is correlated with factors that increase the probability of coming out, including having more LGBTQ peers, location, personality traits, education, age, and the attitudes of friends and family. We are able to provide some insights into the potential role of unobserved characteristics by controlling for a rich set of covariates including the exposure to LGBTQ activists in the network, and the share of LGBTQ-related posts in the pre-campaign period. In addition, the variation of the effects over time may also shed light on the role of unobservable factors. If exposure to peer LGBTQ coming out actions was correlated with characteristics associated with a lower individual cost of coming out, we should observe the strongest effects in the early phases of the campaign. In robustness checks, we control for more distant lags of exposure to LGBTQ coming out actions to alleviate the concern that our results are driven by strong pre-campaign ties with LGBTQ peers, rather than by exposure to peers coming out. To further address this concern, we estimate the duration model for users who participated in the first wave of the campaign, in which we control for their exposure to LGBTQ coming out actions from the second wave of the campaign (coming out actions of LGBTQ users who did not participate in the first wave). Although coming out decisions during the second wave may have been influenced by peers' coming out actions that occurred during the first wave, these effects should be similar regardless of the hour at which the peers joined the first wave of the campaign.

We further address the issue of unobserved heterogeneity by applying the instrumental strategy proposed by Bramoullé et al. (2009) and De Giorgi et al. (2010). This method exploits exogenous shocks to distance 3-nodes by analyzing the peers of the peers of a user who are not the user's peers themselves. This method has been previously used in studies on network effects in consumption (De Giorgi et al., 2020), and peer effects in labor supply (Nicoletti et al., 2018) among others. In the case of our study, we instrument the user's exposure to LGBTQ coming out actions by the exposure of the user's peers to LGBTQ coming out actions, excluding the peers of the peers who are the user's peers themselves. The instrument is given by:

$$NI_{it}^{LGBTQ,IV} = \sum_{j \in \mathbb{I}} \omega_{ji} \frac{\sum_{k \in \mathbb{J}, k \notin \mathbb{I}} \mathbb{G}_{kj} R_{k,t-1}}{\sum_{k \in \mathbb{J}, k \notin \mathbb{I}} \mathbb{G}_{kj}} \tag{3}$$

where ω_{ij} measures the importance of peer j in the user i's network (given by the share of replies sent to the user j), $\mathbb{G}_{kj}$ is the number of replies sent by the user j to the user k in the pre-campaign period, and $R_{k,t-1}$ is a dummy variable that equals one if a user has joined the campaign by hour $t-1$, and equals zero otherwise.

Participants in the campaign were connected to more than 160,000 unique users. In order to compute the exposure of the peers, we need to download data of connected users. For computational reasons, we restricted our analysis to the campaign participants for whom we already had data, and for the most influential peers. In order to select the most influential peers we calculated two variables for each peer: total number of replies sent by the participants to the peer, and the maximum value of importance weight ω_{ji} for the given peer j. The first variable measures the popularity of a peer among the participants, and the second variable measures the maximum importance of the user j. We selected users who were in the top decile of at least one of these variables. We then downloaded their Twitter activity data, and included them in our analysis. On average, the peer data cover around 60 percent of the user's network. Figure B.3 presents the distribution of the network coverage. We were unable to generate instruments for only four users.

6 Results

Exposure to coming out actions in the users' networks had a significant impact on individual coming out decisions, as it substantially increased the probability of LGBTQ individuals publicly coming out (see Table 2). A one standard deviation increase in the exposure to LGBTQ coming out actions increased the hourly probability of coming out by nearly 0.5 percentage points, or 17 percent of the average probability of coming out. The estimates of the effect remained stable after controlling for gender, pre-campaign measures of Twitter activity, and network variables. Hence, these effects were not related to prior Twitter behavior, or to strong connections to media, politics, or LGBTQ activist accounts. Figure 5 descriptively illustrates these results: the probability of coming out following the first peer coming out converged to the probability of coming out for other users (the gap decreased by over 40 percent 10 hours after the first peer came out). The increase in the probability of coming out was particularly large in the hour immediately following the first peer coming out.

Our results could be biased upwards if the exposure to peers coming out as LGBTQ was positively correlated with unobservable factors that facilitate coming out. The exposure to LGBTQ peers is a major unobservable characteristic that may be positively correlated with exposure to peers coming out as LGBTQ. We control for more distant lags of exposure to LGBTQ coming out actions to address this issue. Table 3 shows that more distant lags of exposure have no significant impact on coming out, and our main estimates remain stable. Thus, our estimated peer effects are driven by the coming out actions of peers in the hour immediately preceding the coming out decision, rather than by constantly higher exposure to coming out actions due to the large share of LGBTQ peers in users' networks. In an alternative approach, we show that controlling for the fraction of the network who came out during the second wave of the campaign does not change the estimated peer effects (Table B.6). We also estimate an enhanced specification that includes additional control variables (measures of tweet emotions, additional tweet topics, tweet topics with the list of keywords extended using word embeddings, pre-campaign ties to various type of accounts, grammatical gender and person use) selected using the LASSO procedure (Belloni et al., 2014). The results remain unchanged (see Table B.7). Finally, we investigate the importance of unobservables in the spirit of Oster (2019), and we find that our results are robust to this test (Figure B.2). All these exercises suggest that the selection on unobservables that facilitate coming out is unlikely to affect our results.

The correlation of Twitter activity between peers is another potential source of bias. Groups of peers may be active during a particular time of the day, e.g., after a screening of an episode of a TV show a given group of peers tends to watch. If the exposure to peers coming out as LGBTQ simply reflects stronger activity of peers in a given hour, and there is a positive correlation between the activity of users and their peers, our estimates would be biased upward. We rule out this source of bias in two placebo exercises. First, we estimate the effects of exposure to peer LGBTQ coming out actions on the probability of posting a tweet depending on the topic of the tweet. In addition to our baseline estimates, we also estimate strong effects of the exposure to peers coming out as LGBTQ on the probability of posting LGBTQ-related tweets during the campaign using the fixed-effects estimator (see Table 4). We find no such positive effects on tweets that are unrelated to the LGBTQ topic, which suggests that the exposure to peers coming out as LGBTQ was important only for the coming out tweets. In an additional placebo test, we show that the exposure to peer tweets unrelated to the LGBTQ topic had no impact on coming out, which suggests that the estimated effects do not reflect differences in the overall Twitter activity of peers over time (Table B.9).

Our estimates may be biased downward. Indeed, our estimates should be interpreted as intention-to-treat effects, because we do not observe the users reading peer tweets, and users may, for example, have been offline at the moment when one of their peers came out. We use an instrumental variable strategy that addresses both the problem of incomplete information about peers' coming out actions due to being offline, and the network endogeneity discussed above. We instrument user i's exposure to peers coming out as LGBTQ by the coming out actions of the peers of user i's peers who are not peers in user i's network, following the approach proposed by Bramoullé et al. (2009) and De Giorgi et al. (2010). Using this approach, we obtain estimates that are twice as large as our baseline estimates (Table 5). A one standard deviation increase in the exposure to peers coming out increases the instantaneous probability of coming out by more than 30 percent. The direction of the bias is the same in case of log exposure (Table B.10).

Another source of the potential downward bias is sample selection. We are unable to include all Polish LGBTQ users in the sample, because we observe LGBTQ identity only when it is publicly disclosed. If peer effects are stronger for users with high costs of coming out, our estimates would be biased downward, because we are limited to analyzing users who decided to come out. This intuition is confirmed by our observation that our baseline results estimated on the sample that includes "not-yet-out" participants in the second wave are somewhat larger than the effects estimated using a sample of first wave participants only (Table B.6). This finding, combined with the IV results, suggest that our baseline results provide the lower bound of the effects of peers coming out.

It is likely that LGBTQ users had private knowledge about LGBTQ identity of their peers prior to the campaign. Hence, our estimates primarily capture the effects of learning about LGBTQ peers' decision to publicly come out and not the effects of learning about peers' identity. In other settings, e.g., at schools, learning about peers' identity may further increase the size of peer effects.

Mechanism: information channel

Figure 6 shows that observing anti-LGBTQ posts had no significant influence on the individual decision to come out. We found that posts from straight allies increased the instantaneous probability of coming out, but this effect is weaker than the effect of observing coming out posts from LGBTQ peers. The difference in the size of the effects suggests that the estimated effects are not simply due to the spreading of information about the existence of the campaign through networks.

Next, we study the variation in the magnitude of the peer effects over time. Figure 7 shows that that there was considerable variation in the magnitude of peer effects over time. At the beginning of the campaign, the role of peer effects was limited. This is surprising and again suggests that the spread of information about the existence of the campaign through networks was not the mechanism driving the results. We find largest effects on the second and third days of the campaign. At this point, knowledge about the campaign was widespread, as it was covered by the major media outlets. We also find that the size of the peer effects decreases during the nighttime hours. This can be explained in two ways. First, the nighttime hours were characterized by much lower activity (see Figure 4). Since the effect of observing a peer coming out was decreasing over time, the small peer effects during the nighttime hours could be driven by the low number of coming out tweets during these hours. Second, it is possible that the users who were encouraged by their peers coming out wanted to see their peers' reactions to their own coming out, so they postponed this action until the hours when there were more users on Twitter.

Mechanism: social learning

We have shown that the information channel does not explain the estimated effects. One alternative mechanism proposed in our theoretical model is that perceptions of the costs of coming out change as a result of observing peers disclosing their LGBTQ identity. To provide suggestive evidence of a social learning mechanism, we examine variation in the size of the effects depending on the proximity of peers, and public reactions to peers coming out.

Figure 8 shows that the peer effects increase with peer proximity, measured in several ways. First, we observe that the effects of peers coming out as LGBTQ are somewhat stronger for peers of the same gender than for peers of different genders (Figure 8a). This finding is consistent with the hypothesis that individuals update their beliefs about the costs of coming out by observing their peers deciding to disclose their identity, as the coming out actions of peers of an individual's own gender provide better information about social costs an individual would face after coming out. Second, peer effects are strongest in small circles of friends: the coming out actions of very popular peers has a lower impact on coming out decisions than the coming out action of a peer who is linked to a small number of friends (Figure 8b). Finally, we find that peer effects are solely driven by mutual relationships (Figure 8c). Hence, it seems that the proximity of peers is important in explaining our results.

Next, we study the variation in the effects depending on public reactions to peer coming outs. The social learning mechanism suggests that the peer effects should be particularly strong if the peers' coming out actions received positive reactions. The number of likes received by a coming out tweet is our proxy for public reactions to that coming out action. We can distinguish between the effects of exposure to peers' coming out actions that received only a few likes and effects of exposure to peers' coming out actions that received many likes. Figure 9a shows that the coming out actions of peers that received particularly positive reactions had a substantially greater impact than the coming out posts that received only a few likes. This is due to the popularity of coming out tweets rather than to the pre-campaign popularity of peers, as we find no significant relationship between the pre-campaign popularity of peers and the size of the peer effects (Figure B.4). On the other hand, negative reactions to peers coming out may reduce the size of the peer effects: we find that the coming out actions of peers that were targeted by anti-LGBTQ users had smaller effects than the coming out posts that did not attract hateful comments (Figure 9b). These findings provide suggestive evidence that public reactions to peers' coming out tweets are important in explaining the size of peer effects.

Robustness and heterogeneity

Our results are not driven by the strong networks of public figures, as the effects do not change after excluding journalists, elected officials, and political party members from the sample (Table B.11). We use a more restrictive definition of coming out—requiring users to have attached a photo to their coming out post—and find effects identical to the baseline results (Table B.12). The estimates do not change after excluding not-yet-out users who were not active during the first wave of the campaign (Table B.13). Robust standard errors and two-way clustered standard errors are very similar to our baseline standard errors clustered at the user level (Table B.14). Using unweighted network variables (share of users, instead of share of users weighted by the strength of pre-campaign ties), we estimate slightly smaller but still statistically significant effects (Table B.15). This is expected, as the strength of the pre-campaign ties is one of the sources of variation that we use (for example, the strength of ties increases the probability that the peer's tweet would appear on the user's Twitter timeline). In our main specification, we use cumulative exposure to peers' LGBTQ coming out actions. Using the measures of exposure restricted to the peer coming out actions occurring the preceding hour only (a continuous and a binary variable) yields similar results (Table B.16). Our baseline sample does not include the second and the third hour of the campaign, as the first straight allies and anti-LGBTQ users joined the action during the third hour of the action, and we use standardized lag network variables. Including these hours (and dropping the exposure to peer allies and anti-LGBT peers from the set of control variables) does not change the results (Table B.17). The potential errors in manually classifying coming out posts have no impact on the results. We use ChatGPT to classify coming out posts, restrict the sample to users recognized as LGBTQ by ChatGPT, generate a modified exposure variable, and obtain effects that are identical to the baseline estimates (Table B.18). We estimate the duration model with log network variables to account for their skewed distribution, and the effects are equally strong (Table B.19). Standardizing independent variables is not a common practice. We find clear positive effects using alternative, non-standardized measures of peer exposure: the weighted fraction of the network, the log-weighted fraction of the network, and a dummy variable indicating any exposure (Table B.20). The estimation of a Cox hazard model and parametric survival models yields similar results (Tables B.21-B.23). Finally, the estimated effects are not sensitive to the choice of the sample restriction thresholds (see Figure B.5).

We find no significant variation in the size of the effect of peers coming out as LGBTQ depending on gender, exposure to LGBTQ activists, exposure to the media, emoji use, or pre-campaign Twitter activity levels (see Figures B.6-B.10). While the peer effects seem to be stronger for users who mostly post original tweets than for users who mostly engage in discussions under other users' posts, the difference is statistically insignificant (Figure B.11).

One additional concern is that the peer effects may be limited to viral campaigns characterised by a large number of tweets within a very short time span. It is possible that the estimated effects are the result of the interaction of the exposure to peers coming out as LGBTQ and the existence of the ongoing viral campaign. While most of coming out actions occurred during one of the two waves, a small number of users decided to come out using the hashtag *IamLGBT* during the months between the first and the second waves. These coming out actions were not part of a viral campaign. We calculate the exposure to peers' coming out actions (at the end of the first wave) for two groups of users: those who decided to come out between the two waves, and those who came out during the second wave. Table B.24 shows that users who decided to come out between the two waves had a higher exposure than the participants of the second wave. Although the small sample size reduces the precision of the estimates, the magnitude of the effect is similar to the baseline, as a one standard deviation increase in the exposure is associated with an approximate 20 percent increase in the probability of coming out. This suggests that the influence of peers' coming out actions continues even after the viral phase of the campaign has ended.

Effects on the participation of non-LGBTQ users

The exposure to peers coming out as LGBTQ increases the probability of LGBTQ users disclosing their identity. But does it affect the participation of non-LGBTQ users in a viral campaign as well? The exposure to peers coming out as LGBTQ may encourage supportive statements from non-LGBTQ users for two reasons. First, peers' LGBTQ coming out actions increase the salience of the LGBTQ topic among non-LGBTQ users who had been supportive of LGBTQ people prior to the action. Second, according to the contact hypothesis, the disclosure of a stigmatized identity may reduce the prejudices held by a majority (Allport, 1954). This may lead to an increase in supportive statements, and to a reduction in anti-LGBTQ statements. On the other hand, if they are perceived as a violation of a social norm, LGBTQ coming out actions may evoke discomfort in non-LGBTQ peers (Akerlof and Kranton, 2000), and lead to a backlash.

We find small positive, but statistically insignificant, effects of the exposure to peers coming out as LGBTQ on the participation of straight allies (Table 6). The participation of anti-LGBTQ users is unaffected by observing peers coming out as LGBTQ. Hence, we find no evidence that increased visibility of LGBTQ individuals leads some of their non-LGBTQ peers to engage in anti-LGBTQ behaviors.

Instead, what seems to encourage anti-LGBTQ users to join the campaign is the exposure to tweets posted by their non-LGBTQ peers: straight allies and other anti-LGBTQ users. The participation of straight allies was crucial for the whole campaign, as it significantly increased the participation probabilities of users of all types. Hence, supportive non-LGBTQ users contributed to an increase in the polarization in LGBTQ narratives on Twitter. Nevertheless, the participation of supportive non-LGBTQ users had a net positive impact on the sentiments of the LGBTQ narratives, as the effects on their peers' participation as straight allies were twice as large as the effects on the participation of their anti-LGBTQ peers, and they additionally encouraged their LGBTQ peers to join the campaign.

Effects on post-campaign activity

We have shown that exposure to peers coming out as LGBTQ increases the probability of coming out. Nevertheless, the impact of coming out on the welfare of LGBTQ individuals is ambiguous in our theoretical model. The ultimate effect depends on the trade-off between the costs of concealing one's identity and the costs of discrimination that may occur after disclosure. If coming out induced by peer effects is met with substantial social stigma or discrimination, it may have adverse effects on the welfare of LGBTQ individuals.

While we have no data on users' health, we analyze the effects on various dimensions of Twitter activity: frequency of posting, interactions with other users, as well as tweets' sentiment and topics. To this end, we construct a weekly panel of Twitter activity that covers four weeks before the first wave of the campaign and four weeks following the first wave. We limit the analysis to eight summer weeks, as the tweet content changes substantially after the beginning of the school year. For the campaign week variables, we include only those tweets that do not contain the *IamLGBT* hashtag to avoid mechanical effects.

We use standard difference-in-differences method, in which we use the exposure to peers coming out as LGBTQ as a treatment variable. User i's exposure to peers coming out as LGBTQ at the end of the first wave of the campaign conflates peers' coming out actions that affected user i's coming out decision and peers' coming out actions that were caused by user i's coming out. Therefore, in our sample, we include users who had not decided to come out before the 19th hour of the campaign, and we calculate their exposure to peers coming out as LGBTQ at this hour. In the DiD analysis we use a dummy treatment variable, T_i, which equals zero for users with exposure below the mean, and one for those with exposure equal or greater than mean. The exposure variable strongly predicts the probability of coming out by the end of the first wave (Table B.25).

We estimate a standard difference-in-differences equation using the fixed effects estimator:

$$Y_{i,t} = \theta T_i \times PostCampaign_t + \beta_t + \alpha_i + \epsilon_{i,t} \tag{4}$$

where $PostCampaign_t$ is a dummy variable that is equal to one for the campaign week and three subsequent weeks, and zero for the four weeks preceding the launch of the campaign. We include individual fixed effects, α_i, to account for time-invariant unobserved heterogeneity. Additionally, β_t represents week fixed effects that control for any common time-specific factors that may influence the outcome variable across all individuals. The coefficient of interest, θ, captures the impact of coming out on an outcome variable. All outcome variables are standardized with zero mean and standard deviation of one. We restrict our sample to users who posted at least one tweet in each of the eight weeks to achieve a balanced panel. For all studied outcomes, we cannot reject the parallel trend assumption (Figure B.12).

Our findings indicate minimal and statistically insignificant effects on posting frequency, tweet sentiment, and the frequency of tweets related to illness (see Figure 10). These findings should be interpreted with caution, as predicting users' mental health based on their Twitter activity is challenging (Jaidka et al., 2020; Kelley et al., 2022). Hence, our Twitter health variable likely suffers from a substantial measurement error, and we cannot rule out significant effects on actual health of affected individuals.

Additionally, our analysis indicates that participation in the campaign had no immediate impact on everyday LGBTQ activism, as measured by the frequency of tweets containing LGBTQ-related keywords, nor did it influence network formation. Specifically, we observe that LGBTQ users do not increase their interactions with other openly LGBTQ users following the campaign. However, it is worth noting that coming out may have encouraged users to express themselves more openly about their personal relationships, as we observed marginally significant positive effects on tweets discussing boyfriends and girlfriends. Importantly, for all outcomes, we cannot rule out small effects due to the limited precision of our estimates.

7 Conclusion

The surge in the visibility of LGBTQ individuals in recent decades remains largely unexplained, as data on LGBTQ identity disclosure are scarce. Recently, social media have become important platforms for coming out. We collected unique data from two waves of a spontaneous Twitter coming out campaign in Poland. We use these data to investigate the hypothesis that observing LGBTQ peers coming out increases the probability that an individual would make a decision to disclose her LGBTQ identity. We find significant peer effects that cannot be explained by the information channel. In addition, the estimated effects are not attributable to higher pre-campaign exposure to LGBTQ peers, but rather to immediate reactions to peers' coming out actions. We argue that the effects can be attributed to the social learning mechanism, as we find that the magnitude of the effects is a function of peer proximity and positive reactions to peers' coming out posts.

Our findings suggest that coming out actions of role models have limited immediate effects. The estimated effects are driven solely by mutual relationships, as we found no effects of coming out actions in idol-fan relationships. There are two potential explanations for this observation. First, it is possible that LGBTQ fans do not update their perceived social costs of disclosure after the coming out of a role model because it is not a good signal of what they may experience if they came out themselves. Second, it is possible that the effects of the coming out of a role model are not immediate, unlike the coming out actions of close peers.

The proposed mechanism of social learning may help explain why the number of individuals coming out as LGBTQ has increased so rapidly in recent decades. It was previously shown by Bursztyn et al. (2020) that individuals tend to underestimate changes in restrictive social norms. By coming out, LGBTQ individuals send a signal about social costs to their peers. This, in turn, influences how their peers perceive the costs of coming out themselves. In our study of short-term peer effects, we abstract from changes in the actual levels of social cost, as they could not change during the 54 hours of the Twitter campaign. The extent to which the social learning mechanism may affect the visibility of LGBTQ people in the long run depends on a reduction in the actual levels of discrimination. Importantly, Fernández et al. (forthcoming) suggested that social attitudes toward LGBTQ people may be endogenous to their visibility, as exposure to peers coming out may improve the acceptance of LGBTQ people. This would imply that even in the absence of other favorable conditions, a short-term increase in the visibility of LGBTQ people due to estimated peer effects may decrease the discrimination levels, and lead to a long-term improvement in visibility of LGBTQ people.

References

Acemoglu, Daron, Munther A. Dahleh, Ilan Lobel, and Asuman Ozdaglar (2011). "Bayesian Learning in Social Networks". In: *The Review of Economic Studies* 78.4, pp. 1201–1236. DOI: 10.1093/restud/rdr004.

Akerlof, George A. and Rachel E. Kranton (2000). "Economics and Identity". In: *The Quarterly Journal of Economics* 115.3, pp. 715–753. DOI: 10.1162/003355300554881.

Aksoy, Cevat Giray, Christopher S. Carpenter, and Jeff Frank (2018). "Sexual Orientation and Earnings: New Evidence from the United Kingdom". In: *ILR Review* 71.1, pp. 242–272. DOI: 10.1177/0019793916687759.

Aksoy, Cevat Giray, Christopher S. Carpenter, Jeff Frank, and Matt L. Huffman (2019). "Gay glass ceilings: Sexual orientation and workplace authority in the UK". In: *Journal of Economic Behavior & Organization* 159, pp. 167–180. DOI: 10.1016/j.jebo.2019.01.013.

Allport, Gordon (1954). *The Nature of Prejudice.* Addison-Wesley.

Angrist, Joshua D (2014). "The perils of peer effects". In: *Labour Economics* 30, pp. 98–108.

Badgett, Lee, Christopher S Carpenter, and Dario Sansone (2021). "LGBTQ Economics". In: *Journal of Economic Perspectives* 35.2, pp. 141–70. DOI: 10.1257/jep.35.2.141.

Badgett, MV Lee, Christopher S Carpenter, Maxine J Lee, and Dario Sansone (2024). "A Review of the Economics of Sexual Orientation and Gender Identity". In: *Journal of Economic Literature* 62.3, pp. 948–994.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen (2014). "High-Dimensional Methods and Inference on Structural and Treatment Effects". In: *Journal of Economic Perspectives* 28.2, pp. 29–50. DOI: 10.1257/jep.28.2.29.

Boulianne, Shelley and Yannis Theocharis (2020). "Young People, Digital Media, and Engagement: A Meta-Analysis of Research". In: *Social Science Computer Review* 38.2, pp. 111–127. DOI: 10.1177/0894439318814190.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin (2009). "Identification of peer effects through social networks". In: *Journal of Econometrics* 150.1, pp. 41–55. DOI: 10.1016/j.jeconom.2008.12.021.

Bursztyn, Leonardo, Davide Cantoni, David Y. Yang, Noam Yuchtman, and Y. Jane Zhang (2021). "Persistent Political Engagement: Social Interactions and the Dynamics of Protest Movements". In: *American Economic Review: Insights* 3.2, pp. 233–50. DOI: 10.1257/aeri.20200261.

Bursztyn, Leonardo, Georgy Egorov, Ruben Enikolopov, and Maria Petrova (2019). "Social Media and Xenophobia: Evidence from Russia". In: *Working Paper.*

Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott (2020). "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia". In: *American Economic Review* 110.10, pp. 2997–3029. DOI: 10.1257/aer.20180975.

Carpenter, Christopher S, Maxine J Lee, and Laura Nettuno (2022). "Economic outcomes for transgender people and other gender minorities in the United States: First estimates from a nationally representative sample". In: *Southern Economic Journal* 89.2, pp. 280–304.

Carpenter, Christopher S. (2005). "Self-Reported Sexual Orientation and Earnings: Evidence from California". In: *ILR Review* 58.2, pp. 258–273. DOI: 10.1177/001979390505800205.

CBOS (2021). *Stosunek Polaków do osób homoseksualnych.*

Coffman, Katherine B., Lucas C. Coffman, and Keith M. Marzilli Ericson (2017). "The Size of the LGBT Population and the Magnitude of Anti-Gay Sentiment are Substantially Underestimated". In: *Management Science* 63.10, pp. 3168–3186.

Corno, Lucia, Nicole Hildebrandt, and Alessandra Voena (2020). "Age of Marriage, Weather Shocks, and the Direction of Marriage Payments". In: *Econometrica* 88.3, pp. 879–915. DOI: 10.3982/ECTA15505.

Currie, Janet and Matthew Neidell (2005). "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" In: *The Quarterly Journal of Economics* 120.3, pp. 1003–1030. DOI: 10.1093/qje/120.3.1003.

Dahl, Gordon B., Katrine V. Løken, and Magne Mogstad (2014). "Peer Effects in Program Participation". In: *American Economic Review* 104.7, pp. 2049–74. DOI: 10.1257/aer.104.7.2049.

De Giorgi, Giacomo, Anders Frederiksen, and Luigi Pistaferri (2020). "Consumption Network Effects". In: *The Review of Economic Studies* 87.1, pp. 130–163.

De Giorgi, Giacomo, Michele Pellizzari, and Silvia Redaelli (2010). "Identification of Social Interactions through Partially Overlapping Peer Groups". In: *American Economic Journal: Applied Economics* 2.2, pp. 241–75. DOI: 10.1257/app.2.2.241.

European Union Agency for Fundamental Rights (2020). *A long way to go for LGBTI equality.* European Union Agency for Fundamental Rights. URL: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-lgbti-equality-1_en.pdf.

Fergusson, Leopoldo and Carlos Molina (2019). "Facebook Causes Protests". In: *Working Paper.*

Fernández, Raquel (2013). "Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century". In: *American Economic Review* 103.1, pp. 472–500. DOI: 10.1257/aer.103.1.472.

Fernández, Raquel, Sahar Parsa, and Martina Viarengo (forthcoming). "Coming Out in America: Thirty Years of Cultural Change". In: *Journal of Law, Economics & Organization.*

Gallup (2021). *Gallup Poll Social Series: Values and Beliefs.*

García-Jimeno, Camilo, Angel Iglesias, and Pinar Yildirim (2022). "Information Networks and Collective Action: Evidence from the Women's Temperance Crusade". In: *American Economic Review* 112.1, pp. 41–80. DOI: 10.1257/aer.20180124.

Gaudecker, Hans-Martin von (2014). "Templates for Reproducible Research Projects in Economics". https://github.com/hmgaudecker/econ-project-templates.

Geijtenbeek, Lydia and Erik Plug (2018). "Is there a penalty for registered women? Is there a premium for registered men? Evidence from a sample of transsexual workers". In: *European Economic Review* 109, pp. 334–347. DOI: 10.1016/j.euroecorev.2017.12.006.

González, Felipe (2020). "Collective action in networks: Evidence from the Chilean student movement". In: *Journal of Public Economics* 188, p. 104220. DOI: 10.1016/j.jpubeco.2020.104220.

ILGA-Europe (2021). *Annual Review of the Human Rights Situation of Lesbian, Gay, Bisexual, Trans and Intersex People in Europe and Central Asia.* ILGA-Europe. URL: https://rainbow-europe.org/annual-review.

Jaidka, Kokil, Salvatore Giorgi, H Andrew Schwartz, Margaret L Kern, Lyle H Ungar, and Johannes C Eichstaedt (2020). "Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods". In: *Proceedings of the National Academy of Sciences* 117.19, pp. 10165–10171.

Joulin, Armand, Edouard Grave, Piotr Bojanowski, Matthijs Douze, Hérve Jégou, and Tomas Mikolov (2016). "FastText.zip: Compressing text classification models". In: *arXiv preprint arXiv:1612.03651.*

Kelley, Sean W, Caoimhe Ní Mhaonaigh, Louise Burke, Robert Whelan, and Claire M Gillan (2022). "Machine learning of language use on Twitter reveals weak and non-specific predictions". In: *NPJ Digital Medicine* 5.1, p. 35.

Kudashvili, Nikoloz and Philipp Lergetporer (2022). "Minorities' strategic response to discrimination: Experimental evidence". In: *Journal of Public Economics* 208, p. 104630. DOI: 10.1016/j.jpubeco.2022.104630.

Levy, Ro'ee (2021). "Social Media, News Consumption, and Polarization: Evidence from a Field Experiment". In: *American Economic Review* 111.3, pp. 831–70. DOI: 10.1257/aer.20191777.

Nicoletti, Cheti, Kjell G. Salvanes, and Emma Tominey (2018). "The Family Peer Effect on Mothers' Labor Supply". In: *American Economic Journal: Applied Economics* 10.3, pp. 206–34. DOI: 10.1257/app.20160195.

Oster, Emily (2019). "Unobservable Selection and Coefficient Stability: Theory and Evidence". In: *Journal of Business & Economic Statistics* 37.2, pp. 187–204. DOI: 10.1080/07350015.2016.1227711.

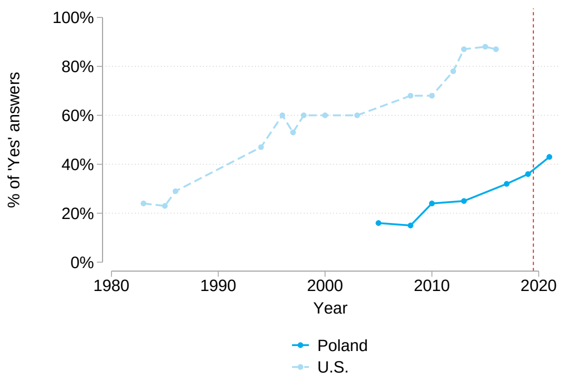
Pew Research Center (2013). *A survey of LGBT Americans: Attitudes, experiences and values in changing times.*

— (2019). *Sizing Up Twitter Users.*

Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov (2020). “Political Effects of the Internet and Social Media”. In: *Annual Review of Economics* 12.1, pp. 415–438. DOI: 10.1146/annurev-economics-081919-050239.

Figures

(a) Do you personally know someone who is gay or lesbian?



(b) Do you think same-sex couples should be allowed to marry?

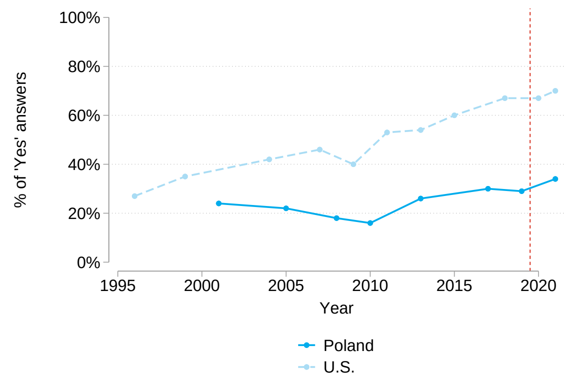
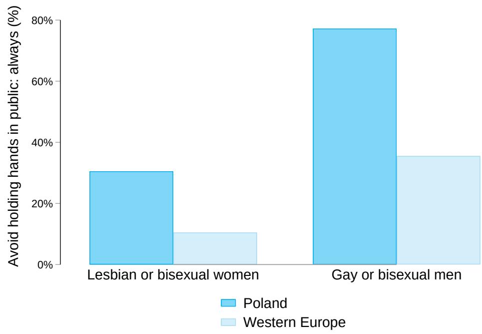


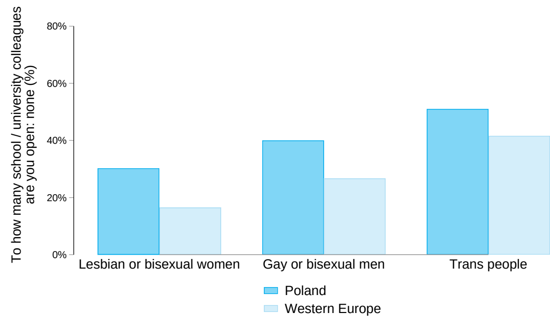
Figure 1: Evolution of the visibility of gay men and lesbians, and the support for same-sex marriage in the U.S. and Poland

Notes: Figure 1a shows the visibility of gay men and lesbians in the U.S. and Poland. In Poland (2005-2021), and in the U.S. (2013-2016) the question asked was: "Do you personally know anyone who is gay or lesbian, or not?" (CBOS, 2021). In the US (1983-2012), the question asked was: "Do you have a friend or acquaintance who is gay or lesbian?" (Fernández et al., forthcoming). Figure 1b shows the support for same-sex marriage in the U.S. and Poland. In Poland, the question asked was: "Do you think same-sex couples should be allowed to marry?" (CBOS, 2021). In the U.S., the question asked was: "Do you think marriages between same-sex couples should or should not be recognized by the law as valid, with the same rights as traditional marriages?" (Gallup, 2021). The vertical red line denotes the beginning of the *#IamLGBT* campaign.
Data: CBOS, Pew Research Center, Gallup, Fernández et al. (forthcoming).

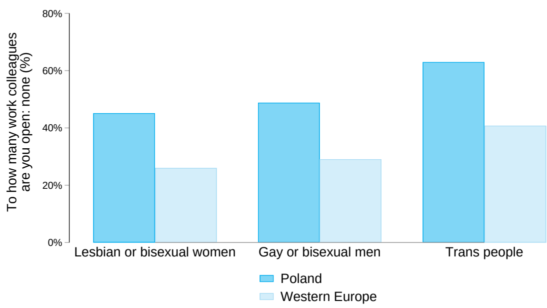
(a) Avoid holding hands in public with same-sex partner



(b) Being open to noone: school / university



(c) Being open to noone: workplace



(d) Being open to noone: family

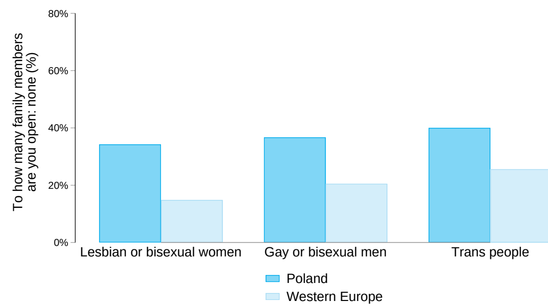


Figure 2: Concealment of the LGBTQ identity in Poland and Western Europe (2019)

Notes: Figure 2a shows the percentage of respondents who selected the answer "Never" to the question: "Do you avoid holding hands in public with a same-sex partner for fear of being assaulted, threatened or harassed?". Figure 2b shows the percentage of respondents who selected the answer "None" to the question: "To how many schoolmates / university co-students are you open about being LGBTI?". Figure 2c shows the percentage of respondents who selected the answer "None" to the question: "To how many work colleagues are you open about being LGBTI?". Figure 2d shows the percentage of respondents who selected the answer "None" to the question: "To how many family members (other than your partner(s)) are you open about being LGBTI?". Western Europe bars (light blue) show the average values for five countries: Austria, Belgium, France, Germany, and the Netherlands.
Data: 2019 Survey on LGBTI people in the EU and North Macedonia and Serbia, European Union Agency for Fundamental Rights.

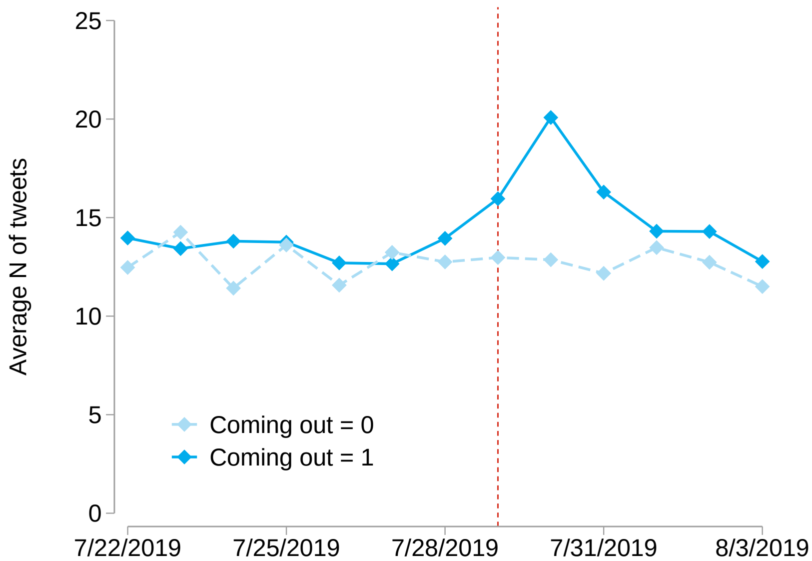


Figure 3: Twitter activity by coming out decision

Notes: Figure shows the daily average number of tweets for two groups of users: those who joined the first wave of the campaign and those who did not join the campaign in the period from July 22, 2019 to August 4, 2019.

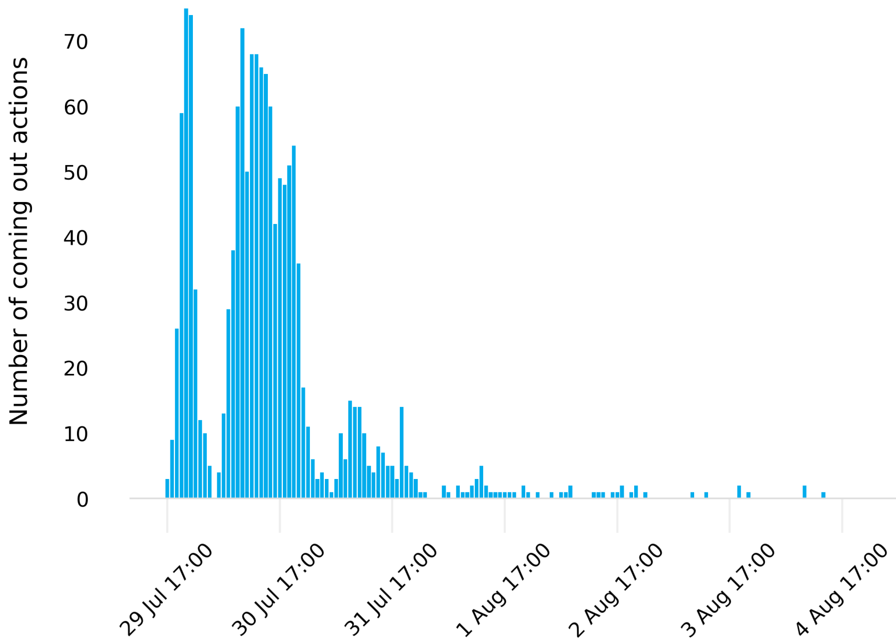


Figure 4: The development of the first wave of the LGBTQ coming out campaign

Notes: Figure displays the number of all recorded LGBTQ coming out tweets per hour from July 29 17:00:00 to August 5 00:00:00. Figure B.1 shows the number of tweets with the hashtag #IamLGBT over time for all types of users.

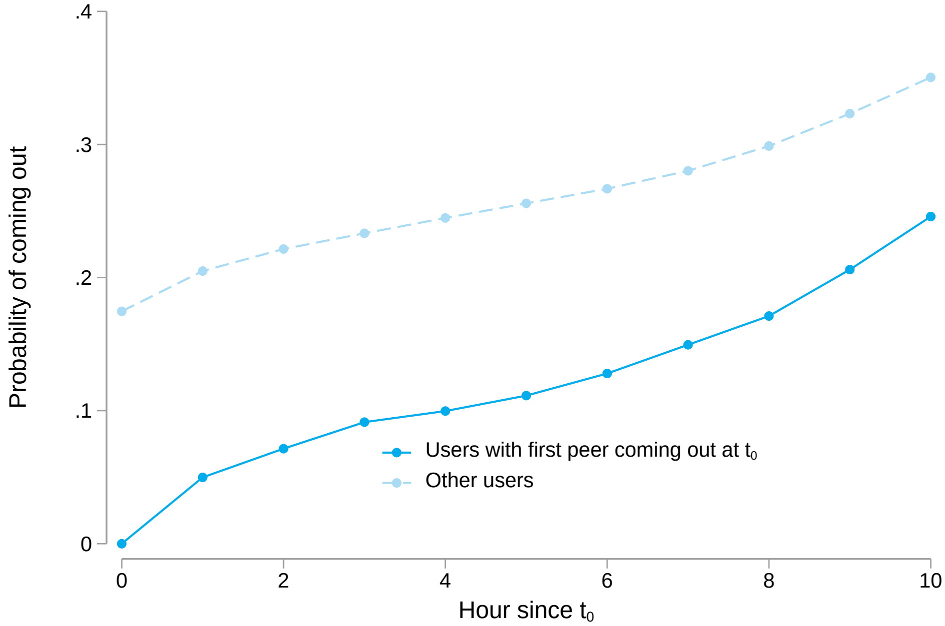


Figure 5: Visual representation of peer effects

Notes: Figure plots the average probability of coming out for users with no past exposure to peers coming out (solid line), where zero is the hour of the first peer coming out in the user's network. This is limited to users who observed the first peer coming out when they were not yet out. For comparison, we also plot the hourly probability of coming out of other users (dashed line) in the same hour.

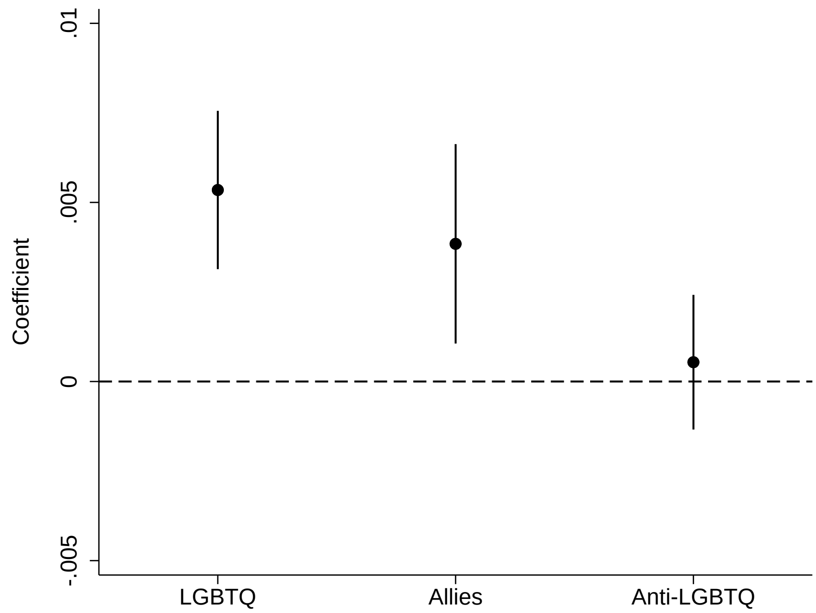


Figure 6: Peer effects and coming out: other user types

Notes: Figure shows coefficients from an OLS estimation of the effects of peers' participation in the Twitter campaign (posting a tweet with the hashtag #IamLGBT) on the hourly probability of coming out. The "LGBTQ" coefficient measures the effect of the exposure to peers coming out as LGBTQ. The "Allies" coefficient measures the effect of the exposure to peer posts by straight allies. The "Anti-LGBTQ" coefficient measures the effect of the exposure to peer posts by anti-LGBTQ users. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. In the regression, we control for gender, pre-campaign measures of Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

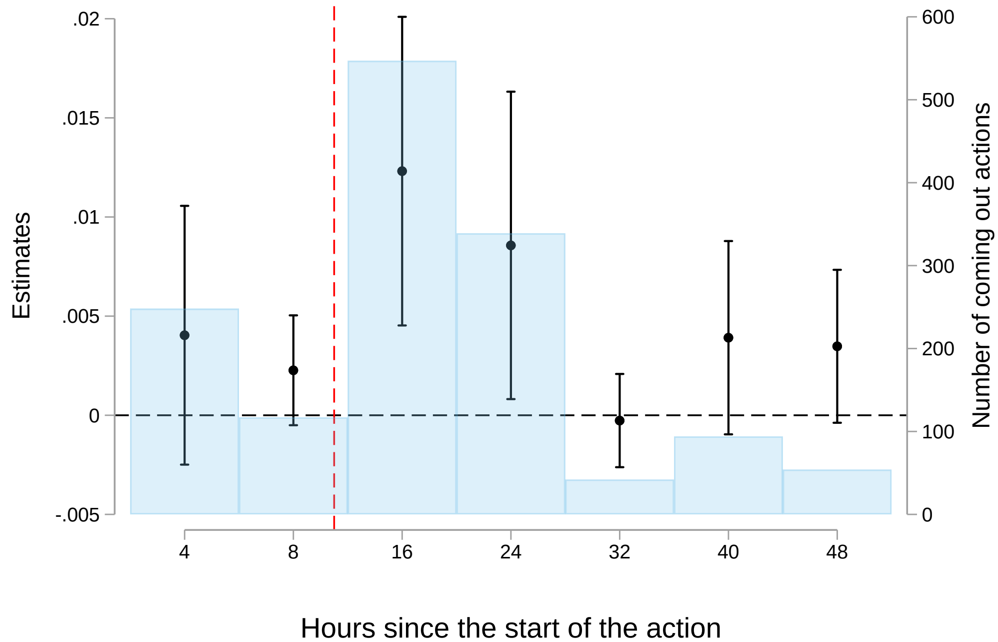
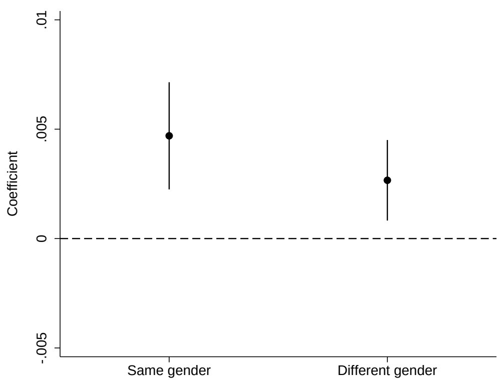


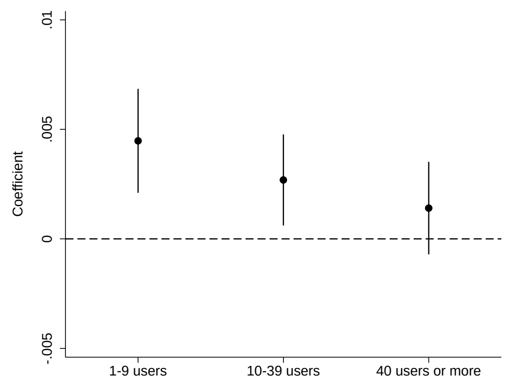
Figure 7: Peer effects and coming out over time

Notes: Figure shows coefficients from an OLS estimation of the effects of peers' participation in the Twitter campaign (posting a tweet with the hashtag #IamLGBT) on the hourly probability of coming out. We allow the effects to vary over eight-hour intervals (12AM to 7AM, 8AM to 3PM, 4PM to 11PM). The first interval includes only four hours, as our analysis starts at 8PM on July 29th. The last interval lasts seven hours as our analysis ends at 10PM on July 31st. We control for gender, pre-campaign measures of Twitter activity, and network variables. Blue bars represent the number of all recorded LGBTQ coming out tweets in the given interval. 95% confidence intervals are constructed based on standard errors clustered at the user level. The dashed vertical line denotes the time when the major Polish newspaper posted a tweet about the campaign.

(a) Peers' gender



(b) Peers' outdegree centrality



(c) Mutual vs. non-mutual relationships

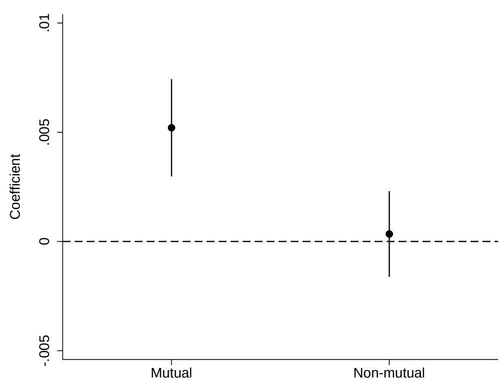


Figure 8: Heterogeneity of peer effects: peer proximity

Notes: Graphs show the OLS estimates of a duration model: point estimates and their 95% confidence intervals. Figure 8a, shows the effects of peers' coming out depending on the peers' gender. Figure 8b shows the peer effects depending on the peers' outdegree centrality: the number of users who have a given peer in their network. In Figure 8c, the 'Mutual' coefficient shows the effects of the identity disclosure of peers coming out who had a mutual relationship with the user (there was at least one reply tweet from the peer to the user), and the 'Non-mutual' coefficient shows the effects of the identity disclosure of peers who had never posted a reply tweet to the user. In all regressions, we control for gender, pre-campaign measures of Twitter activity, and network variables. 95% confidence intervals are based on standard errors clustered at the user level.

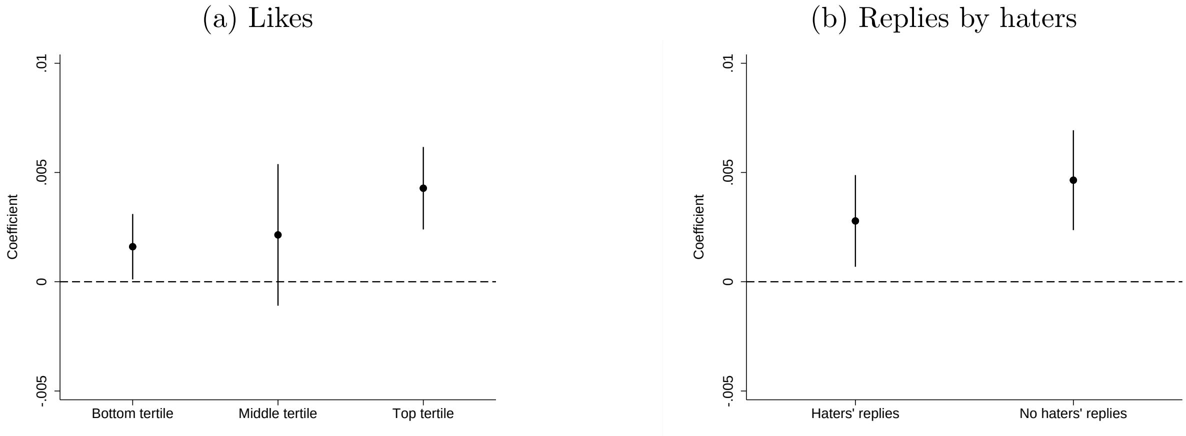


Figure 9: Peer effects and coming out: reactions to coming out tweets

Notes: Figure 9a shows coefficients from an OLS estimation of the effects of peers' LGBTQ coming out actions depending on the number of likes received by the peers' coming out tweets. The bounds of the categories represent tertiles of the distribution of the coming out tweets' likes. The bottom tertile coefficient shows the effects of peers coming out as LGBTQ if their coming out tweets received from zero to four likes. The middle tertile coefficient shows the effects of peers coming out as LGBTQ if their coming out tweets received from five to fifteen likes. The top tertile coefficient shows the effects of peers coming out as LGBTQ if their coming out tweets received at least sixteen likes. Figure 9b shows coefficients from an OLS estimation of the effects of peers' LGBTQ coming out actions for peers who were replied by anti-LGBTQ haters during the campaign, and for peers who did not experience any replies from anti-LGBTQ haters during the campaign. In both regressions, we control for gender, pre-campaign measures of Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

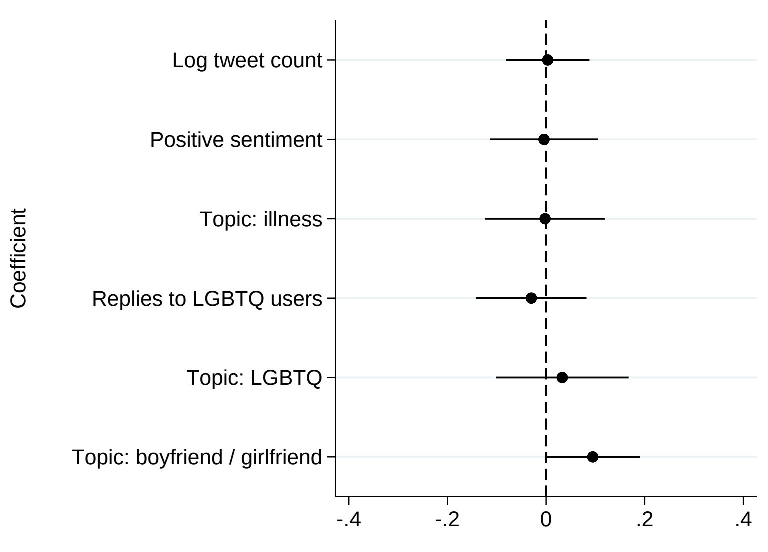


Figure 10: Peer effects and post-campaign activity

Notes: Figure shows the reduced-form DiD estimates of the effects of the exposure to peers coming out as LGBTQ on six outcomes. The treatment variable equals zero for users with exposure below the mean, and one for those with exposure equal or greater than mean. The weekly panel consists of eight weeks surrounding the campaign and two groups of users: those participated in the first wave but came out after the 19th hour of the campaign and those who did not participate in the first wave and came out during the second wave. We control for user and week fixed effects. 95% confidence intervals are constructed based on standard errors clustered at the user level. See Figure B.12 for event study plots.

Tables

Table 1: Descriptive statistics by coming out wave

Variable	1st wave	2nd wave	Obs	1st - 2nd std diff	p-val
Network: LGBTQ coming out	0.084	0.052	1,412	0.28	0.00
Network: allies	0.011	0.007	1,412	0.16	0.00
Network: anti-LGBTQ	0.007	0.004	1,412	0.12	0.04
Gender: female	0.740	0.817	1,412	-0.18	0.00
Gender: male	0.230	0.165	1,412	0.15	0.01
Gender: transgender / non-binary	0.030	0.017	1,412	0.07	0.21
log Tweets count	6.740	6.676	1,412	0.04	0.49
Average tweet length	11.241	9.314	1,412	0.41	0.00
Replies (% of all tweets)	0.451	0.453	1,412	-0.01	0.82
Hashtag use	0.062	0.056	1,412	0.07	0.25
Emoji use	0.039	0.047	1,412	-0.20	0.00
LGBTQ-related tweets	0.016	0.010	1,412	0.19	0.00
Emotional words share	0.088	0.093	1,412	-0.15	0.02
Positive tweet sentiment	0.630	0.629	1,412	0.01	0.91
log Network size	4.639	4.597	1,412	0.03	0.58
Network: media	0.014	0.005	1,412	0.22	0.00
Network: politics	0.016	0.008	1,412	0.16	0.01
Network: LGBTQ activists	0.005	0.002	1,412	0.14	0.02

Notes: This table presents the comparison of the average values of variables used in the analysis for two groups: participants in the first wave, and participants in the second wave (who did not come out during the first wave). The first column shows the averages of the variables for users who came out during the first wave of the campaign. The second column shows the averages of the variables for users who came out during the second wave of the campaign. The third column shows the number of users in the sample. The fourth column shows the standardized difference in the variables between the two groups. The fifth column shows the p-value of a t-test of the equality of averages in the two groups. The sources and a description of the variables can be found in Tables C.1-C.2. Table B.1 presents additional statistics for the whole sample of LGBTQ users. Table B.2 shows the comparison for additional variables that describe tweet emotions, tweet topic, and grammatical gender and person use. Table B.3 shows the comparison of the variables for different types of users.

Table 2: Peer effects and coming out

	(1)	(2)	(3)	(4)	(5)
Network: LGBTQ coming out	0.005***	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
Adj. R-Squared	0.00	0.02	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03	0.03	0.03
Number of clusters	1412	1412	1412	1412	1412
Observations	36128	36128	36128	36128	36128

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. We control for gender (female, male, transgender / non-binary), pre-campaign measures of Twitter activity (log tweets count, average tweet length, hashtag use, emoji use, share of LGBTQ-related tweets, replies as percentage of all tweets, emotional words use, positive tweet sentiment), and network characteristics (log network size, replies to media, politics, LGBTQ activist accounts, the exposure to peer posts by straight allies, and anti-LGBTQ users in the network). Full results are presented in Table B.5. We use standard errors clustered at the user level.
* $p < .10$; ** $p < .05$; *** $p < .01$

Table 3: Peer effects and coming out: controlling for lagged exposure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Network: LGBTQ coming out	0.005***	0.016**	0.011**	0.007**	0.005**	0.005**	0.016*
	(0.001)	(0.008)	(0.005)	(0.003)	(0.003)	(0.002)	(0.008)
Network: LGBTQ coming out (t-2)		-0.011					-0.009
		(0.008)					(0.010)
Network: LGBTQ coming out (t-3)			-0.006				-0.007
			(0.005)				(0.006)
Network: LGBTQ coming out (t-4)				-0.001			0.002
				(0.003)			(0.003)
Network: LGBTQ coming out (t-5)					0.000		0.001
					(0.002)		(0.002)
Network: LGBTQ coming out (t-6)						0.001	0.002
						(0.002)	(0.002)
Hour FE	yes	yes	yes	yes	yes	yes	yes
Gender	yes	yes	yes	yes	yes	yes	yes
Twitter activity	yes	yes	yes	yes	yes	yes	yes
Network	yes	yes	yes	yes	yes	yes	yes
Adj. R-Squared	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Number of clusters	1220	1220	1220	1220	1220	1220	1220
Observations	32073	32073	32073	32073	32073	32073	32073

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. Depending on the specification, we control for more distant lags of the network variable. In all regressions, we control for gender (female, male, transgender / non-binary), pre-campaign measures of Twitter activity (log tweets count, average tweet length, hashtag use, emoji use, share of LGBT-related tweets, replies as percentage of all tweets, emotional words use, positive tweet sentiment), and network characteristics (log network size, replies to media, politics, LGBTQ activist accounts, the exposure to peer posts by straight allies and anti-LGBTQ users in the network). Our sample includes observations from the seventh to the 54th hour of the first wave of the campaign in order to be able to estimate more distant lags. We use standard errors clustered at the user level.
* $p < .10$; ** $p < .05$; *** $p < .01$

Table 4: Peer effects by topic of the tweet

	(1) LGBTQ coming out OLS	(2) LGBTQ-related tweets Panel FE	(3) Non-LGBTQ-related tweets OLS	(4) Non-LGBTQ-related tweets Panel FE
Network: LGBTQ coming out	0.005***	0.009***	-0.004*	-0.000
	(0.001)	(0.003)	(0.002)	(0.003)
Mean of outcome	0.03	0.03	0.11	0.11
Number of clusters	1412	1412	1412	1412
Observations	36128	36128	36128	36128

Notes: Column 1 shows the OLS estimates of a duration model of the effects of the exposure to peers coming out as LGBTQ on the hourly probability of coming out (baseline estimates). Column 2 shows the panel fixed effects estimates of the effects of the exposure to peers coming out as LGBTQ on the probability of posting tweets with LGBTQ-related words. Column 3 shows the OLS estimates of the effects of the exposure to peers coming out as LGBTQ on the probability of posting tweets without LGBTQ-related words. Column 4 shows the panel fixed effects estimates of the effects of the exposure to peers coming out as LGBTQ on the probability of posting tweets without LGBTQ-related words. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. In all regressions, we control for the exposure to peer posts by straight allies and anti-LGBTQ users in the network. In columns 1 and 2, we additionally control for the baseline set of time-invariant user characteristics (see Table 2 for the list of control variables). In columns 2 and 4, we instead control for user fixed effects. We use standard errors clustered at the user level.
* $p < .10$; ** $p < .05$; *** $p < .01$

Table 5: Peer effects and and probability of coming out: IV estimates

Panel A: Second stage

	LGBTQ coming out				
	(1)	(2)	(3)	(4)	(5)
Network: LGBTQ coming out	0.009***	0.009***	0.009***	0.010***	0.011***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
F-statistic	16.16	16.14	15.71	14.83	14.28
Number of clusters	1408	1408	1408	1408	1408
Observations	36048	36048	36048	36048	36048

Panel B: First stage

	Network: LGBTQ coming out				
	(1)	(2)	(3)	(4)	(5)
Peers of peers: LGBTQ coming out	0.344***	0.344***	0.339***	0.323***	0.302***
	(0.086)	(0.086)	(0.086)	(0.084)	(0.080)
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
F-statistic	16.16	16.14	15.71	14.83	14.28
Number of clusters	1408	1408	1408	1408	1408
Observations	36048	36048	36048	36048	36048

Notes: Panel A shows the estimates of the two-stage least-squares estimation of a duration model of the hourly probability of coming out. Panel B shows the results of the first-stage regressions. The exposure to peers coming out as LGBTQ was instrumented by the exposure of peers to LGBTQ coming out actions of their peers who were not user's peers themselves. For each hour, network variables and the instrument are standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. All regressors are based on Twitter activity in the period between January 1, 2019 and July 28, 2019 (one day before the start of the campaign). We use standard errors clustered at the user level.
* $p < .10$; ** $p < .05$; *** $p < .01$

Table 6: Peer effects and probability of participation in the campaign

	(1) LGBTQ	(2) Allies	(3) Anti-LGBTQ
Network: LGBTQ coming out	0.005***	0.002	-0.000
	(0.001)	(0.002)	(0.001)
Network: allies	0.004***	0.008***	0.004**
	(0.001)	(0.002)	(0.002)
Network: anti-LGBTQ	0.001	0.003	0.002
	(0.001)	(0.002)	(0.002)
Hour FE	yes	yes	yes
Gender	yes	yes	yes
Twitter activity	yes	yes	yes
Network	yes	yes	yes
Adj. R-Squared	0.02	0.03	0.03
Mean of outcome	0.03	0.03	0.03
Number of clusters	1430	431	552
Observations	36721	11072	13560

Notes: Table shows the OLS estimates of a duration model of the probability of joining the first wave of the #IamLGBT campaign. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. 'Network: allies' measures the fraction of the network who joined as straight allies before a given hour. 'Network: anti-LGBTQ' measures the fraction of the network who joined as anti-LGBTQ users before a given hour. In column 1, we show the results for the probability of joining action as a LGBTQ user (baseline results). In column 2, we show the results for the probability of joining the campaign as a straight ally. In column 3, we show the results for the probability of joining the campaign as an anti-LGBTQ users. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. We control for gender, pre-campaign measures of Twitter activity, and network characteristics. We use standard errors clustered at the user level.
* $p < .10$; ** $p < .05$; *** $p < .01$

Appendix A Theoretical model

Consider a society consisting of N individuals with a concealable LGBTQ identity. Each individual is connected to a certain number of individuals (peers) and can observe what decisions her peers are making. The connections are described by a connected graph $\mathbb{G}$, where $\mathbb{G}_{i,j}$ is a weight of a directed edge from individual i to peer j, capturing the relative influence of peer j in individual i's decision-making. These weights may be asymmetric, meaning some peers have a stronger influence than others. For simiplicity of notation we also introduce $\mathbb{I}$ - a set of neighbors of player i.

In each period t, all concealing individuals decide whether to conceal or to irreversibly disclose their stigmatized identities (i.e., to publicly come out). Two factors affect the decision to come out. First, individual-level characteristics determine net personal benefits of the coming out, $P_i \sim \mathcal{N}(\bar{P}, \epsilon)$, which vary across individuals. Individuals have perfect information about their own net personal benefits. High concealment costs (such as anxiety or the inability to be oneself) increase P_i, while factors like shyness or a strong attachment to a homophobic family decrease it. Second, net social cost (D) driven by discrimination reduces the likelihood of coming out. Negative social reactions (e.g., employer discrimination, bullying, stigmatization) increase net social cost, and positive social reactions (e.g., financial or non-financial support from other LGBTQ individuals) reduce net social cost. These net social costs are identical for all individuals. For simplicity, the actual level of net social cost can take only two values (high and low). Importantly, the level of the social cost is revealed only if the individual discloses their identity. Hence, concealing individuals do not know the actual level of social cost they will experience if they come out. Instead, they form their beliefs about net social cost levels (i.e., the probability of net social cost levels being high), $D(y_{i,t})$, based on coming outs they have observed in their network ($y_{i,t}$), as we explain below. Individuals maximize their utility described by the following function:

$$\max_{R_{i,t}} U(R_{i,t}) = R_{i,t} \times P_i - R_{i,t} \times D(y_{i,t}), \tag{5}$$

where $R_{i,t}$ equals zero for a concealing individual, and one for an individual who disclosed his identity. An individual will decide to come out only if the individual-level benefits are at least as large as the expected net social cost:

$$P_i - D(y_{i,t}) \geq 0. \tag{6}$$

A crucial aspect of this decision-making process is how individuals update their beliefs about net social costs. When individuals observe others coming out, two opposing forces influence their perception of risk. First, individuals interpret others' coming out as a signal about the level of net social cost. They reason that those who came out likely had information or experiences suggesting discrimination is lower than previously feared. Second, observing immediate positive reactions to peer disclosures may increase perceived probability of the low levels of net social cost due to public support. As a result, individuals update their beliefs using a Bayesian learning process.

In each period t, each concealing individual observes the coming out actions of her peers, denoted as R_{jt} (equal to 1 if a peer j has come out and 0 otherwise). The measure of the weighted exposure to peers coming out as LGBTQ is given by:

$$y_{i,t} = \frac{\sum_{j \in \mathbb{I}} \mathbb{G}_{ji} R_{jt}}{\sum_{j \in \mathbb{I}} \mathbb{G}_{ji}}, \tag{7}$$

It is the share of peers participating in the campaign weighted by the strength of the user i's connection to peer j. Individuals use the observations to update their beliefs about net social costs in the society, $D(y_{i,t})$.

Since concealing individuals do not experience the actual level of net social cost themselves, they rely on observing peer behavior to infer it. For simplicity, the actual level of net social cost D in society can take only two values, $D \in \{D_L, D_H\}$, where $D_L < D_H$. The individual i's prior belief about the level of net social cost is given by the log-likelihood ratio:

$$\lambda_0 = \ln\left(\frac{Pr(D = D_L)}{Pr(D = D_H)}\right). \tag{8}$$

After observing the state of the world through $y_{i,t}$, they update their belief using Bayes' rule. The learning mechanism is similar to one proposed by Acemoglu et al. (2011), where individuals notice the share of others taking the risky action, and assess the probability that they will experience a high level of net social cost.

Individuals estimate the probability of coming out under each net social cost level:

$$Pr_x := Pr(R_{i,t}|D = D_x) = 1 - \Phi\left(\frac{D_x - \bar{P}}{\sqrt{\epsilon}}\right).$$

Since individuals do not observe others' personal benefits P_i, their best estimate is based on the mean private benefit of coming out.

Given the Pr_x, each individual looks at the probability that a share of the society $y_{i,t}$ would come out under each net social cost level:

$$Pr(y_{i,t}|D = D_x) = \binom{n}{y_{i,t}n} \Pr(R_{i,t}|D = D_x)^{ny_{i,t}}(1 - Pr(R_{i,t}|D = D_x))^{n(1-y_{i,t})}.$$

Using those, individuals obtain the new log-likelihood ratio given by:

$$\lambda_{i,t+1}(y_{i,t+1}) = \lambda_0 + \ln\left(\frac{Pr(y_{i,t}|D = D_L)}{Pr(y_{i,t}|D = D_H)}\right) \tag{9}$$

$$= \lambda_0 + \ln\left(\left(\frac{1 - \Phi(\frac{D_L - \bar{P}}{\sqrt{\epsilon}})}{1 - \Phi(\frac{D_H - \bar{P}}{\sqrt{\epsilon}})}\right)^{ny_{i,t}} \left(\frac{\Phi(\frac{D_L - \bar{P}}{\sqrt{\epsilon}})}{\Phi(\frac{D_H - \bar{P}}{\sqrt{\epsilon}})}\right)^{n(1-y_{i,t})}\right). \tag{10}$$

Individuals face uncertainty about discrimination and rely on peer observations to update their beliefs. The decision to come out is driven by personal characteristics and the perceived risk of discrimination, which evolves as individuals interpret the actions of others in their network.

PROPOSITION 1. *The log-likelihood ratio $\lambda_{i,t}$ is linearly increasing in $y_{i,t}$.*

Proof. We take the first derivative of $\lambda_{i,t}$ with respect to $y_{i,t}$ and obtain a result showing how the increase in the observed share of society coming out changes player i's perceived likelihood of discrimination being of a low type.

$$\frac{\partial \lambda_{i,t}(y_{i,t})}{\partial y_{i,t}} = n\left(ln\left(\frac{Pr_L}{Pr_H}\right) - ln\left(\frac{1-Pr_L}{1-Pr_H}\right)\right),$$

as $1 - Pr_L < 1 - Pr_H$ or $\Phi(\frac{D_L - \bar{P}}{\sqrt{\epsilon}}) < \Phi(\frac{D_H - \bar{P}}{\sqrt{\epsilon}})$ it is always greater than 0. This means that the higher the weighted share of the player's network who declares their participation in the movement, the more likely the player is to perceive low levels of discrimination in the society as a whole.

The right-hand side of the equation remains constant as neither n, Pr_L, nor Pr_H change during the game. Therefore the increase of $\lambda_{i,t}$ w.r.t. $y_{i,t}$ is linear. □

Proposition 1 reveals that observing peers coming out makes individuals more likely to believe that the true level of net social cost is low. This follows an intuition that an individual will see risky decisions made by her peers as an indication that they perceive the level of net social cost to be low. This may alter the person's own beliefs about the costs of coming out, as she becomes aware that her prior beliefs may be wrong. Importantly, it is irrelevant if the user i had prior private knowledge about the LGBTQ identity of their peers. As players use the information to gain knowledge about the level of net social cost, only the public action is relevant, not the previous private disclosure of the individual characteristics. Proposition 1 yields the first testable prediction: an increase in the exposure to peers coming out should lead to an increase in the probability of coming out.

Now we can derive a threshold of the exposure to peers coming out, y_i^*, at which an individual changes her mind about whether to come out. A strategy profile given by a vector y^* of individual minimum levels of network participation y_i^* is a pure strategy perfect Bayesian equilibrium of this game of social learning as for each $i \in \mathbb{I}$, observation $y_{i,t} > y_i^*$ maximizes the expected pay-off of individual i given the strategies of other individuals y_{-i}^*.

PROPOSITION 2. *There exists a unique pure strategy perfect Bayesian equilibrium y^* of the proposed game.*

Proof. As a first step let's introduce p_i^*, a critical probability of $D = D_L$ i.e.

$$p_i^* D_L + (1 - p_i^*) D_H = P_i \implies p_i^* = \frac{P_i - D_H}{D_L - D_H} = \frac{D_H - P_i}{D_H - D_L}. \tag{11}$$

Players will only participate in the campaign if the probability of discrimination being of the low type is at least p_i^*. From that we can conclude the y_i^*.

$$\lambda_{i,t}(y_i^*) = ln\left(\frac{p_i^*}{1 - p_i^*}\right) \iff$$

$$ln\left(\frac{Pr(D = D_L)}{Pr(D = D_H)}\right) + ln\left(\left(\frac{Pr_L}{Pr_H}\right)^{ny_i^*}\left(\frac{1 - Pr_L}{1 - Pr_H}\right)^{n(1-y_i^*)}\right) = ln\left(\frac{\frac{P_i - D_H}{D_L - DH}}{1 - \frac{P_i - D_H}{D_L - DH}}\right).$$

This condition is only satisfied if:

$$y_i^* = \frac{n \cdot ln\left(\frac{1-Pr_L}{1-Pr_H}\right) - ln\left(\frac{Pr(D=D_H)}{Pr(D=D_L)}\frac{D_H - P_i}{D_L - P_i}\right)}{n \cdot ln\left(\frac{Pr_H}{Pr_L}\frac{1-Pr_L}{1-Pr_H}\right)}, \tag{12}$$

which is the lowest possible value of participation among player i's neighbors at which she would come out as a part of the campaign. For each player, this level is unique and depends only on parameters of the model. Neither of the parameters on the right-hand side of the equation depends on the players' decision, therefore, the y_i^* above is unique for each player, and stays constant throughout the game.

A vector y^* of individual thresholds y_i^* is an equilibrium in the proposed game. It is a unique equilibrium given by the closed-form solution in equation 12. □

Proposition 2 reveals that each individual has a fixed net social cost level she is willing to accept while still choosing to come out. As individuals' beliefs about the level of net social cost in society are a function of the share of their peers who are coming out, we can find the share that would convince each individual to come out as part of the campaign. Proposition 2 gives us an individual-specific threshold of observed coming out actions, after which a given individual in the society would decide to come out himself.

The equilibrium has a closed-form solution, therefore, it has to be unique. It only depends on individual costs, the possible net social cost levels in the society, and the likelihood of them being the actual state of the world. As we obtain such a threshold for each player, we have a vector y^* that defines the actions of each individual.

In the following propositions, we study the equilibrium threshold y^* to obtain further testable predictions. First, we check how the individual-level time-invariant benefits of coming out, P_i, affect the level of exposure needed to change the coming out decision.

PROPOSITION 3. *The high individual-level time-invariant benefit of coming out, P_i, leads to the lower threshold y_i^*. Among two individuals with otherwise identical characteristics and network connections, the one with the higher private benefit of coming out P_i will have the lower threshold y_i^*, and will therefore be more likely to come out during the campaign; and will, holding everything else constant, come out more quickly than the individual with lower P_i.*

Proof. As $\lambda_{i,t}(y_i^*) = ln\left(\frac{p_i^*}{1-p_i^*}\right)$ and λ is strictly increasing we know that as p^* increases y^* increases as well. Therefore to show that y_i^* decreases, it is enough to show that the higher the P_i, the lower the p_i^*.

From equation 11 we have that $p_i^* = \frac{D_H - P_i}{D_H - D_L}$. This gives us $\frac{\partial p_i^*}{\partial P_i} = \frac{1}{D_L - D_H}$, which is always negative. Therefore, an increase in P always leads to a lower probability of low discrimination being required for a player i to participate in the campaign. Theorem 1 tells us that the perceived probability of $D = D_L$ increases linearly in $y_{i,t}$, therefore with lower p_i^*, y_i^* will be lower too. $\square$

Proposition 3 means that individuals with low personal costs of coming out or a high perceived benefit from doing so would be more likely to come out as part of the campaign. This yields another testable prediction: if peer effects are driven by the social learning mechanism, they should be smaller during the first hour of the campaign than during subsequent hours. This is because the individuals who come out at the beginning of the campaign are those who have the lowest individual-level cost or the highest benefit. For them, according to Proposition 2, peer effects are less important than they are for users with low P_i, who need to make sure that the risk of high discrimination is low.

Peer impact on the decision by a player

Given the structure of the game, players only change their strategy if the peer effect is strong enough to convince them to join the campaign, i.e. $y_i >= y_i^*$. Therefore only the neighbor who increases player i's observed measure of participation y_i above the threshold is having an impact on his decision and ex-interim utility. This is similar to the role of the median voter in games of elections. It may also influence the observed behavior of players as the peer effect's influence on a strategy is a step function.

As we analyze the equilibrium of the game, we look at the influence of each parameter on y^*. This might not impact the outcome of the game because even if y_i^* decreases, individuals may not observe enough of their peers' coming out actions to change their action. It means, however, that they are more likely to change their action and would do so more quickly if the coming out actions happened in their network.

Welfare effects

From equation 5, the utility of individuals who decide not to come out is zero, resulting in no change in their welfare. For individuals who decide to take the risky coming out action, their ex-post utility is given by $P_i - D$. Hence, the effects on their welfare are ambiguous. In the non-trivial scenario where $D_L < P_i < D_H$, the welfare effects are positive if $D = D_L$ and negative if $D = D_H$.

In the trivial case where $D_L > P_i$, the individual i would never come out, as any realization of discrimination would result in a negative utility. On the other extreme, if $D_H < P_i$, the individual i would always decide to come out.

In the proposed model, the discrimination and the idiosyncratic benefit from coming out remain constant. Therefore, potential changes in the network structure and individual decisions do not impact the ex post utility, only the ex ante expectations.

Other findings

PROPOSITION 4. *The strength of the relation between coming outs observed by an individual i and the probability of i's coming out described in the Proposition 1 is negatively correlated with the variance of P_i - ϵ.*

Proof. Let's look at a fragment of the equation for $Pr(y_{i,t}|D = D_x)$, where x is any L or H.

$$\frac{\partial\Phi(\frac{D_x-\bar{P}}{\sqrt{\epsilon}})}{\partial\epsilon} = -\frac{(D_x-\bar{P})\phi(\frac{D_x-\bar{P}}{\sqrt{\epsilon}})}{2\epsilon^{3/2}}. \tag{13}$$

It is always greater than zero for D_L and less than zero for D_H. Now, let's put it back into the $Pr(y_{i,t}|D = D_x)$.

$$\begin{aligned}\frac{\partial Pr(y_{i,t}|D = D_x)}{\partial\Phi(\frac{D_x-\bar{P}}{\sqrt{\epsilon}})} = \binom{n}{y_{i,t}n}&((1-y_{i,t})n(1-\Phi(\frac{D_x-\bar{P}}{\sqrt{\epsilon}}))^{y_{i,t}n}\Phi(\frac{D_x-\bar{P}}{\sqrt{\epsilon}})^{(1-y_{i,t})n-1}\\ &-y_{i,t}n(1-\Phi(\frac{D_x-\bar{P}}{\sqrt{\epsilon}}))^{y_{i,t}n-1}\Phi(\frac{D_x-\bar{P}}{\sqrt{\epsilon}})^{(1-y_{i,t})n} > 0\\ &\iff 1-y_{i,t} > \Phi(\frac{D_x-\bar{P}}{\sqrt{\epsilon}})^{(1-y_{i,t})n}.\end{aligned}$$

Thus, for low levels of $y_{i,t}$, with an increase of ϵ, the perceived probability of lower discrimination levels go down as well. □

Proposition 4 reveals that peer effects are stronger in societies with a low variation of individual benefits from coming out (or costs of concealing) than in societies with a high variance of individual benefits. This follows the intuition that the extent to which individuals rely on the signal from their peers depends on how similar their peers are. Individuals who are connected to a more homogeneous crowd tend to understand the signals and decisions better. This leads to a greater certainty that their peers' decisions are driven by the low level of net social cost, as opposed to peers' private benefits being very high.

COROLLARY 1. *Increasing any of the social costs an individual bears when participating in the campaign - D_L or D_H - decreases the chances of an individual to participate.*

Proof. Following the reasoning from previous propositions, we take the first derivatives of p^* with respect to both levels of discrimination individuals may face:

$$\frac{\partial p_i^*}{\partial D_H} = \frac{P_i - D_L}{(D_H - D_L)^2},$$

$$\frac{\partial p_i^*}{\partial D_L} = \frac{D_H - P_i}{(D_H - D_L)^2}.$$

Both those derivatives are positive. Therefore, higher costs of participation in the campaign cause the critical probability for an individual to join to increase. This means that as the discrimination increases, the less likely individuals are to come out. □

The corollary above and the propositions in the main part of the paper reflect the direct impact of parameters in the utility function. With a higher benefit of participating in the campaign or a higher cost of not doing so (P_i), an individual will be more likely to participate. Conversely, if the costs of participation increase (D), individuals will be less likely to join the movement. The second part of this finding is not directly testable in the data, as levels of discrimination do not change during a short campaign on social media. We may, however, estimate individuals' expected personal costs and benefits of their risky decision to come out.

In the proposed model, the decision to participate in a campaign is based solely on individuals' expected utility of doing so. After they make their choice to come out, they do not have more choices to make. However, their decision impacts other individuals' expectations in a way that is impacted by their out-degree in relation to the in-degree of their peers.

The impact of player i's coming out on j's $y_{j,t}$ is given by i's ratio in j's in-degree I_{ij}:

$$I_{ij} = \frac{\mathbb{G}_{ij}}{\sum_{k \in \mathbb{I}} \mathbb{G}_{kj}}, \tag{14}$$

where $\mathbb{G}_{ij}$ is the strength of the edge from i to j, and $\sum_{k \in \mathbb{I}} \mathbb{G}_{kj}$ is j's in-degree. A total, direct effect of i's coming out on her peers is given by:

$$I_i = \sum_{j \in \mathbb{I}} \frac{\mathbb{G}_{ij}}{\sum_{k \in \mathbb{I}} \mathbb{G}_{kj}}. \tag{15}$$

Therefore, individual i's impact on the whole campaign is given by her out-degree, through which she influences other individuals. Indirectly, individual i's coming out may set off a domino effect by causing her neighbor to come out, which would, in turn, cause individuals in her network to join the campaign and so forth. Therefore, an individual's final impact on the success of the campaign can be measured using the eigenvalue centrality, but using the intensity matrix I instead of the traditional degrees.

Appendix B Additional results

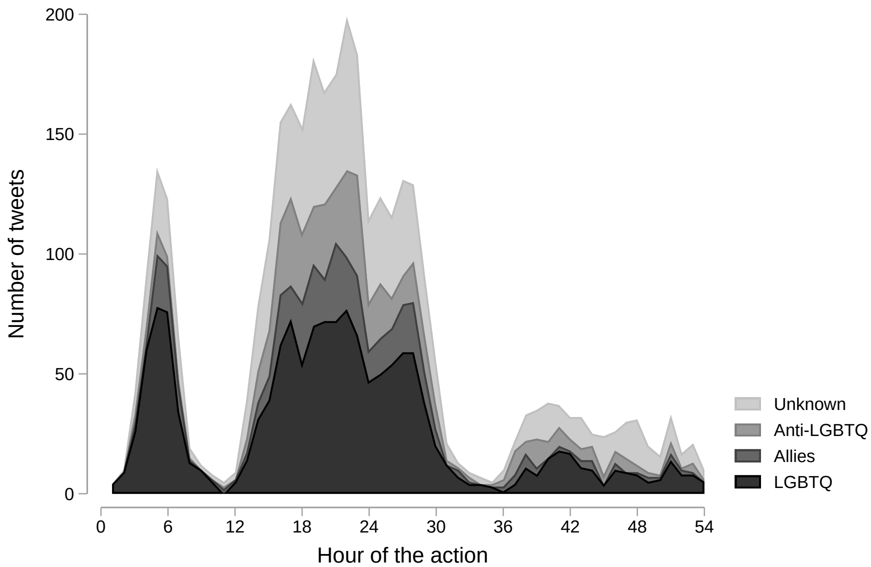


Figure B.1: The number of tweets per hour during the first wave of the campaign, by user type

Notes: Figure displays the number of tweets with the *IamLGBT* hashtag per hour during the first 54 hours of the first wave of the campaign by user type.

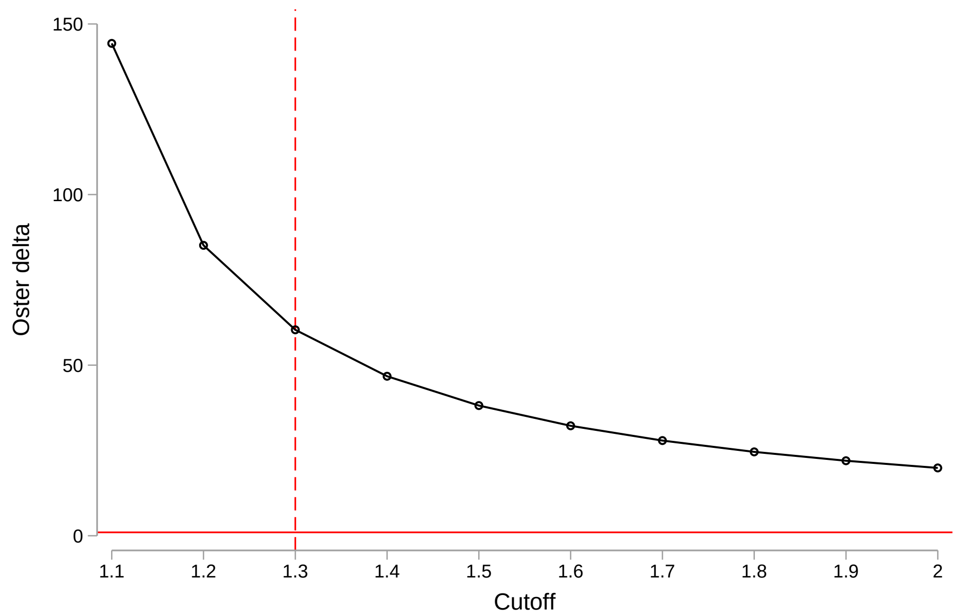


Figure B.2: Robustness of Oster's δ to alternative assumptions about R^2_{max} cutoffs.

Notes: Figure presents the value of Oster's δ as a function of R^2_{max} from $R^2_{max} = 1.1\bar{R}^2$ to $R^2_{max} = 2\bar{R}^2$, where $\bar{R}^2$ is the R-squared from the baseline regression. δ statistic indicates how much more important unobservables need to be compared to observables to fully explain our results by omitted variable bias. The red vertical line denotes the cutoff suggested by Oster (2019). The red horizontal line denotes $\delta = 1$, which indicates equal selection on observables and unobservables, a threshold suggested by Oster (2019).

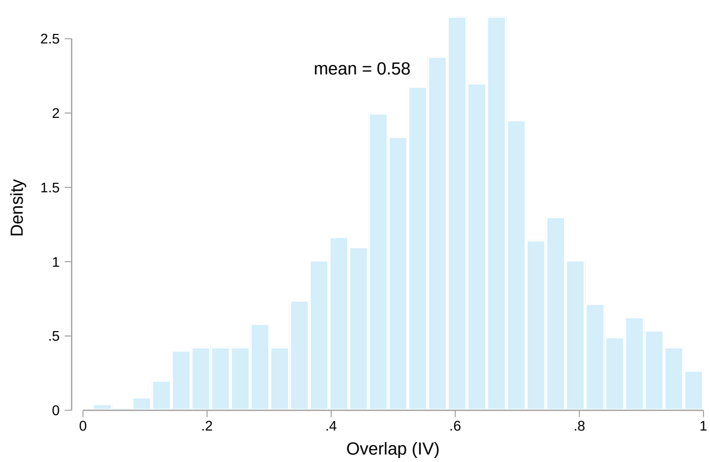


Figure B.3: Peer of peer data coverage

Notes: Figure shows the distribution and the mean of the network coverage of the data used to generate peers of peers instruments. The coverage measures the share of users' peers for whom we downloaded Twitter data (participants in two waves of the campaign and the most influential peers).

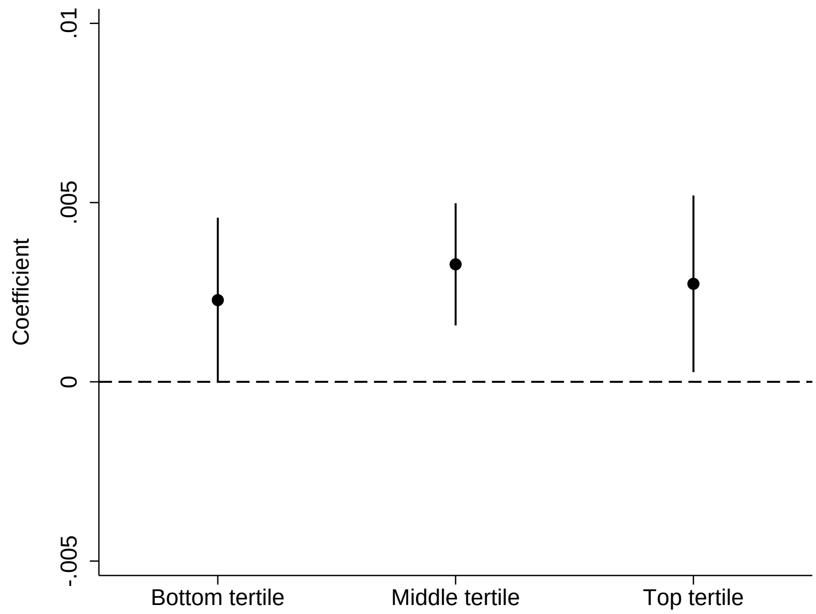


Figure B.4: Peer effects and coming out: pre-campaign popularity of peers

Notes: Figure shows coefficients from an OLS estimation of the effects of peers' LGBTQ coming out actions depending on the average number of likes received by the peers' tweets in the pre-campaign period. The bounds of the categories represent tertiles of the distribution. The 'Bottom tertile' coefficient shows the effects of LGBTQ coming out actions of peers from the bottom tertile (less than approximately 0.87 likes per tweet). The 'Middle tertile' coefficient shows the effects of LGBTQ coming out actions of peers from the second tertile (from approximately 0.87 to 2.09 likes per tweet). The 'Top tertile' coefficient shows the effects of LGBTQ coming out actions of peers from the top tertile (more than approximately 2.09 likes per tweet). In the regression, we control for gender, pre-campaign measures of Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

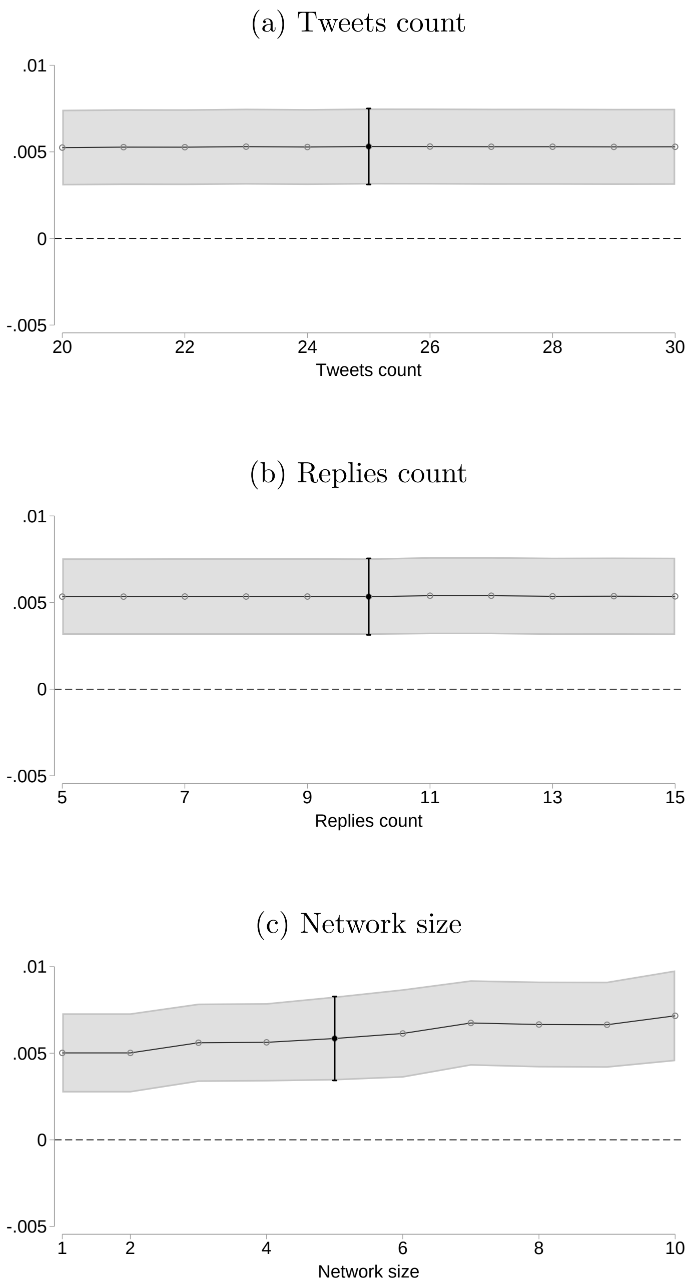


Figure B.5: Sensitivity of the results to changes in sample restrictions

Notes: Figure shows coefficients from an OLS estimation of the effects of peers' LGBTQ coming out actions in the Twitter campaign on the probability of coming out for varying sample restrictions. In Figure B.5a, we show the results with varying cutoffs of the minimum number of tweets during the pre-treatment period (January 1, 2019 - July 28, 2019). In Figure B.5b, we show the results with varying cutoffs of the minimum number of replies during the pre-treatment period. In Figure B.5c, we show the results with varying cutoffs of the minimum network size (the number of unique users to whom the user replied at least once) during the pre-treatment period. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. In all regressions, we control for gender, pre-campaign measures of Twitter activity, and network characteristics. 95% confidence intervals are constructed based on standard errors clustered at the user level.

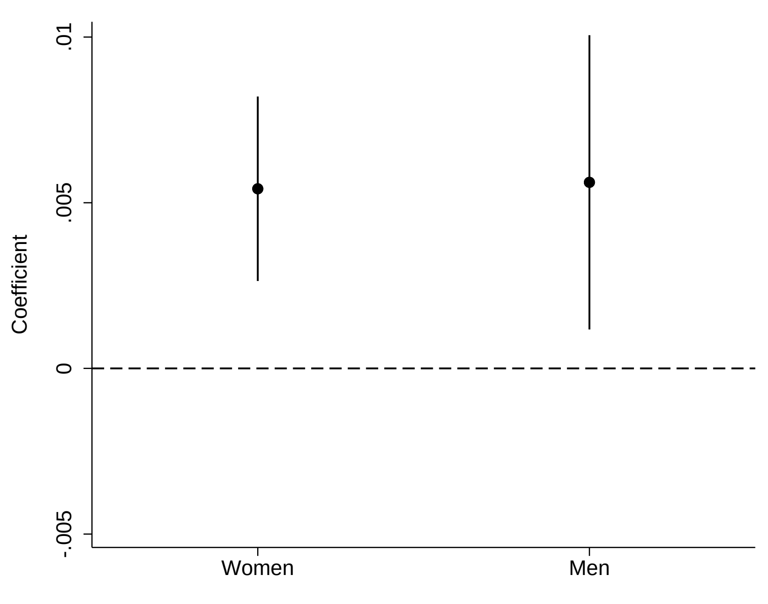


Figure B.6: Heterogeneity of the effect: gender

Notes: Figure shows coefficients from an OLS estimation of the effects of peers' LGBTQ coming out actions in the Twitter campaign on the time of coming out for women and men. The low number of observations does not allow us to study the effects separately for transgender / non-binary users (there are only four users who did not come out in the first wave in the sample). For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. In all regressions, we control for gender, pre-campaign measures of Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

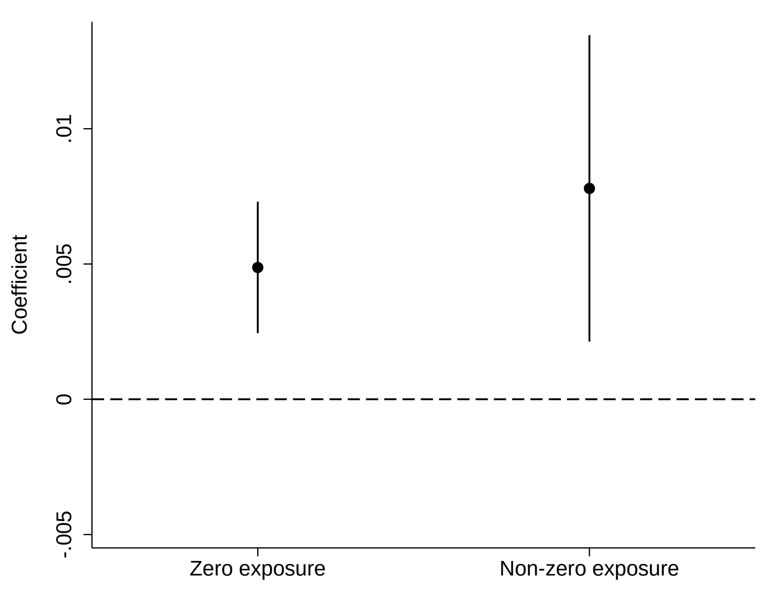


Figure B.7: Heterogeneity of the effect: exposure to LGBT activists in the network

Notes: Figure shows coefficients from an OLS estimation of the effects of peers' LGBTQ coming out actions in the Twitter campaign on the time of coming out for users with zero and non-zero pre-campaign exposure to LGBTQ activists in their network. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. In all regressions, we control for gender, pre-campaign measures of Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

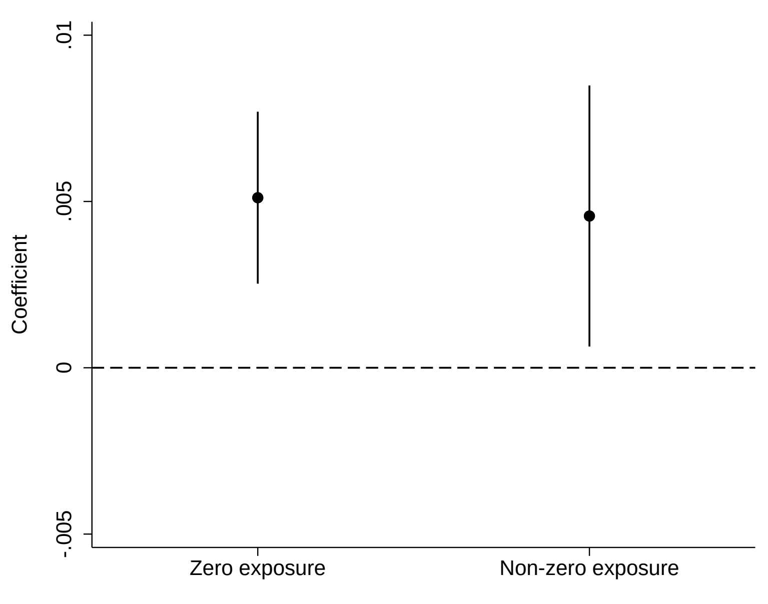


Figure B.8: Heterogeneity of the effect: exposure to media accounts

Notes: Figure shows coefficients from an OLS estimation of the effects of peers' LGBTQ coming out actions in the Twitter campaign on the time of coming out for users with zero and non-zero pre-campaign exposure to media accounts in their network. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. In all regressions, we control for gender, pre-campaign measures of Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

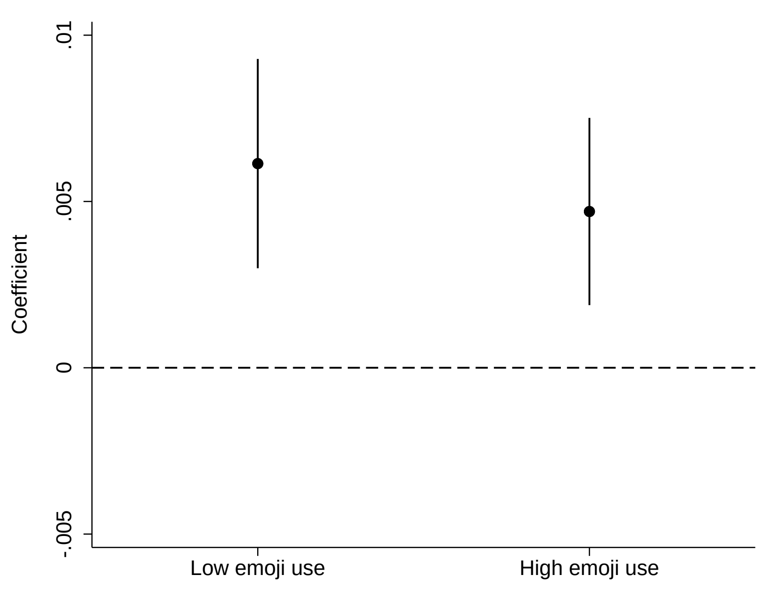


Figure B.9: Heterogeneity of the effect: emoji use

Notes: Figure shows coefficients from an OLS estimation of the effects of peers' LGBTQ coming out actions in the Twitter campaign on the time of coming out for users with emoji use below or equal to median (Low emoji use), and users with emoji use above median (High emoji use). For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. In all regressions, we control for gender, pre-campaign measures of Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

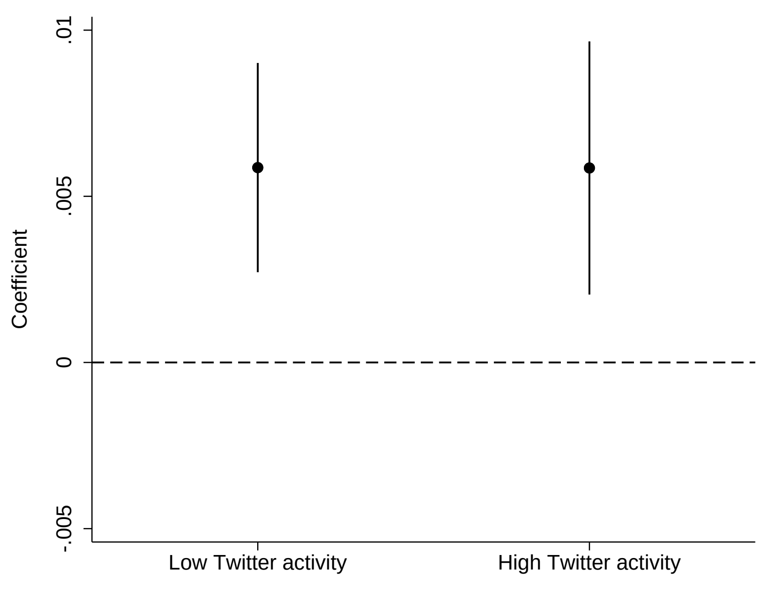


Figure B.10: Heterogeneity of the effect: Twitter activity level

Notes: Figure shows coefficients from an OLS estimation of the effects of peers' LGBTQ coming out actions in the Twitter campaign on the time of coming out for users whose number of posted tweets is below or equal to median (Low Twitter activity), and users whose number of posted tweets is above median (Twitter activity). The number of posted tweets is the number of tweets a user posted between January, 1 2019 and July, 28 2019. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. In all regressions, we control for gender, pre-campaign measures of Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

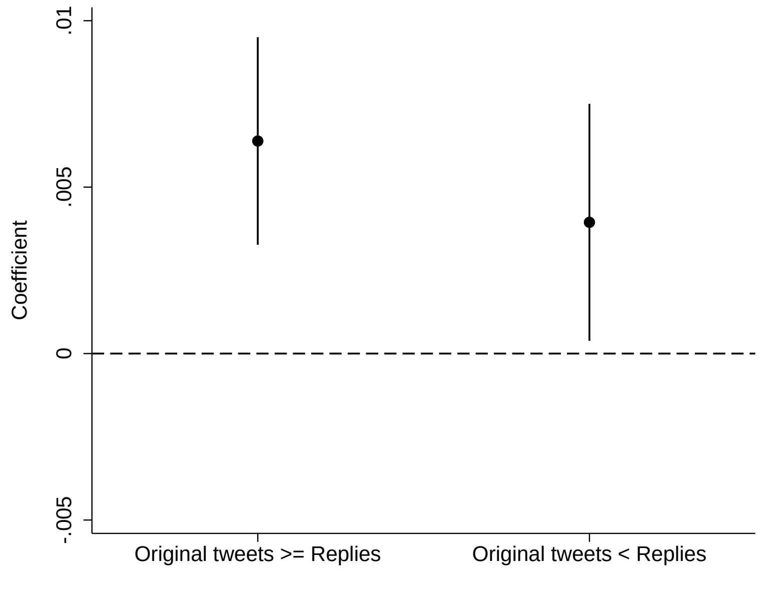


Figure B.11: Heterogeneity of the effect: reply behavior

Notes: Figure shows coefficients from an OLS estimation of the effects of peers' LGBTQ coming out actions in the Twitter campaign on the time of coming out for two groups: users whose reply share of tweets is less than or equal to 0.5, and users whose reply share of tweets is greater than 0.5. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. In all regressions, we control for gender, pre-campaign measures of Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

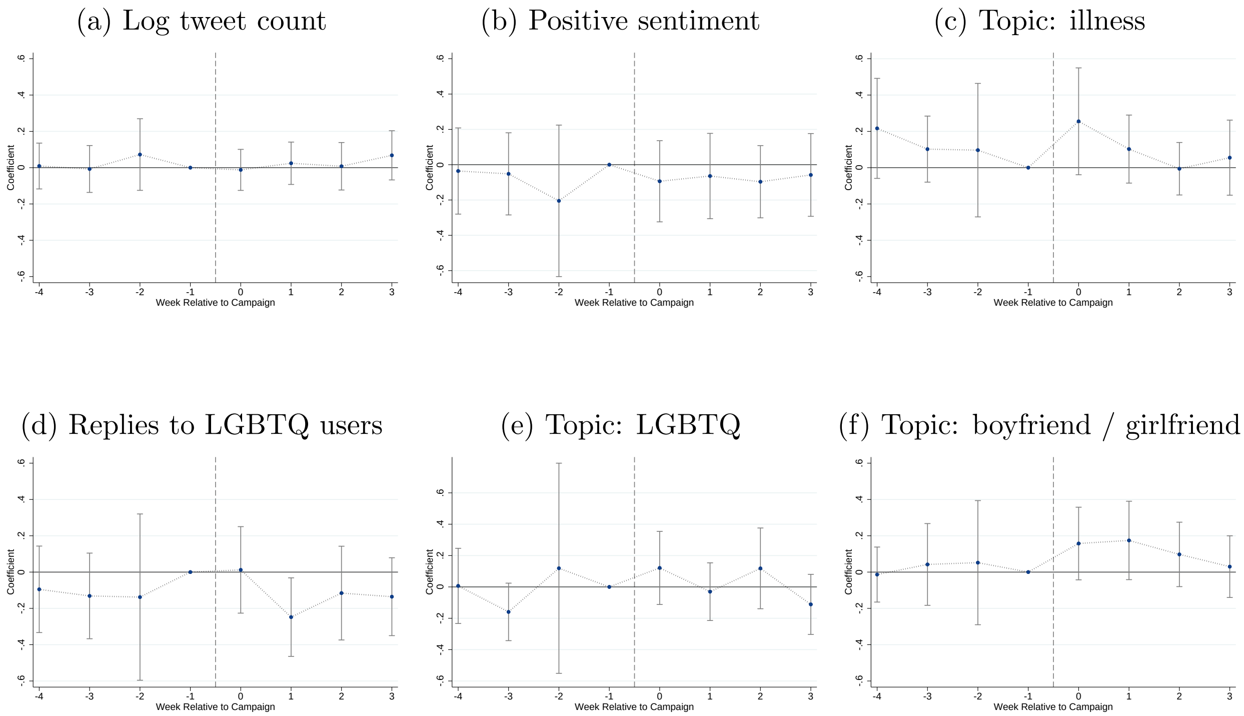


Figure B.12: Peer effects and post-campaign activity: event study approach

Notes: Figure shows the leads and lags of the effects of exposure to peers coming out as LGBTQ on six outcomes. The treatment variable equals zero for users with exposure below the mean, and one for those with exposure equal or greater than mean. The weekly panel consists of eight weeks surrounding the campaign and two groups of users: those participated in the first wave but came out after the 19th hour of the campaign and those who did not participate in the first wave and came out during the second wave. We control for user and week fixed effects. 95% confidence intervals are constructed based on standard errors clustered at the user level.

Table B.1: Descriptive statistics

	Obs.	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Network: LGBTQ coming out	36128	0.00	1.00	-0.68	22.94
Gender: woman	36128	0.77	0.42	0.00	1.00
Gender: man	36128	0.20	0.40	0.00	1.00
Gender: transgender / non-binary	36128	0.03	0.17	0.00	1.00
log Tweets count	36128	6.62	1.52	3.18	9.95
Average tweet length	36128	10.39	4.44	2.03	37.04
Replies (% of tweets)	36128	0.46	0.20	0.02	1.00
Hashtag use	36128	0.06	0.09	0.00	0.96
Emoji use	36128	0.04	0.05	0.00	0.67
LGBTQ-related words use	36128	0.01	0.03	0.00	0.59
Emotional words share	36128	0.09	0.04	0.02	0.34
Positive tweet sentiment	36128	0.63	0.12	0.24	1.00
Network: allies	36128	-0.00	1.00	-0.61	31.02
Network: anti-LGBTQ	36128	-0.00	1.00	-0.34	34.80
log Network size	36128	4.56	1.24	1.61	7.44
Network: media	36128	0.01	0.04	0.00	0.40
Network: politics	36128	0.01	0.04	0.00	0.49
Network: LGBTQ activists	36128	0.00	0.01	0.00	0.19

Notes: This table presents the following statistics for each variable: Number of Observations, Average Value, Standard Deviation, Maximum and Minimum Value. The sources and a description of the variables can be found in Tables C.1-C.2. Network variables are standardized with a mean of zero and a standard deviation of one.

Table B.2: Descriptive statistics by coming out wave, including additional variables

Variable	1st wave	2nd wave	Obs	1st - 2nd	
				std diff	p-val
Network: LGBTQ coming out	0.084	0.052	1,412	0.28	0.00
Network: allies	0.011	0.007	1,412	0.16	0.00
Network: anti-LGBTQ	0.007	0.004	1,412	0.12	0.04
Gender: female	0.740	0.817	1,412	-0.18	0.00
Gender: male	0.230	0.165	1,412	0.15	0.01
Gender: transgender / non-binary	0.030	0.017	1,412	0.07	0.21
log Tweets count	6.740	6.676	1,412	0.04	0.49
Average tweet length	11.241	9.314	1,412	0.41	0.00
Replies (% of all tweets)	0.451	0.453	1,412	-0.01	0.82
Hashtag use	0.062	0.056	1,412	0.07	0.25
Emoji use	0.039	0.047	1,412	-0.20	0.00
LGBTQ-related tweets	0.016	0.010	1,412	0.19	0.00
Emotional words share	0.088	0.093	1,412	-0.15	0.02
Positive tweet sentiment	0.630	0.629	1,412	0.01	0.91
Tweet sentiment: negative	0.370	0.371	1,412	-0.01	0.91
Tweet emotions: joy	0.310	0.300	1,412	0.12	0.06
Tweet emotions: surprise	0.032	0.033	1,412	-0.03	0.63
Tweet emotions: fear	0.015	0.012	1,412	0.20	0.00
Tweet emotions: sadness	0.106	0.104	1,412	0.06	0.34
Tweet emotions: anger	0.154	0.161	1,412	-0.08	0.21
Swear words use	0.051	0.058	1,412	-0.15	0.02
Tweet topic: Bialystok	0.013	0.009	1,328	0.10	0.10
Tweet topic: politics	0.025	0.011	1,412	0.23	0.00
Tweet topic: protests	0.004	0.003	1,412	0.13	0.03
Tweet topic: police	0.002	0.001	1,412	0.14	0.01
Tweet topic: Jews	0.001	0.000	1,412	0.14	0.01
Tweet topic: religion	0.007	0.004	1,412	0.18	0.00
Tweet topic: culture	0.027	0.026	1,412	0.04	0.54
Tweet topic: school	0.029	0.029	1,412	0.00	0.96
Tweet topic: family	0.029	0.027	1,412	0.12	0.06
Tweet topic: friends	0.013	0.011	1,412	0.19	0.00
Tweet topic: love	0.013	0.015	1,412	-0.21	0.00
Tweet topic: boy(friends) and girl(friends)	0.014	0.012	1,412	0.10	0.08
Tweet topic: pets	0.009	0.008	1,412	0.14	0.03
Tweet topic: alcohol	0.004	0.004	1,412	0.02	0.78
Tweet topic: illness	0.002	0.002	1,412	0.14	0.01
Feminine verbs share	0.723	0.812	1,355	-0.23	0.00
1st person verbs share	0.907	0.910	1,360	-0.03	0.64
Singular verbs share	0.937	0.949	1,360	-0.14	0.03
log Network size	4.639	4.597	1,412	0.03	0.58
Network: media	0.014	0.005	1,412	0.22	0.00
Network: politics	0.016	0.008	1,412	0.16	0.01
Network: LGBTQ activists	0.005	0.002	1,412	0.14	0.02
Network: pro-government	0.011	0.007	1,412	0.09	0.12
Network: anti-government	0.023	0.007	1,412	0.25	0.00
Network: NGOs	0.001	0.000	1,412	0.12	0.02
Network: public figures	0.016	0.006	1,412	0.22	0.00
Network: music	0.009	0.020	1,412	-0.31	0.00
Network: K-pop	0.004	0.009	1,412	-0.15	0.03
Network: youtubers	0.004	0.006	1,412	-0.07	0.33
Network: movies	0.001	0.001	1,412	0.08	0.15
LGBTQ coming out: GPT	0.911	0.984	1,372	-0.26	0.00

Notes: This table presents the comparison of the average values of the variables used in the analysis for two groups: participants in the first wave, and participants in the second wave (who did not come out during the first wave). The first column shows the averages of the variables for users who came out during the first wave of the campaign. The second column shows the averages of the variables for users who came out during the second wave of the campaign. The third column shows the number of users in the sample. The fourth column shows the standardized difference in the variables between two groups. The fifth column shows the p-value of a t-test of the equality of averages in the two groups. The description of the variables can be found in Tables C.1-C.5.

Table B.3: Descriptive statistics by user type

	LGBTQ	**Allies**	**Anti-LGBTQ**
Network: LGBT coming out	0.077	0.030	0.012
Network: allies	0.010	0.029	0.014
Network: anti-LGBT	0.006	0.016	0.069
Gender: woman	0.759	0.552	0.158
Gender: man	0.214	0.448	0.842
Gender: transgender / non-binary	0.027	0.000	0.000
log Tweets count	6.724	6.724	7.039
Average tweet length	10.770	12.523	13.876
Replies (% of tweets)	0.451	0.583	0.707
Hashtag use	0.061	0.105	0.098
Emoji use	0.041	0.041	0.036
LGBTQ-related words use	0.015	0.009	0.017
Emotional words share	0.090	0.069	0.050
Tweet sentiment: positive	0.630	0.641	0.553
Tweet sentiment: negative	0.370	0.359	0.447
Tweet emotions: joy	0.308	0.339	0.292
Tweet emotions: surprise	0.032	0.034	0.032
Tweet emotions: fear	0.014	0.018	0.023
Tweet emotions: sadness	0.106	0.116	0.143
Tweet emotions: anger	0.156	0.118	0.142
Swear words use	0.053	0.026	0.017
Tweet topic: Białystok	0.012	0.014	0.012
Tweet topic: politics	0.022	0.087	0.137
Tweet topic: protests	0.004	0.006	0.011
Tweet topic: police	0.002	0.003	0.005
Tweet topic: Jews	0.001	0.001	0.010
Tweet topic: religion	0.006	0.016	0.016
Tweet topic: culture	0.027	0.023	0.013
Tweet topic: school	0.029	0.030	0.029
Tweet topic: family	0.029	0.028	0.033
Tweet topic: friends	0.013	0.012	0.010
Tweet topic: love	0.013	0.008	0.004
Tweet topic: boy(friends) and girl(friends)	0.013	0.010	0.004
Tweet topic: pets	0.009	0.007	0.006
Tweet topic: alcohol	0.004	0.004	0.005
Tweet topic: illness	0.002	0.002	0.002
Feminine verbs share	0.744	0.522	0.167
1st person verbs share	0.908	0.869	0.804
Singular verbs share	0.940	0.904	0.871
log Network size	4.628	5.130	5.701
Network: media	0.012	0.053	0.105
Network: politics	0.014	0.039	0.069
Network: LGBTQ activists	0.004	0.004	0.004
Network: pro-government	0.010	0.036	0.083
Network: anti-government	0.019	0.085	0.103
Network: NGOs	0.000	0.001	0.002
Network: public figures	0.013	0.035	0.052
Network: music	0.011	0.008	0.000
Network: K-pop	0.005	0.001	0.000
Network: youtubers	0.004	0.011	0.000
Network: movies	0.001	0.001	0.000
LGBTQ coming out: GPT	0.928	0.129	0.399

Notes: This table presents the averages of the individual characteristics for three types of users who participated in two waves of the #IamLGBT campaign. The sources and a description of the variables can be found in Tables C.1-C.2, and C.3-C.5.

Table B.4: Correlates of the exposure to peers coming out as LGBTQ

	(1)	(2)	(3)
	Network: LGBTQ coming out	Network: LGBTQ coming out	Network: LGBTQ coming out
Gender: female	-0.178**	-0.192**	-0.162**
	(0.072)	(0.079)	(0.079)
Gender: transgender / non-binary	0.047	-0.007	0.019
	(0.188)	(0.196)	(0.192)
log Tweets count		-0.068***	-0.003
		(0.023)	(0.040)
Average tweet length		-0.002	-0.004
		(0.008)	(0.008)
Replies (% of all tweets)		-0.359***	-0.123
		(0.135)	(0.165)
Hashtag use		-1.052***	-1.001***
		(0.307)	(0.274)
Emoji use		-1.015***	-1.188***
		(0.380)	(0.385)
LGBTQ-related tweets		0.146	-0.668
		(0.919)	(0.816)
Emotional words share		-1.233*	-1.113
		(0.738)	(0.696)
Positive tweet sentiment		-0.169	-0.175
		(0.227)	(0.223)
Network: allies			0.000
			(0.013)
Network: anti-LGBTQ			-0.024*
			(0.013)
log Network size			-0.082
			(0.058)
Network: media			-2.074***
			(0.628)
Network: politicians			0.569
			(0.865)
Network: LGBTQ activists			13.717***
			(1.899)
Hour FE	yes	yes	yes
Adj. R-Squared	0.00	0.03	0.07
Number of clusters	1412	1412	1412
Observations	36128	36128	36128

Notes: Table shows the OLS estimates of a model where the dependent variable is exposure to peers coming out as LGBTQ. The dependent variable measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. * $p < .10$; ** $p < .05$; *** $p < .01$

Table B.5: Peer effects and coming out: full results

	(1)	(2)	(3)	(4)	(5)
Network: LGBTQ coming out	0.005***	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Gender: female			-0.004	-0.002	-0.001
			(0.002)	(0.002)	(0.002)
Gender: transgender / non-binary			-0.008*	-0.008*	-0.008*
			(0.005)	(0.005)	(0.005)
log Tweets count				0.003***	0.003**
				(0.001)	(0.001)
Average tweet length				0.001***	0.001***
				(0.000)	(0.000)
Replies (% of all tweets)				-0.001	-0.004
				(0.005)	(0.006)
Hashtag use				0.006	0.004
				(0.010)	(0.010)
Emoji use				0.004	0.003
				(0.015)	(0.016)
LGBTQ-related tweets				0.040	0.039
				(0.032)	(0.032)
Emotional words share				-0.019	-0.002
				(0.025)	(0.025)
Positive tweet sentiment				0.004	0.005
				(0.008)	(0.008)
Network: allies					0.004***
					(0.001)
Network: anti-LGBTQ					0.001
					(0.001)
log Network size					-0.000
					(0.002)
Network: media					0.035
					(0.029)
Network: politicians					0.020
					(0.029)
Network: LGBTQ activists					-0.036
					(0.080)
Hour FE	no	yes	yes	yes	yes
Adj. R-Squared	0.00	0.02	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03	0.03	0.03
Number of clusters	1412	1412	1412	1412	1412
Observations	36128	36128	36128	36128	36128

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. All regressors are based on Twitter activity in the period between January, 1 2019 and July, 28 2019 (one day before the start of the campaign). We use standard errors clustered at the user level.
* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.6: Peer effects and coming out: excluding users who did not participate in the first wave of the campaign

	(1)	(2)	(3)	(4)	(5)	(6)
Network: LGBTQ coming out	0.004***	0.004**	0.004**	0.004**	0.004**	0.004**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Hour FE	no	yes	yes	yes	yes	yes
Gender	no	no	yes	yes	yes	yes
Twitter activity	no	no	no	yes	yes	yes
Network	no	no	no	no	yes	yes
Network: second wave	no	no	no	no	no	yes
Adj. R-Squared	0.00	0.03	0.03	0.04	0.04	0.04
Mean of outcome	0.06	0.06	0.06	0.06	0.06	0.06
Number of clusters	1067	1067	1067	1067	1067	1067
Observations	18533	18533	18533	18533	18533	18533

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. "Network: second wave" is the fraction of the network who came out as an LGBTQ person during the second wave of the campaign and not during the first wave of the campaign. The remaining control variables are described in the note of Table 2. All regressors are based on Twitter activity in the period between January, 1 2019 and July, 28 2019 (one day before the start of the campaign). The sample includes only users who participated in the first wave of the campaign. We use standard errors clustered at the user level.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.7: Peer effects and coming out: covariates selected by LASSO

	(1) Baseline	(2) LASSO Baseline variables	(3) LASSO Extended set of variables	(4) LASSO Extended, incl. verb forms
Network: LGBTQ coming out	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Number of potential control variables		67	111	114
Number of selected control variables		60	72	71
Mean of outcome	0.03	0.03	0.03	0.03
Observations	36128	36128	33669	32760

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out, where the set of covariates is selected using a double-selection LASSO procedure (see Belloni et al., 2014). 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. In each specification, we control for the full set of hour fixed-effects. Column 1 presents the results for the baseline set of covariates. Column 2 presents the results for the LASSO-selected variables from the pool of baseline covariates. Column 3 presents the results for the extended set of covariates (baseline variables and variables listed in Tables C.3-C.5). Column 4 presents the results for the extended set of covariates including verb forms variables (for which some observations are missing due to not using gender-specific verb forms). All covariates are based on Twitter activity in the period between January, 1 2019 and July, 28 2019 (one day before the start of the campaign). We exclude journalists, elected officials, and political party members from the sample. We use standard errors clustered at the user level. Table B.8 shows the list of selected covariates in each specification.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.8: LASSO: lists of selected covariates

Column	**Variables Selected by LASSO**
(2)	Emoji use; Gender: woman; Hashtag use; log Network size; Replies (% of tweets); Network: anti-LGBT; Network: LGBTQ activists; Network: media; Average tweet length; Emotional words share; Tweet sentiment: positive
(3)	Emoji use; Gender: woman; Hashtag use; log Network size; Network: anti-LGBT; Network: K-pop; Network: LGBTQ activists; Network: media; Network: movies; Network: music; Network: NGOs; Network: pro-government; Network: public figures; Network: youtubers; Average tweet length; Tweet topic: Białystok; Tweet topic: culture (extended); Tweet topic: illness; Tweet topic: pets (extended); Tweet topic: police (extended); Tweet topic: politics (extended); Tweet topic: protests; Swear words use
(4)	Emoji use; Gender: woman; Hashtag use; log Network size; Network: anti-LGBT; Network: K-pop; Network: LGBTQ activists; Network: media; Network: music; Network: NGOs; Network: pro-government; Network: public figures; Network: youtubers; Average tweet length; Singular verbs share; Tweet topic: Białystok; Tweet topic: culture; Tweet topic: illness; Tweet topic: pets; Tweet topic: police (extended); Tweet topic: politics (extended); Swear words use

Notes: Table shows the list of covariates selected in the LASSO double-selection procedure presented in Table B.7. The values in the first column correspond to the column numbers of Table B.7, and the second column shows the list of selected covariates. We do not report the hour fixed-effects which are included in each specification. The description of the variables can be found in Tables C.1-C.5.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.9: Placebo: effects of other peer tweets (not LGBTQ-related)

	(1)	(2)	(3)	(4)	(5)
Network: any peer tweet	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Network: LGBTQ coming out	yes	yes	yes	yes	yes
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
Adj. R-Squared	0.00	0.02	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03	0.03	0.03
Number of clusters	1408	1408	1408	1408	1408
Observations	36048	36048	36048	36048	36048

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. 'Network: any peer tweet' measures the fraction of the network who posted at least one tweet in the preceding hour. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. The control variables are described in the note of Table 2. We use standard errors clustered at the user level.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.10: Peer effects and coming out: IV estimates, log network variable

Panel A: Second stage

	LGBTQ coming out				
	(1)	(2)	(3)	(4)	(5)
Network (log): LGBTQ coming out	0.008***	0.008***	0.008***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
F-statistic	208.52	208.25	202.00	184.56	186.72
Number of clusters	1408	1408	1408	1408	1408
Observations	36048	36048	36048	36048	36048

Panel B: First stage

	Network: LGBTQ coming out				
	(1)	(2)	(3)	(4)	(5)
Peers of peers: LGBTQ coming out	0.487***	0.487***	0.489***	0.469***	0.441***
	(0.034)	(0.034)	(0.034)	(0.035)	(0.032)
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
F-statistic	208.52	208.25	202.00	184.56	186.72
Number of clusters	1408	1408	1408	1408	1408
Observations	36048	36048	36048	36048	36048

Notes: Panel A shows the estimates of the two-stage least-squares estimation of a duration model of the hourly probability of coming out. Panel B shows the results of the first-stage regressions. The log exposure to peers coming out as LGBTQ was instrumented by the log exposure of peers to LGBTQ coming out of their peers who were not user's peers themselves. For each hour, the network variables and the instrument are standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. All regressors are based on Twitter activity in the period between January, 1 2019 and July, 28 2019 (one day before the start of the campaign). We use standard errors clustered at the user level.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.11: Peer effects and coming out: excluding journalists, elected officials, and members of political parties

	(1)	(2)	(3)	(4)	(5)
Network: LGBTQ coming out	0.005***	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
Adj. R-Squared	0.00	0.02	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03	0.03	0.03
Number of clusters	1381	1381	1381	1381	1381
Observations	35446	35446	35446	35446	35446

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. All regressors are based on Twitter activity in the period between January, 1 2019 and July, 28 2019 (one day before the start of the campaign). We exclude journalists, elected officials, and political party members from the sample. We use standard errors clustered at the user level.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.12: Peer effects and coming out: excluding coming out posts without any photos

	(1)	(2)	(3)	(4)	(5)
Network: LGBTQ coming out	0.005***	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
Adj. R-Squared	0.00	0.02	0.02	0.02	0.02
Mean of outcome	0.02	0.02	0.02	0.02	0.02
Number of clusters	1102	1102	1102	1102	1102
Observations	30681	30681	30681	30681	30681

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. All regressors are based on Twitter activity in the period between January, 1 2019 and July, 28 2019 (one day before the start of the campaign). We exclude users who did not attach any photo to their coming out post. We use standard errors clustered at the user level.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.13: Peer effects and coming out: only those active during the first wave of the campaign

	(1)	(2)	(3)	(4)	(5)
Network: LGBTQ coming out	0.006***	0.006***	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
Adj. R-Squared	0.00	0.02	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03	0.03	0.03
Number of clusters	1410	1410	1410	1410	1410
Observations	35493	35493	35493	35493	35493

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. The sample includes only users who were active during the first wave of the Twitter campaign (posted at least one tweet in the period from 29 July 17:00 to 4 August 23:59). We use standard errors clustered at the user level.
* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.14: Peer effects and coming out: alternative standard errors

	(1)	(2)	(3)
	Clustered: user	Robust	Two-way clustered: user and hour
Network: LGBTQ coming out	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)
Hour FE	yes	yes	yes
Gender	yes	yes	yes
Twitter activity	yes	yes	yes
Network	yes	yes	yes
Adj. R-Squared	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03
Number of clusters	1412		51
Observations	36128	36128	36128

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. All regressors are based on Twitter activity in the period between January, 1 2019 and July, 28 2019 (one day before the start of the campaign). In column 1, we show the baseline results with standardard errors clustered at the user level. In column 2, we show results with robust standard errors. In column 3, we show results with two-way clustered standard errors at the level of the user and the hour.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.15: Peer effects and coming out: unweighted network variables

	(1)	(2)	(3)	(4)	(5)
Network (unweighted): LGBTQ coming out	0.003**	0.003**	0.002**	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
Adj. R-Squared	0.00	0.02	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03	0.03	0.03
Number of clusters	1412	1412	1412	1412	1412
Observations	36128	36128	36128	36128	36128

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. 'Network (unweighted): LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. Instead of using the measure weighted by the strength of pre-campaign ties, we use calculate the unweighted measure by dividing the number of peers who came out by hour $t-1$ by the number of all peers in the user's network. Then, For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. We use standard errors clustered at the user level.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.16: Peer effects and coming out: exposure in the preceding hour only

	(1)	(2)	(3)	(4)
Network: LGBTQ coming out	0.003^{**}	0.003^{**}		
	(0.001)	(0.001)		
Non-zero exposure to peers coming out			0.013^{***}	0.012^{***}
			(0.004)	(0.004)
Hour FE	yes	yes	yes	yes
Gender	no	yes	no	yes
Twitter activity	no	yes	no	yes
Network	no	yes	no	yes
Adj. R-Squared	0.02	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03	0.03
Number of clusters	1412	1412	1412	1412
Observations	33233	33233	36128	36128

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. In columns 1 and 2, 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ in a preceding hour. In columns 3 and 4, "Non-zero exposure to peers' coming out" is binary variable that equals 1 if the user experienced at least one peer's coming out actions in the preceding hour, and zero otherwise. For each hour, this variable is standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. We use standard errors clustered at the user level.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.17: Peer effects and coming out: including second and third hour of the campaign

	(1)	(2)	(3)	(4)	(5)
Network: LGBTQ coming out	0.005***	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
Adj. R-Squared	0.00	0.02	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03	0.03	0.03
Number of clusters	1447	1447	1447	1447	1447
Observations	39013	39013	39013	39013	39013

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. Compared to the baseline analysis, we include the second hour of the campaign. We control for variables described in the note of Table 2 except for ally and anti-LGBT network variables because no posts from allies and anti-LGBTQ users were recorded during first two hours of the campaign. All regressors are based on Twitter activity in the period between January, 1 2019 and July, 28 2019 (one day before the start of the campaign).

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.18: Peer effects and coming out: users classified by ChatGPT

	(1)	(2)	(3)	(4)	(5)
Network: LGBTQ coming out	0.006***	0.006***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
Adj. R-Squared	0.00	0.02	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03	0.03	0.03
Number of clusters	1273	1273	1273	1273	1273
Observations	32331	32331	32331	32331	32331

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. Compared to the baseline analysis, the sample consists of an intersection of a set of users classified as LGBTQ by ChatGPT and a baseline set of users manually classified as LGBTQ. Similarly, the exposure variable, 'Network: LGBTQ coming outs', is based on the intersection of these two sets.
* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.19: Peer effects and coming out: log network variables

	(1)	(2)	(3)	(4)	(5)
Network (log): LGBTQ coming out	0.005^{***}	0.005^{***}	0.005^{***}	0.005^{***}	0.005^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hour FE	no	yes	yes	yes	yes
Gender	no	no	yes	yes	yes
Twitter activity	no	no	no	yes	yes
Network	no	no	no	no	yes
Adj. R-Squared	0.00	0.02	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03	0.03	0.03
Number of clusters	1412	1412	1412	1412	1412
Observations	36128	36128	36128	36128	36128

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out. 'Network (log): LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. We take logs of the network variable after adding a small number (0.01). Then, For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. We use standard errors clustered at the user level.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.20: Peer effects and coming out: non-standardized alternative exposure definitions

	(1)	(2)	(3)
Network (fraction): LGBTQ coming out	0.087***		
	(0.019)		
Network (log fraction): LGBTQ coming out		0.006***	
		(0.001)	
Network (any): LGBTQ coming out			0.007***
			(0.002)
Hour FE	yes	yes	yes
Gender	yes	yes	yes
Twitter activity	yes	yes	yes
Network	yes	yes	yes
Adj. R-Squared	0.02	0.02	0.02
Mean of outcome	0.03	0.03	0.03
Number of clusters	1412	1412	1412
Observations	36128	36128	36128

Notes: Table shows the OLS estimates of a duration model of the hourly probability of coming out for alternative measures of peer exposure. Unlike in the baseline, we do not standardize the exposure variables. 'Network (fraction): LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. 'Network (log fraction): LGBTQ coming outs' measures the log of 0.01 plus the fraction of the network who came out as LGBTQ before a given hour. 'Network (any): LGBTQ coming outs' is equal to zero if a user experienced no peer coming outs before a given hour and to one if a user experienced at least one peer coming out before a given hour. Control variables are described in the note of Table 2. We use standard errors clustered at the user level.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.21: Peer effects and coming out: Cox model

	(1)	(2)	(3)	(4)
Network: LGBTQ coming out	1.125***	1.121***	1.127***	1.132***
	(0.021)	(0.021)	(0.021)	(0.022)
Gender	no	yes	yes	yes
Twitter activity	no	no	yes	yes
Network	no	no	no	yes
Number of clusters	1412	1412	1412	1412
Observations	36128	36128	36128	36128

Notes: Table shows the estimates of Cox proportional hazards regressions. The coefficients are reported in terms of proportional hazards (exponentiated coefficients). 'Network: LGBTQ coming outs' measures the fraction of the network who came out as LGBTQ before a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. We use standard errors clustered at the user level.
* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.22: Peer effects and coming out: parametric proportional hazards model

	(1)	(2)	(3)	(4)
Network: LGBTQ coming out	1.118***	1.114***	1.119***	1.123***
	(0.023)	(0.024)	(0.023)	(0.024)
Gender	no	yes	yes	yes
Twitter activity	no	no	yes	yes
Network	no	no	no	yes
Number of clusters	1412	1412	1412	1412
Observations	36128	36128	36128	36128

Notes: Table shows the estimates of parametric proportional hazards random effects model with exponential survival distribution. The coefficients are reported in terms of proportional hazards (exponentiated coefficients). 'Network: LGBTQ coming outs' measures the fraction of the network who came out as an LGBTQ person by a given hour. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. We use standard errors clustered at the user level.
* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.23: Peer effects and coming out time: parametric AFT model

	(1)	(2)	(3)	(4)
Network: LGBTQ coming out	-0.095***	-0.094***	-0.092***	-0.102***
	(0.022)	(0.022)	(0.021)	(0.022)
Gender	no	yes	yes	yes
Twitter activity	no	no	yes	yes
Network	no	no	no	yes
Number of clusters	1412	1412	1412	1412
Observations	36128	36128	36128	36128

Notes: Table shows the coefficients from an AFT random effects model with log-normal survival distribution where the dependent variable is the natural log of the hour of the coming out for a given individual, and 'Network: LGBTQ coming outs' measures the fraction of the network who came out as an LGBTQ person by a given hour. Negative estimates imply that the independent variable accelerates the decision to come out. For each hour, the network variables are standardized with a mean of zero and a standard deviation of one. Control variables are described in the note of Table 2. We use standard errors clustered at the user level.

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.24: Peer effects and the probability of coming out beyond the viral campaign

	(1)	(2)	(3)	(4)
Network: LGBTQ coming out	0.029	0.025	0.026	0.026
	(0.021)	(0.022)	(0.022)	(0.023)
Gender	no	yes	yes	yes
Twitter activity	no	no	yes	yes
Network	no	no	no	yes
Adj. R-Squared	0.01	0.11	0.14	0.15
Mean of outcome	0.12	0.12	0.12	0.12
N	382	382	382	382

Notes: Table shows the effects of the exposure to peers coming out as LGBTQ on the probability of coming out after the viral phase of the campaign had ended. 'Network: LGBTQ coming outs' measures the fraction of the network who came out as an LGBTQ person by the end of the first wave. The dependent variable is a dummy variable which is equal to one for users who decided to come out between August, 10 2019 and May, 26 2020, and is equal to zero for users who came out during the second wave of the campaign (May, 27-29 2020). We control for gender (female, male, transgender / non-binary), pre-campaign measures of Twitter activity (log tweets count, average tweet length, hashtag use, emoji use, share of LGBTQ-related tweets, replies as percentage of all tweets, emotional words use, positive tweet sentiment), and network characteristics (log network size, replies to media, politics, LGBTQ activist accounts, the exposure to peer posts by straight allies, and anti-LGBTQ users in the network).

* $p < .10$; ** $p < .05$; *** $p < .01$

Table B.25: Peer effects and coming out: DiD binary treatment variable

	(1)	(2)	(3)	(4)
Network: LGBTQ coming out	0.142^{***}	0.132^{***}	0.116^{***}	0.123^{***}
	(0.037)	(0.038)	(0.038)	(0.040)
Gender	no	yes	yes	yes
Twitter activity	no	no	yes	yes
Network	no	no	no	yes
Adj. R-Squared	0.01	0.02	0.04	0.04
Mean of outcome	0.77	0.77	0.77	0.77
Observations	772	772	772	772

Notes: Table shows the effects of the exposure to peers coming out as LGBTQ on the probability of coming out in the first wave of the campaign. The exposure variable is a dummy variable that is equal to one for users with the exposure greater or equal to the average, and to zero otherwise. The sample consists of participants of the first wave who joined the campaign after the 19th hour as well as users who joined the campaign only during the second wave. The exposure variable measures cumulative exposure up until the 19th hour. We use standard errors clustered at the user level.

* $p < .10$; ** $p < .05$; *** $p < .01$

Appendix C Data Appendix

Lists of tweets with the hashtag #IamLGBT

The list of users who participated in the first wave of the Twitter campaign was constructed manually on October 17-18, 2019 (two months after the campaign). We searched for tweets from the first three days of the campaign (July, 29 - July, 31) that included the *#jestemLGBT* search term, and manually wrote a file with the screen names of users participating in the campaign, as well as their user type and gender. The list of users who participated in the second wave of the Twitter campaign was downloaded first on May 28, 2020 (one day after the start of the campaign), and then extended for users who participated in May 28-29 period on May 30, 2020 (one day after the end of the campaign). This time, we automated the process of retrieving data, as the data were downloaded using the *GetOldTweets3* library. This library scrapes the Twitter search results for a given term.

On July 13, 2021, we additionally scraped first-wave tweets with the *#jestemLGBT* hashtag from July 29 to August 4 (using *snscrape*, as the *GetOldTweets3* library became obsolete). This scrape included tweets that we could have skipped during the initial manual download, and tweets from days 4-7 of the campaign. Initially, we skipped these days because the number of participants during this period was much smaller than the number of users participating during the first three days of the campaign (see Figure 4). Nevertheless, we decided to add those users to make sure that we were not misclassifying them as "not yet out" individuals.

On June 5, 2023, we additionally scraped the list of tweets with the *#jestemLGBT* hashtag, which were posted between the two waves (August 10, 2019 - May 26, 2020). The data on these additional users is used only in robustness Table B.24.

Classification of user types and genders

After downloading the list of tweets, we manually coded two variables: user type and gender. There were four types of users: users who came out as LGBTQ, non-LGBTQ allies, anti-LGBTQ haters, and unknown. Users who came out as LGBTQ were detected based on two patterns. First, some users explicitly added information about their sexual orientation or gender identity in a tweet (e.g., they mentioned that they are bisexual or that they have a partner of the same gender). These direct statements were sometimes captioned with attached photos. Second, some users wrote "#IamLGBT and I am", ending the sentence with a description of their occupation or personal characteristics. See examples of the *IamLGBT* tweets of users classified as LGBTQ below:

> We come from the same town. I am LGBT - I am not an ideology - **I am a lesbian** and what @Janusz1967 writes is blatant homophobia that is discrimination based on psychosexual orientation. I was taught by the church to love my fellow human beings. I see the priest was taught hatred!

> **I'm LGBT and I'm** the one who's always the laughing girl from your school who's gonna defend that younger kid they're bullying, help the old lady with her shopping, but no matter how hard I try, I'm still a pedophile in people's eyes. I've had enough.

Allies were detected based on a supportive statement that included an explicit declaration that they themselves are not LGBTQ (the most common pattern was the phrase "nie #jestemLGBT ale" which means "I am not LGBT but" followed by a supportive statement. See an example below:

> **I am not LGBT**, but I support you, I cheer you on.

Anti-LGBTQ haters were detected in several ways. Some tweets included explicitly homophobic or transphobic hate speech. Some tweets contained explicit homophobic or transphobic hate speech. Some tweets contained obscene images or pictures of totalitarian leaders. Some trolling tweets included pictures of well-known politicians and their false coming out statements. Some posts criticized the campaign for sharing private information (e.g., "nothing to brag about"), accused foreign powers of organizing the campaign, or made other negative comments about the campaign. See examples below:

> And I'm not LGBT and I also want to feel normal among #IamLGBT and you don't interest me at all **until you are vulgar and want to impose your**

customs on me. It takes pity when I look at this contrived, unnatural otherness.

I am LGBT **and I am also a loyal SA-man, a proud Aryan son of the Reich**. I will defend your borders. I serve all Germans, no difference if poor or rich.

The remaining tweets were classified as "unknown." This category includes tweets that did not contain any additional text beyond the hashtag, tweets disseminating information about the campaign, positive or neutral tweets in which users did not disclose their LGBTQ or non-LGBTQ identity, and tweets about topics unrelated to the campaign (using the popularity of the hashtag to draw attention to themselves). We manually detected accounts of organizations (media, NGOs, European Commission) that posted tweets, so we can account for these tweets, but we may have classified some posts by LGBTQ individuals as "unknown." However, we believe that participation in the IamLGBT campaign did not necessarily require LGBTQ users to come out. Therefore, tweets that merely spread messages or expressed support without revealing the user's identity were classified as "unknown" and not as representing the coming out of LGBTQ individuals.

During both waves of the campaign, the largest group of tweets were classified as coming out tweets by LGBTQ people (Figure C.1). In the first wave of the campaign, we recorded over 2,200 coming out actions by LGBTQ users, and over 1,500 such tweets in the second wave of the campaign. There were significantly more anti-LGBTQ and unknown users in the first wave of the campaign than in the second wave of the campaign. More than 1,000 LGBTQ participants in the second wave did not participate in the first wave; nearly 300 LGBTQ users tweeted statements about coming out in both waves of the campaign, and 27 users participated in the first wave of the campaign but did not explicitly disclose their identities (see Figure C.2).[8] In our analysis, the timing of an LGBTQ person's coming out corresponds to the timing of the first tweet with the hashtag *IamLGBT*. Therefore, users

[8]For the remaining 79 users, we were unable to retrieve their unique Twitter IDs, so we do not know whether they participated in the first wave. Moreover, it is possible that some users participated in the first wave of the campaign, but we classify them as non-participants if they deleted their coming out tweet before we retrieved the list of participants in the first wave.

who posted tweets about LGBTQ people coming out during both waves of the campaign were classified as having come out during the first wave of the campaign. Participants in the second wave of the campaign who joined the first wave without posting an LGBTQ coming out statement, or who did not participate in the first wave of the campaign were classified as having come out during the second wave of the campaign.

We used ChatGPT API (*gpt-3.5-turbo*) to validate our classification. ChatGPT is a state-of-the-art large language model, which is designed to understand and generate human-like text based on the input it receives.[9] The model is trained on a diverse range of internet text, allowing it to acquire knowledge and language patterns from a wide variety of sources. ChatGPT utilizes a deep learning architecture, which enables it to capture complex dependencies and long-range contextual information within the text. One of the key features of ChatGPT is its ability to engage in conversational interactions, understanding and generating text in a dialogue format. For each participant, we generate an API request that consists of the phrase "Does this person identify as LGBT? Answer with one word, yes or no:" and the text of the coming out post (the first post with the hashtag *IamLGBT*). Hence, we rely on the interpretation of the coming out post by the ChatGPT model. The API response is either yes or no. 93 percent of users, whom we manually classified as LGBTQ, were found to be LGBTQ by ChatGPT (see Table B.3). The precision is slightly lower for allies than for LGBTQ users (13 percent of straight allies are recognized as LGBTQ by ChatGPT). The precision of ChatGPT is lowest when it comes to identifying anti-LGBTQ users, as it recognizes 40 percent of them as LGBTQ. This discrepancy is largely due to ChatGPT's difficulty in recognizing sarcasm.

We classified users' gender based on their declarations and text analysis of their tweets. First, we examined whether users explicitly stated their gender in their coming out tweets. Second, we exploited the fact that verbs in Polish have gender inflection. Specifically, gender can be inferred from the past tense ending of verbs written in the first person singular (morphological endings also allow us to distinguish the singular of verbs in the first person singular). Verbs in the masculine gender end in "łem" and verbs in the feminine gender end in "łam". The non-binary endings include "łxm" and "łom". However, the use of these

[9] https://platform.openai.com/docs/models/gpt-3-5

endings is still quite rare, and non-binary individuals may also use one of the traditional forms. We compared the frequency of verbs with these endings to determine the gender of individuals. In the few cases in which it was impossible to determine gender (users did not use the first person singular forms of verbs, or they used them just as often), we classified their gender as "unknown".

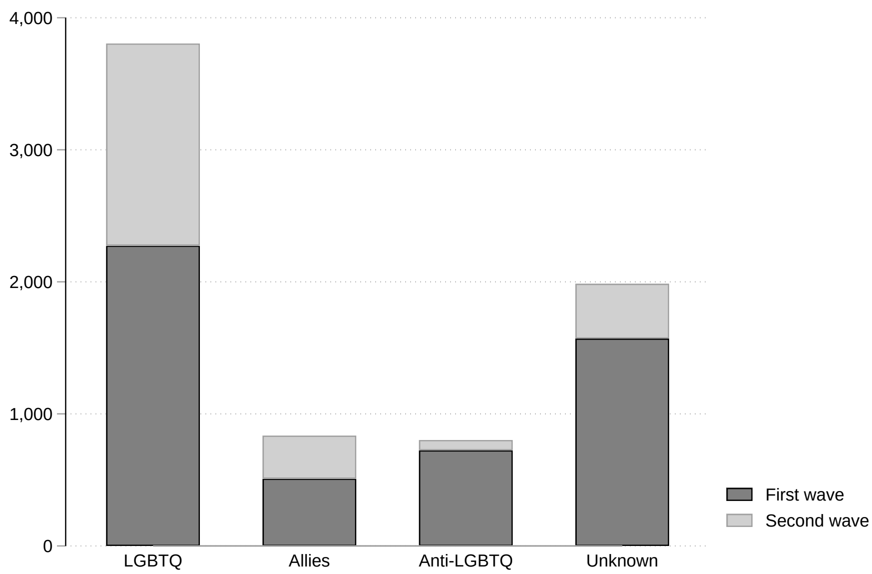


Figure C.1: The number of users participating in the *IamLGBT* campaign, by wave of the campaign and user type

Notes: Figure shows the number of all recorded users who participated in at least one of the two waves of the *IamLGBT* campaign (users who participated in both waves are recorded twice).

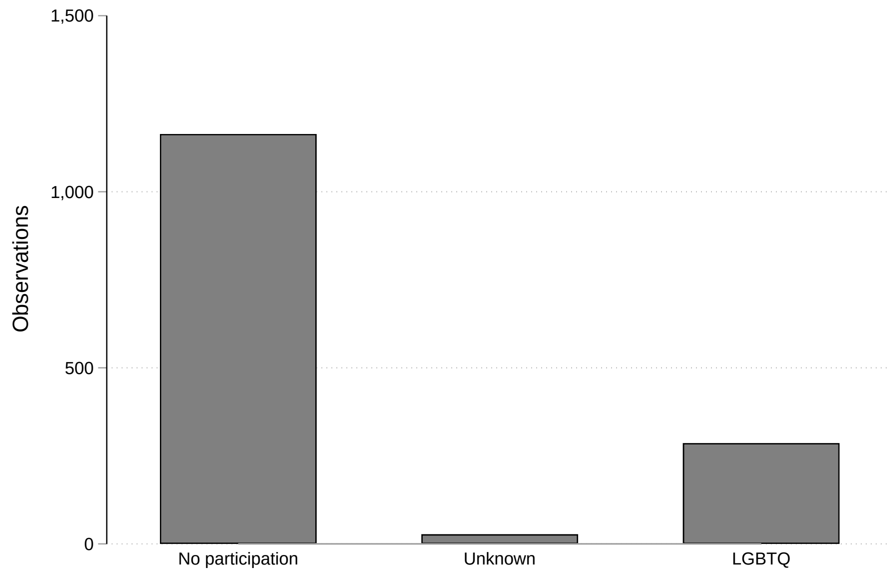


Figure C.2: The number of users participating in the second wave of the *IamLGBT* campaign, by participation in the first wave of the *IamLGBT* campaign

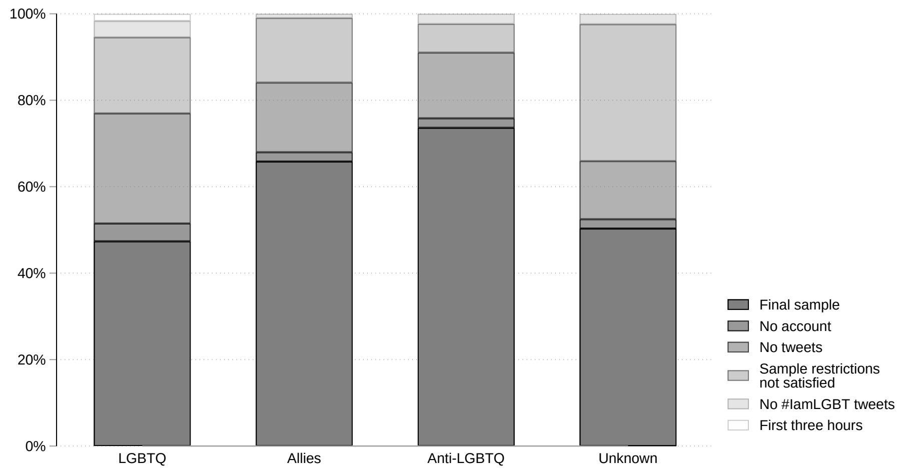
Notes: Figure shows the number of all recorded users who participated in the second wave of the *IamLGBT* campaign, depending on their participation in the first wave of the campaign. The "No participation" bar shows the number of second wave participants who did not participate in the first wave. The "Unknown" bar shows the number of second wave participants who participated in the first wave of the campaign as the unknown type (did not directly disclose their identity). The "LGBTQ" bar shows the number of second wave participants who had already disclosed their LGBTQ identity during the first wave.

Twitter activity data

Twitter activity data were extracted using the library *snscrape*. For each user, we downloaded up to 30,000 recent tweets published between January 1, 2019 and November 1, 2019. The user activity file contains the following information about each published tweet: unique tweet ID, timestamp, content, number of likes, number of retweets, number of replies, whether the tweet is a reply or an original tweet, the screen name of the reply recipient, and the unique account ID of the reply recipient (for reply tweets only). The Twitter activity data were downloaded in August 2021. We initially downloaded Twitter activity data for users who participated in the first wave of the campaign in November 2019 and for users who participated in the second wave of the campaign in June 2020. However, we did not retrieve the account ID information of the reply recipients (we only retrieved their screen names). Because screen names can be changed by users, our network variables may have underestimated the contribution of users who participated in the first wave to the networks of users who participated in the second wave (if users changed their screen names between November 2019 and June 2020). Therefore, we use the data downloaded again for all users in August 2021, which contains account ID information that is consistent over time. The drawback of this approach is that we lose the Twitter activity information of users whose accounts had been deleted, suspended, or made private. The Twitter activity data may also be incomplete for users who deleted some of their tweets from the study period.

Figure C.3 shows the fraction of all recorded users who were not included in the sample for all user types and the two waves of shares. We see that about 50 percent of all recorded LGBTQ users from the first wave of the campaign were included in the final sample. The most common reasons for not being included in the final sample were a lack of tweets in the pre-campaign period and a failure to meet the sample restrictions. The lack of tweets in the pre-campaign period could be due to Twitter suspending the Twitter account, changing the account type to private, creating an account during the campaign, reactivating an old account for the purposes of participating in the campaign, and deleting tweets from the pre-campaign period. The final sample included less than 40 percent of all recorded LGBTQ users from the second wave of the campaign. The most common reason for not being included in the final sample was a lack of tweets in the period before the campaign started. This problem was more severe for the second wave than the first wave, because an additional reason for missing tweets was that the user created an account after the first wave of the campaign.

(a) First wave



(b) Second wave

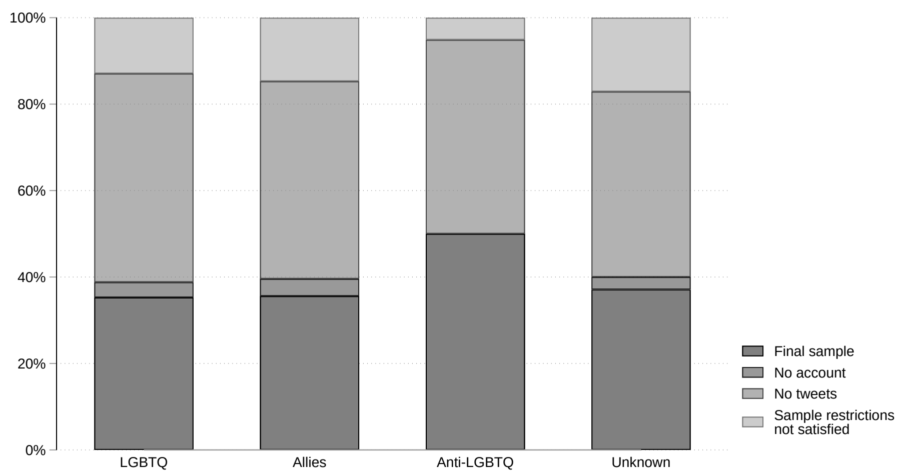


Figure C.3: Users recorded vs. users in the sample

Notes: Figure shows the percentage of users who were recorded and were included in the final sample, and the percentage of users who were recorded but not included in the final sample, broken down by the reason for not being included. "No account" indicates users for whom a unique Twitter account ID was not found (those who removed their account or changed their screen name). "No tweets" indicates users for whom we found no tweets in the pre-campaign period. "Sample restrictions not satisfied" indicates users whose pre-campaign activity did not satisfy the sample restrictions. "No #IamLGBT" tweets indicates the number of participants in the first wave of the campaign who removed their *IamLGBT* posts (or their posts were removed by Twitter admins). "First three hours" denotes the number of participants in the first wave of the campaign who joined the campaign during first three hours of the campaign.

Weekly panel data: effects of coming out

The weekly panel used in the DiD analysis consists of eight weeks. A week starts on Monday at 12:00 AM and ends on Sunday on 11:59 PM. The campaign week starts on July 29th (the day of the start of the campaign) and ends on August 4th. The pre-campaign period consists of four weeks starting on July 1st and ending on July 28th. The post-campaign period includes the campaign week and three weeks between August 5th and August 25th.

For each user, we calculate the value of a given variable in a given week based on all tweets she posted during a given week. The only exception are tweets, which contain the hashtag *IamLGBT*. We exclude these tweets from calculating the Twitter activity variables. We calculate the following variables:

Log tweet count: log count of tweets posted in a given week.

Positive sentiment: the number of words that are associated with a positive sentiment divided by the number of all words that express emotions posted in a given week.

Topic: illness: the number of tweets that include illness-related words divided by the count of all tweets posted in a given week.

Replies to LGBTQ users: the number of replies to users who decided to come out during the first wave of the campaign divided by the total number of replies in a given week.

Topic: LGBTQ: the number of tweets that include LGBTQ-related words divided by the count of all tweets posted in a given week.

Topic: boyfriend / girlfriend: the number of tweets that include keywords related to the boy(friends) and girl(friends) divided by the count of all tweets posted in a given week.

Variable descriptions

Table C.1: Variable descriptions (i.)

Variable	Description
Dependent variables	
R_i	a binary variable that equals one if the user came out during the first wave of the campaign, and equals zero if the user came out during the second wave but not during the first wave (hand coded)
Gender	
Gender: female	a binary variable that equals one if the user uses female pronouns / verb forms and equals zero otherwise (hand coded)
Gender: male	a binary variable that equals one if the user uses male pronouns / verb forms and equals zero otherwise (hand coded)
Gender: transgender / non-binary	a binary variable that equals one if the user disclosed transgender or non-binary identity and equals zero otherwise (hand coded)
Twitter activity	
Tweets count	the count of tweets
Average tweet length	average number of characters of tweets
Replies (% of all tweets)	the count of replies divided by the count of all tweets
Hashtag use	the count of hashtags divided by the word count of all tweets
Emoji use	the count of emojis divided by the word count of all tweets
LGBTQ-related words use	the number of tweets that include LGBTQ related keywords in Polish or English (word stems of "LGBT", "homosexual", "lesbian", "gay", "bisexual", "transgender", "transsexual", "non-binary", "homophobia", "transphobia", "pansexual", "asexual") divided by the count of all tweets
Emotional words share	the number of words that express emotions divided by the count of all words (excl. stop words)
Positive tweet sentiment	the number of words that are associated with a positive sentiment divided by the number of all words that express emotions

Notes: Description of the variables used in the baseline analysis. All variables except for R_i and 'Gender: transgender / non-binary' are based on Twitter activity before the first wave of the Twitter campaign (January 1, 2019 - July 28, 2019). The network variables are based on replies during the pre-campaign period as well, while the information about the user type is obtained from tweets that were a part of the "IamLGBT" Twitter campaign.

Table C.2: Variable descriptions (ii.)

Variable	Description
Network	
Network size	the count of all users to whom the user replied at least once (replies)
Network: LGBTQ coming out	the number of replies to LGBTQ users who came out during the first wave of the campaign ($R_i = 1$) divided by the total count of replies (LGBTQ user type dummy variable hand coded)
Network: allies	the number of replies to allies ('I am not LGBTQ but I support this campaign') who participated in the first wave of the campaign divided by the total count of replies (ally user type dummy variable hand coded)
Network: anti-LGBTQ	the number of replies to anti-LGBTQ users who participated in the first wave of the campaign divided by the total count of replies (anti-LGBTQ user type dummy variable hand coded)
Network: media	the number of replies to journalists and news accounts (participants of the two waves of the campaign and those among the most popular accounts in the network) divided by the total count of replies (media account type dummy variable hand coded)
Network: politicians	the number of replies to elected officials, members of political parties and parties accounts (participants of the two waves of the campaign and those among the most popular accounts in the network) divided by the total count of replies (politician account type dummy variable hand coded)
Network: LGBTQ activists	the number of replies to LGBTQ activists and LGBTQ organization accounts who participated in the first wave of the campaign (participants of the two waves of the campaign and those among the most popular accounts in the network) divided by the total count of replies (LGBTQ activist account type dummy variable hand coded)

Notes: Description of the variables used in the baseline analysis. All variables are based on Twitter activity before the first wave of the Twitter campaign (1 January 2019 - 28 July 2019). The network variables are based on replies during the pre-campaign period as well, while the information about the user type is obtained from tweets that were a part of the "IamLGBT" Twitter campaign.

Table C.3: Extended list of covariates for LASSO (i.)

Variable	Description
Twitter activity	
Tweet sentiment: negative	the number of words that are associated with a positive sentiment divided by the number of all words that express emotions
Tweet emotions: joy	the number of words that express joy divided by the number of all words that express emotions
Tweet emotions: surprise	the number of words that express surprise divided by the number of all words that express emotions
Tweet emotions: fear	the number of words that express fear divided by the number of all words that express emotions
Tweet emotions: sadness	the number of words that express sadness divided by the number of all words that express emotions
Tweet emotions: anger	the number of words that express anger divided by the number of all words that express emotions
Swear words use	the number of tweets that include swear words divided by the count of all tweets
Tweet topic: Białystok	the number of tweets that include Białystok-related words divided by the count of all tweets (post-Białystok violence - 20th July-28th July only)
Tweet topic: politics	the number of tweets that include politics-related words divided by the count of all tweets
Tweet topic: protests	the number of tweets that include protests-related words divided by the count of all tweets
Tweet topic: Jews	the number of tweets that include Jews-related words divided by the count of all tweets
Tweet topic: religion	the number of tweets that include religion-related words divided by the count of all tweets
Tweet topic: culture	the number of tweets that include culture-related words divided by the count of all tweets
Tweet topic: school	the number of tweets that include school-related words divided by the count of all tweets
Tweet topic: family	the number of tweets that include family-related words divided by the count of all tweets
Tweet topic: friends	the number of tweets that include friends-related words divided by the count of all tweets

Notes: Description of the variables used in the double-selection LASSO procedure. All the variables are based on the text analysis of Polish- and English-language tweets. See Appendix D for the detailed description of the text analysis as well as emotions and topic keywords dictionaries.

Table C.4: Extended list of covariates for LASSO (ii.)

Variable	**Description**
Twitter activity	
Tweet topic: boy(friends) and girl(friends)	the number of tweets that include keywords related to the boy(friends) and girl(friends) divided by the count of all tweets
Tweet topic: alcohol	the number of tweets that include alcohol-related words divided by the count of all tweets
Tweet topic: illness	the number of tweets that include illness-related words divided by the count of all tweets
Tweet topic: politics (extended)	the number of tweets that include politics-related words divided by the count of all tweets, extended keywords list including most similar words (word embeddings)
Tweet topic: protests (extended)	the number of tweets that include protests-related words divided by the count of all tweets, extended keywords list including most similar words (word embeddings)
Tweet topic: Jews (extended)	the number of tweets that include Jews-related words divided by the count of all tweets, extended keywords list including most similar words (word embeddings)
Tweet topic: religion (extended)	the number of tweets that include religion-related words divided by the count of all tweets, extended keywords list including most similar words (word embeddings)
Tweet topic: culture (extended)	the number of tweets that include culture-related words divided by the count of all tweets, extended keywords list including most similar words (word embeddings)
Tweet topic: school (extended)	the number of tweets that include school-related words divided by the count of all , extended keywords list including most similar words (word embeddings)
Tweet topic: family (extended)	the number of tweets that include family-related words divided by the count of all tweets, extended keywords list including most similar words (word embeddings)
Tweet topic: friends (extended)	the number of tweets that include friends-related words divided by the count of all tweets, extended keywords list including most similar words (word embeddings)
Tweet topic: boy(friends) and girl(friends) (extended)	the number of tweets that include keywords related to the boy(friends) and girl(friends) divided by the count of all tweets, extended keywords list including most similar words (word embeddings)
Tweet topic: alcohol (extended)	the number of tweets that include alcohol-related words divided by the count of all tweets, extended keywords list including most similar words (word embeddings)
Tweet topic: illness (extended)	the number of tweets that include illness-related words divided by the count of all tweets, extended keywords list including most similar words (word embeddings)

Notes: Description of the variables used in the double-selection LASSO procedure. The topic variables are based on the text analysis of Polish- and English-language tweets. The verb forms variables are based on the text analysis of Polish-language tweets. See Appendix D for more details.

Table C.5: Extended list of covariates for LASSO (iii.)

Variable	Description
Twitter activity	
Feminine verbs share	the number of verbs used in the feminine form divided by the number of all verbs inflected for gender
First person verbs share	the number of verbs used in the first person form (singular and plural) divided by the number of all verbs inflected for the grammatical person
Singular verbs share	the number of verbs used in the singular form divided by the number of all verbs inflected for the grammatical person
Network: pro-government	the number of replies to pro-government accounts (the most popular accounts in the network) divided by the total count of replies (pro-government account type dummy variable hand coded)
Network: anti-government	the number of replies to anti-government accounts (the most popular accounts in the network) divided by the total count of replies (anti-government account type dummy variable hand coded)
Network: music	the number of replies to musicians' accounts (the most popular accounts in the network) divided by the total count of replies (musician account type dummy variable hand coded)
Network: K-pop	the number of replies to K-pop artists' and fan clubs' accounts (the most popular accounts in the network) divided by the total count of replies (musician account type dummy variable hand coded)
Network: youtubers	the number of replies to youtubers' accounts (the most popular accounts in the network) divided by the total count of replies (youtuber account type dummy variable hand coded)
Network: movies	the number of replies to movie stars' accounts (the most popular accounts in the network) divided by the total count of replies (movies account type dummy variable hand coded)

Notes: Description of the variables used in the double-selection LASSO procedure. The topic variables are based on the text analysis of Polish- and English-language tweets. The verb forms variables are based on the text analysis of Polish-language tweets. In the network analysis, the most popular accounts are accounts with the outdegree centrality of at least one (517 accounts). See Appendix D for more details.

Appendix D Text analysis

Text preparation

For each user, we generated a list of tweets based on her pre-campaign Twitter activity. First, we selected the user's tweets that were posted in the pre-campaign period (January 1, 2019 - July 8, 2019). Next, we converted the tweets to strings, and lowercase these strings. Then, we removed the URL links and the user mentions from the tweet texts. The list of tweets prepared in this way was used for generating the following variables: "Tweets count", "Average tweet length", "Hashtag use", and "Emoji use". For the remaining variables, we additionally removed the tweets without text (e.g., tweets consisting of emojis or URLs only).

Language detection

We used a model included in the *fastText* library to detect the language of tweets (Joulin et al., 2016).[10] For each tweet, the function returns a list of detected languages and the respective confidence of the sentence belonging to those languages. We defined the language with the highest confidence score as the language of the tweet. We used this information in two ways. First, for each user, we generated a list of tweets in Polish language, and a list of tweets in English language, as these lists would be important in the next steps of the text analysis. Second, we generated two variables: the share of Polish tweets, and the share of English tweets.

Most common lemmas

We generated the list of the most common lemmas in the following way (separately for Polish and English language): we pooled all users' pre-campaign tweets, tokenized them, and calculated the number of occurrences for each token. We refer to tokens as words, although tokens include misspelled words, and various forms of lemmas. Then, we selected words with more than 10 occurrences. We obtained 200,000 unique words in Polish and 36,000 words in English. Next, we used the *Spacy* library to lemmatize all the words.[11]

[10] For the full documentation, see `https://fasttext.cc/docs/en/language-identification.html`

[11] For the full documentation, see `https://spacy.io/`.

We obtained 90,000 lemmas in Polish and 22,000 lemmas in English. For each lemma, we obtained the number of its occurrences, and all used forms of the lemma.

Sentiment and emotion analysis

First, we obtained a dictionary of lemmas associated with positive and negative sentiment. To this end, we utilized the Valence Aware Dictionary and sEntiment Reasoner (VADER). It is a lexicon and rule-based sentiment model that is specifically attuned to sentiments expressed in social media.[12] For each lemma, the model returns compound score, which ranges from -1 to 1, where -1 indicates extremely negative sentiment, 0 indicates neutral sentiment, and 1 indicates extremely positive sentiment. Over 92 percent of lemmas have compound scores equal to zero. We classified lemmas with compound scores above 0.3 as lemmas related to a positive sentiment, and lemmas with compound scores below -0.3 as lemmas expressing a negative sentiment. By combining all the positive- and negative-sentiment lemmas, we formed a set of lemmas defined by us as emotional words, which constitute approximately 5 percent of the total number of lemmas.

Next, we use a pre-trained Bidirectional Encoder Representations from Transformers model (BERT) to assess the association between lemmas and six distinct emotions: fear, anger, surprise, sadness, love, and joy.[13]. The model returns a score indicating the likelihood of a given text (in our case, lemmas) expressing a specific emotion, with the scores for six emotions always summing up to one. Within our dictionary of emotion-related keywords, we include lemmas that meet two criteria: their emotion scores surpass the 95th percentile for the corresponding emotion, and they are identified as emotional words (either positive or negative) by the *VADER* model.

For Polish lemmas, we first translated them into English using the *Google Translate* library, and then we assigned them to sentiments/emotions using the *VADER* and *BERT* models. The *Google Translate* library is particularly useful in the case of the Polish language words, as it is able to translate even misspelled words (including missing diacritics). The final dictionaries included all the forms of lemmas.

[12] For the full documentation, see `https://pypi.org/project/vaderSentiment/`

[13] For the full documentation, see `https://huggingface.co/bhadresh-savani/bert-base-uncased-emotion`

Next, for each user, we pooled the person's pre-campaign tweets (Polish and English separately), and tokenized them. Then, we removed stopwords (the list of stopwords was obtained from the *stop_words* library[14]). We calculated the "Emotional words share variable" by dividing the number of emotional words in a given language divided by the number of all words in a given language. We calculated the remaining sentiment and emotion variables by dividing the number of words associated with a given sentiment or emotion by the number of all emotional words. Finally, for each variable, we calculate the weighted average of Polish- and English-language variables, with the weight being calculated by dividing the number of tweets in a given language by the number of Polish and English tweets. The full dictionaries are available online (`https://www.jgromadzki.com/software/gromadzki_siemaszko_dictionaries_keywords.zip`).

Verb forms

The Polish verbs are inflected for gender in the past tense. This allowed us to study how often a given user used feminine and masculine forms. The Polish verbs are also inflected depending on the grammatical person, which made it easy to study how often the users tweet in the first person as opposed to in the second person, and how often they used singular vs. plural forms. There are following singular suffixes: "łem" for the masculine first person, "łam" for the feminine first person, "łeś" and "łaś" for the masculine and feminine second person, respectively. There are following plural suffixes: "liśmy" for the masculine first person, "łyśmy" for the feminine first person, "liście" and "łyście" for the masculine and feminine second person, respectively.[15] These suffixes are almost exclusively used in verbs. Hence, for each user, we pooled the person's tweets, tokenized them, and calculated the occurences of words that ended with suffixes listed above (with diacritics and without diacritics, as some users ignore diacritics). Next, we calculated the "Feminine verbs share" by dividing the number of first person feminine verbs by the number of all first person verbs (feminine or maskuline). Then, we calculate the "first person verbs share" by dividing the number of first person verbs (singular or plural) by the number of all first and second

[14] For the full documentation, see: `https://pypi.org/project/stop-words/`

[15] For more details, see Verbs sections in `https://en.wikipedia.org/wiki/Polish_morphology#Verbs` and `https://en.wikipedia.org/wiki/Polish_grammar#Verbs`.

person. Finally, we calculate the "Singular verbs share" variable by dividing the number of all singular verbs (first and second person) by the number of all plural verbs (first and second person). As the suffixes for the third person are less unique, we are not able to detect third person verbs.

Topic analysis

Finally, we conducted a topic analysis using dictionary method. The keyword lemmas were manually selected in the following way:

- We selected the most often frequently used lemmas (top 1000 lemmas for the Polish dictionary, and top 500 lemmas for the English dictionary).
- We manually classified these lemmas into 12 topics (most of the lemmas remained unassigned to any of the topics). We also found lemmas of swear words.
- Then, we manually searched for other lemmas that had same stems. For example, to the political topic's lemma "president", we added lemmas with the same stem: "presidential", "presidency". The additional lemmas were searched for among lemmas with at least 100 occurences in users' tweets. This was particularly important for the Polish dictionary, as the Spacy lemmatizer of the Polish language is not as precise as the English one, and the users often ignored diacritics (e.g, they used "kosciol" instead of the correct form "kościół").

Next, for each user, we tokenized each of her tweets. Then, for each topic, we calculated the number of tweets that included at least one topic keyword. The final variable was calculated by dividing the number of tweets with at least one occurrence of a topic keyword divided by the number of all tweets.

The full dictionaries that include all forms of lemmas (obtained by using Spacy lemmatizers) are available online (`https://www.jgromadzki.com/software/gromadzki_siemaszko_dictionaries_keywords.zip`). All strings were lowercased, and the lemmas are delimited by commas.

In robustness, we extend the list of keywords using pretrained word embeddings models from *fasttext*. The models that transforms word into vectors were trained on Common Crawl and Wikipedia and are available for many languages, including Polish and English.[16] We use cosine similarity as a metric of similarity of two word vectors. First, for each manually selected keyword in a given topic, we find ten most similar lemmas in the model's corpus. We obtain a long list of similar lemmas for the given topic (ten times the original number of keywords less duplicates). Next, for each similar lemma, we calculate its average similarity with each of the topic's keywords. In our final dictionary, we include lemmas with the highest average similarity and the size of the final list of most similar lemmas is twice the original size. The extended dictionary includes both the original and the most similar lemmas.

While the use of word embeddings reduces the sparsity and the arbitrariness of the list of keywords, it may complicate the interpretation of the variable based on an extended dictionary. Word embeddings are generated from a corpus of text that reflects a particular cultural context, and they can be influenced by existing stereotypes within that culture. For instance, the cosine similarity between words islam and terrorism in the English-language model is three times larger than the similarity between words catholic and terrorism. It is also larger than the similarity between words islam and catholic, despite both describing religions. This bias is evident in the context of LGBTQ topics. While the English-language model finds similar words such as "lgbtqia", "homosexuals", and "queer", the similar words according to the Polish-language model include words such as "pedophile" and "neonazi".

[16] For full documentation, see `https://fasttext.cc/docs/en/crawl-vectors.html`

Appendix E Examples of tweets

I'm LGBT and I'm pansexual, not that I'm attracted to pans.

I'm LGBT and my goal is to control my hair, not the world, although success encourages expansion.

I'm LGBT and I'm living in Poland even though I don't know for how long. I'm not an ideologist. I love and I feel. And unfortunately, I am also increasingly feeling fear and helplessness.

my name is vincent and I am LGBT. I'm not a pedophile; I'm only 16 years old and I have my whole life ahead of me. My own father doesn't tolerate who I am. I was mocked for not being "normal" but still I am LGBT and I'm proud of it.

I'm LGBT and I'm the one who's always the laughing girl from your school who's gonna defend that younger kid they're bullying, help the old lady with her shopping, but no matter how hard I try, I'm still a pedophile in people's eyes. I have had enough.

I'm very supportive of the whole campaign and I'm LGBT, but I don't want to put my pictures here, because our country is shit and if someone who shouldn't see it were to, they might not let me live.

I'm LGBT and I believe there will come a time when I won't be afraid for my life by saying it out loud.

I'm not LGBT, but I was crying at the Christmas table after a fight with my uncle about it, seeing how much hatred and ignorance can be in people, so I can also help you to raise your hashtag, kisses xx

Every day I worry about my parents finding out about my orientation and kicking me out of the house, and to marry the person I love I'd have to leave the country, just because I'm LGBT.