#IamLGBT: Social networks and coming out in a hostile environment*

Jan Gromadzki†, Przemysław Siemaszko‡

Preliminary – please do not quote

January 22, 2022

Abstract

Recent decades have witnessed a remarkable increase in the number of people disclosing their LGBTQ identity. We propose a model of a binary-action supermodular game on a network with social learning to investigate the role of peer effects in coming out decisions and analyse the perfect Bayesian equilibrium we obtain. We collected unique data on coming outs which occurred during two spontaneous Twitter actions in Poland. We use these data to empirically test the hypothesis that observing peers coming out increases the probability that an individual will make a decision to disclose their LGBTQ identity. The estimated peer effects are strong and statistically significant. The spread of information about the existence of the action through networks does not explain the results. Instead, we argue that these effects are due to changing beliefs about the costs of disclosure.

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

JEL Codes: J15, D85, D74, P16, Z13

*We are grateful to Joanna Franaszek, Nelson Mesker, and Danisz Okulicz for very helpful discussions. We thank participants at the AEA/ASSA Annual Meeting 2022, the 2021 Discrimination and Diversity Workshop, and the 2021 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.

†SGH Warsaw School Economics and Institute for Structural Research, Email: `jan.gromadzki[at]ibs[dot]org[dot]pl`

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

1 Introduction

A rapid increase in the number of openly LGBTQ individuals is one of the most spectacular examples of cultural change in recent decades. In the US, the proportion of respondents who stated they knew someone who was gay or lesbian rose from less than 25% in the early 1980s to nearly 90% in 2013. This surge in coming out decisions remains largely unexplained. In this paper, we use a unique setting to examine the role of network effects in LGBTQ individuals' decisions to come out in a hostile environment. We show that observing peer disclosure significantly increases the probability of LGBTQ identity disclosure.

What is the impact of social networks on individual decisions to disclose a concealable stigmatized identity? Our hypothesis is that peer effects matter in this case because they affect perceived levels of discrimination. Individuals learn about the costs of coming out by observing their peers. They update their beliefs about the negative consequences of disclosing a stigmatized identity both immediately (if my friends came out, they must have known it wouldn't be so bad) and by observing the consequences after a period of time.

We develop a theoretical model of social learning on networks in games of binary action and demonstrate the mechanisms by which network effects operate. We show the unique equilibrium in the proposed game and provide comparative statics examining the effects of changing state of the world and learning mechanisms. A large, spontaneous coming out action on Twitter allowed us to collect the unique data needed to empirically investigate the role of peer effects. In 2019, after violent attacks against a Pride march in Poland, thousands of Polish users posted coming-out tweets with the hashtag 'IamLGBT' (*#jestemLGBT*) to show that they are human beings and not an ideology. A few months later, the action with the hashtag 'IamLGBT' took place again, and some users who did not join the first action, this time decided to come out. Thus, using data on revealed preferences, we are able to determine the exact time of the costly coming out action of hundreds of LGBTQ users. We downloaded Twitter activity data for all participants, which allowed us to elicit their network and generate additional variables. We exploit the variation in the intensity of LGBTQ peer coming outs in users networks to estimate the effects of peers' coming outs on individual probability of coming out.

Our paper contributes to the growing literature on cultural change. Giuliano (2020) points out that the literature has long focused on the persistence of social norms, while theoretical and empirical evidence on cultural change is scarce. Our theoretical model is very closely related to the model of social learning developed by Fernández (2013) and we combine it with a network approach. In empirical studies, strong network effects were found in female labor supply (Nicoletti et al., 2018), paternity leave take-up (Dahl et al., 2014), and participation in student protests (Bursztyn, Cantoni, et al., 2021; F. González, 2020). Bursztyn, A. L. González, et al. (2020) showed that correcting misperceptions of others' beliefs influences individual preferences regarding women at work. Fernández et al. (2021) studied the impact of the AIDS epidemic on changing social attitudes toward gay men and lesbians. They argue that the change in social attitudes was due to the increased visibility of homosexuals, as the epidemic led to the mobilization of the LGBTQ community. We, on the other hand, examine the determinants of the decision to come out at the individual level. An external threat (the AIDS epidemic in the U.S., violent attacks on the LGBTQ community in Poland) in both cases serves as a trigger for collective action, but compared to Fernández et al. (2021), we are interested in the micro-level determinants of participation rather than the effects of the action.

We also contribute to the literature on LGBTQ people. Recent literature documents persistent discrimination against LGBTQ people (Badgett et al., 2021). Existing research has found that gay men in same-sex couples earn less than men in opposite-sex couples, and lesbians in same-sex couples earn more than women in opposite-sex couples (Aksoy, Carpenter, and Frank, 2018; Aksoy, Carpenter, Frank, and Huffman, 2019; Carpenter, 2005). Plug et al. (2014) found that LGBTQ people try to avoid discrimination by sorting into occupations with low levels of anti-LGBTQ sentiment. Transgender workers also face discrimination (Campbell et al., 2021; Geijtenbeek and Plug, 2018). Importantly, recent studies show that institutional changes have sizable effects on the situation of LGBTQ people: legalization of same-sex marriages reduces discrimination in the labor market, improves partnership stability, and mental health (Chen and van Ours, 2020, 2021; Sansone, 2019). Empirical research on the micro-level determinants of LGBTQ identity disclosure decisions is lacking, most likely due to limited data availability. We overcome these challenges by using our unique dataset to identify coming out actions and elicit networks of LGBTQ individuals. While most existing studies rely on survey data with stated preferences, we extend the literature that uses revealed preferences data (e.g., Geijtenbeek and Plug, 2018). Furthermore, while existing

research focuses almost exclusively on relatively liberal Western countries, we examine coming out decisions in a society with very high levels of anti-LGBTQ sentiment, similar to what prevailed in Western countries in the early 1990s.

Furthermore, we add a theoretical contribution to the literature on weighted majority games or global games. We propose a model of a binary-action game of strategic complements on a network (M. O. Jackson and Zenou, 2015) with Bayesian learning about the true state of the world, that is also supermodular. Despite multiple possible outcomes in terms of actions of the game, we find a unique equilibrium of the game in terms of thresholds (Oyama and Takahashi, 2020), expressed as a belief operator, which happens to have a closed form solution. We provide comparative statics of the equilibrium with respect to the model's parameters and show how personal and social aspects, as well as learning mechanisms influence individual decisions regarding joining a collective, risky action (Frankel et al., 2003; de Martí and Zenou, 2015).

Finally, we contribute to literature on social media and collective action. Existing studies showed that social media increased engagement in civic and political life (Boulianne and Theocharis, 2020; Fergusson and Molina, 2019). Some studies showed that online platforms may be particularly useful for spreading xenophobic and populist message and may contribute to political polarization (Bursztyn, Egorov, et al., 2019; Levy, 2021; Zhuravskaya et al., 2020). We show that social media may be used by LGBTQ people to increase their visibility, and we investigate the mechanism of a collective action on Twitter in detail.

We construct an hourly panel dataset, and estimate Cox proportional hazards model to assess the role of LGBTQ peers' coming outs on individual decisions to disclose LGBTQ identity. We overcome the reflection problem by using information on the exact time of the coming out. We find that coming out by peers had a strong impact on individual decisions to disclose LGBTQ identity. We find little variation in the peer effects depending on gender or past Twitter behavior. Peer effects were stronger for peers of the same gender than for peers of a different gender which may be linked to the fact that coming out of peers of the same gender provides better information about the level of discrimination an individual will face. Participation by straight allies had a significant but weaker impact on probability of coming out than observing coming out of LGBTQ

peers which suggests that the effects are not simply due to spreading information about the action. The second reason for ruling out this mechanism is that the effects were strongest after the action was covered in the national media. Although we cannot directly account for the role of unobserved characteristics, we show that the peer effects were relatively small at the beginning of the action, and hence were not driven by users with highest willingness to come out. Moreover, our estimates are unaffected by controlling for a rich set of covariates which further suggests that unobserved heterogeneity should not undermine our results.

The remainder of this paper is structured as follows. In Section 2, we introduce a theoretical model of coming out decisions. Section 3 outlines the situation of LGBTQ people in Poland and details of the Twitter action. In Section 4, we describe our data and econometric strategy. We present our empirical results on the peer effects in coming out decisions in Section 5. Section 6 concludes.

2 Theoretical model

The society consists of set of n individuals $\mathbb{I} = 1, 2, 3, ..., n$, who are connected by a connected graph $\mathbb{G}$, where $\mathbb{G}_{i,j} \in [0, \inf)$ is a weight of a directed edge from player i to j. Weights of edges indicate the influence a decision of player i has on a network effect experienced by player j. All members face a decision to come out with their orientation to their network by joining a collective movement on social media. In each period, LGBTQ individuals decide to perform an action $R_i = 0$ of concealing their type or a risky, irreversible decision to come out $R_i = 1$. The timeline of the model looks as follows:

- 0 • Players' characteristics and network are drawn by nature
- 1 • Players obtain their perceived cost of coming out
- 2 • Players decide to come out or not
- 3 • Information about players' decisions are transmitted through the network
- 4 • Players update their beliefs about the cost of coming out through the network effect
- 5 • Steps 2 - 4 are repeated until no player wants to join the movement.

2.1 Utility maximisation

Players decide to come out at each time t by maximizing their utility,

$$\max_{R_{it}} U(R_{it}) = R_{it} \times P_i - R_{it} \times D(y_{i,t}), \tag{1}$$

where P_i is the expected payoff of coming out (likes, new followers, etc.) summed with an individual cost of concealing (stress, depression, blackmail) given by $P_i \sim \mathcal{N}(\bar{P}, \epsilon)$. Normal distribution of K was chosen, as biological response to stress and victimization is correlating with different hormones (such as cortizol) levels, which are normally distributed across population, however some of them are skewed. Finally, $D(y_{i,t})$ is player i's expectation of level of discrimination given the decisions of other players. The decisions taken by the individuals happen in continuous time and any player is can decide to join the movement at any stage of the game.

A player will only decide to come out if expected utility of coming out is higher than the cost of concealing:

$$P_i - \hat{D}_i \geq 0 \iff P_i - D(y_{i,t}) \geq 0. \tag{2}$$

At each period, if player didn't decide to come out, she observes the actions of other players in her network. She obtains a measure of participation in the action

$$y_{i,t} = \frac{\sum_{j \in \mathbb{I}} \mathbb{G}_{ji} \alpha R_{jt}}{\sum_{j \in \mathbb{I}} \mathbb{G}_{ji} \alpha} \tag{3}$$

which is a share of neighbors participating in the action weighted by their connection strength and a multiplier α which is positive for any member of the LGBTQ community in player's network and 0 otherwise. Players use the observations to update their beliefs of discrimination in the society $\hat{D}_i$. As α and $\mathbb{G}$ are positive, the game is supermodular.

For simplicity the actual level of discrimination D in society can take only two values $D \in D_L, D_H$ where $D_L < D_H$. It is unknown to player's before they make the decision. The problem becomes trivial in cases when $D_H < P_i$, as player i will always choose to come out, or when $D_L > P_i$, when she would never participate in the action. In the rest of the paper we only analyze the case when $D_H \geq Pi$ and $D_L \leq P_i$.

If we look at individual costs of not coming out we can find out the interval that allows player to ever participate in the action. The upper and lower bound of P_i are given respectively:

$$P_{UB} = D_L, \tag{4}$$

$$P_{LB} = D_H. \tag{5}$$

Players' prior belief about the level of discrimination is given by the log likelihood ratio $\lambda_t = ln\left(\frac{Pr(D=D_L)}{Pr(D=D_H)}\right)$. After observing the state of the world y_{it} they update their belief using the Bayes's rule. The learning mechanism is similar to one proposed by Daron Acemoglu and Ozdaglar (2011), where players notice the share of others taking the risky action and judge what is the probability that the discrimination level is of high or low type. The new log likelihood ratio is given by:

$$\lambda_{i,t+1}(y_{i,t+1}) = \lambda_t + ln\left(\frac{Pr(y_{i,t}|D=D_L)}{Pr(y_{i,t}|D=D_H)}\right) = \lambda_t + ln\left(\left(\frac{1-\Phi(\frac{D_L-\bar{P}}{\sqrt{\epsilon}})}{1-\Phi(\frac{D_H-\bar{P}}{\sqrt{\epsilon}})}\right)^{ny_{i,t}}\left(\frac{\Phi(\frac{D_L-\bar{P}}{\sqrt{\epsilon}})}{\Phi(\frac{D_H-\bar{P}}{\sqrt{\epsilon}})}\right)^{n(1-y_{i,t})}\right), \tag{6}$$

where Φ is a standard normal cdf. Here, $y_{i,t}$ is multiplied by n, as players project their network onto the whole society - their opinions are only influenced by their neighbors, but they draw conclusions about society-wide discrimination. The result above follows from the probability that any player i would take a risky action given discrimination level D_x

$$Pr_x := Pr(R_{it}|D=D_x) = Pr(P_i - D_x \geq 0) = 1 - \Phi(\frac{D_x - \bar{P}}{\sqrt{\epsilon}}),$$

and the probability that a share of population $y_{i,t}$ would participate given this individual probability

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

THEOREM 1. *The log likelihood ratio λ_t is linearly increasing in $y_{i,t}$.*

Proof.

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

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. It means that the higher weighed share of player's network declares their participation in the movement, the more likely they perceive the low levels of discrimination in the society as a whole. □

Now we can look for such an $y_{i,t}$ that a player i would feel certain enough about low level of discrimination to participate in the action. Let's call y_i^* the critical value of the weighted share of player i's network participating in the action i.e. $D(y_i^*) = P_i$.

DEFINITION 1. *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 $y_{i,t} > y_i^*$ maximises expected pay-off of agent i given the strategies of other agents y_{-i}^*.*

THEOREM 2. *There exists a unique equilibrium y^* of the proposed game.*

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

$$p_i^* D_L + (1 - p_i^*) D_H = P_i \implies p_i^* = \frac{P_i - D_H}{D_L - D_H} = \frac{D_H - P_i}{D_H - D_L}. \tag{7}$$

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

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

This condition is only satisfied if:

$$y_i^* = \frac{ln\left(\frac{Pr(D=D_H)}{Pr(D=D_L)}\frac{D_H - P_i}{P_i - D_L}\right) - nln\left(\frac{1-Pr_L}{1-Pr_H}\right)}{n\left(ln\left(\frac{Pr_L}{Pr_H}\right) - ln\left(\frac{1-Pr_L}{1-Pr_H}\right)\right)}, \tag{8}$$

which is the lowest possible value of participation among player i's neighbors at which she would come out as a part of the action. For each player such level is unique and depends only on parameters of the model. □

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 the equation 8.

2.2 Comparative statics

In the following section we analyze how the equilibrium threshold y_i^* for each player changes as parameters of the models change. First we want to see how initial state of the game changes the outcomes.

COROLLARY 1. *Higher initial belief in low discrimination levels, or it's log ratio to high levels, λ_t decreases the required threshold of participating users y_i^*.*

Proof. Taking a derivative wrt $Pr(D = D_L)$ gives us a proof:

$$\frac{\partial y_i^*}{\partial Pr(D = D_L)} = \frac{1}{Pr(D = D_L)n\left(ln\left(\frac{1-Pr_L}{1-Pr_H}\right) - ln\left(\frac{Pr_L}{Pr_H}\right)\right)}. \tag{9}$$

This is always lower than 0. Therefore, as player's prior suggests the low levels of discrimination are more likely, she'll require less confirmation from her network to participate in the action. □

COROLLARY 2. *Increasing the benefit a player draws from participating in the action - P, increases the chances of a player to participate, or decreases the threshold y_i^*.*

Proof. From equation (6) we have that $p_i^* = \frac{D_H - P_i}{D_H - D_L}$. This gives us $\frac{\partial p_i^*}{\partial P} = \frac{1}{D_L - D_H}$, which is always negative. Therefore increase in P always leads to lower probability of low discrimination required for a player i to participate in the action. Theorem 1 tells us that perceived probability of $D = D_L$ increases linearly in $y_{i,t}$, therefore with lower p_i^*, y_i^* will be lower too. □

COROLLARY 3. *Increasing any of the costs a player bears when participating in the action - D_L or D_H, decreases the chances of a player to participate, or increases the threshold y_i^*.*

Proof. Following the reasoning from previous corollaries:

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

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

Both those derivative are positive. Therefore higher costs of participation in action decreases the probability a player would join. □

The corollaries above reflect the impact of parameters directly in the utility function. With higher benefit of participating in the action or the cost of not doing so (P_i), player will be more likely to do so. Oppositely, if the costs of participation increase (D) players will be less likely to join the movement.

COROLLARY 4. *Players will be more willing to come out as a part of the action as an effect of increasing size of the population n if the probability of low discrimination levels is low i.e.*

$$Pr(D = D_L) < Pr(D = D_H)\frac{D_H - P_i}{P_i - D_L}. \tag{10}$$

Proof. In the derivative of y_i^* wrt n

$$\frac{\partial y_i^*}{\partial n} = \frac{ln\left(\frac{Pr(D=D_H)}{Pr(D=D_L)}\frac{D_H-P_i}{P_i-D_L}\right)}{n^2\left(ln\left(\frac{1-Pr_L}{1-Pr_H}\right) - ln\left(\frac{Pr_L}{Pr_H}\right)\right)}$$

, the denominator is always negative. For the whole effect to remain negative,numerator has to stay positive. As it is a logarithm, it is enough to show the condition for it's inner value to be higher than 1. □

The meaning of it is that in the model, players would be more willing to come out in bigger societies if their perceived level of discrimination was high. As the population gets bigger the perceived strength of private signals received from the network is lower. If a player receives information about high discrimination being more likely, she would be more inclined to join the action if the representativeness of that signal was low.

OBSERVATION 1. *If new players join player i's network, they fall directly into $y_{i,t}$. Therefore they will encourage player i to join the movement if they themselves are already a part of it, otherwise they will decrease $y_{i,t}$ and make player i less likely to come out.*

LEMMA 1. *To show that a parameter of the model decreases the equilibrium threshold y_i^* it is enough to show that it increases $\frac{Pr(y_{i,t}|D=D_L)}{Pr(y_{i,t}|D=D_H)}$.*

Proof. y_i^* is the minimum value at which player believes the low discrimination levels are probable enough to participate in the action. Her beliefs are updated by adding $ln\left(\frac{Pr(y_{i,t}|D=D_L)}{Pr(y_{i,t}|D=D_H)}\right)$ to the prior, so if an increase in a parameter increases the value inside the logarithm it also increases the perceived probability of low discrimination levels, which allows for lower y_i^* to satisfy the condition. □

It follows the reasoning of the Corollary 1, which shows that if initial belief is more "low", we need lower y_i^*. Lemma 1 shows that if the updating speed is higher with each increase in $y_{i,t}$, y_i^* will be lower too.

COROLLARY 5. *If ϵ increases, or the variance of individual costs of hiding own identity or perceived benefit from coming out becomes greater, players will be more likely to come out only if the share of players already participating is not high enough.*

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.

$$\frac{\partial\Phi(\frac{D_x-\bar{P}}{\sqrt{\epsilon}})}{\partial\epsilon} = -\frac{(D_x-\bar{P})\phi(\frac{D_x-\bar{P}}{\sqrt{\epsilon}})}{2\epsilon^{3/2}}. \tag{11}$$

It is always greater than 0 for D_L and lower for D_H. Now let's put it back into the $Pr(y_{i,t}|D = D_x)$.

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

So for low levels of $y_{i,t}$ with increase of ϵ perceived probability of lower discrimination levels go down as well. Following Lemma 1, it shows that at the beginning of the movement higher uncertainty regarding intrinsic costs of not coming out lead to lower threshold y_i^*. □

The last corollary follows the logic from the corollary 4. If not many players do come out as part of the action, players will be more likely to participate if they believe this signal is less meaningful, which happens when ϵ is higher.

2.3 Theoretical predictions

Having the theorems and corollaries above, we are able to judge whether new individuals would join an action, or would it stop developing. At any given point we just need to establish if there is any person i for whom $y_i \geq y_i^*$.

2.3.1 Beginning

After the initial player starts an action, next participants has to have a very low threshold y_i^* to follow. The first ones would be players with very few connections to others than the initial player. They would immediately observe very high y_i and would significantly increase their $\lambda_{i,t}$. Another important factor for early participants is high cost of concealing included in P_i.

Therefore, such action has the highest chance of starting and not dying at a first participant among a smaller, groups of people with similar characteristics (namely, high cost of concealing). That is a typical description of an echo chamber that appears on social media, where individuals of similar opinions interact with each other.

In such echo chamber, players would ideally have high K, leading to much lower y_i^*, and a group of members would have few or no connections outside of the echo chamber. That would lead to quick increase of their y_i that would surpass their individual thresholds. Their observations porbably would not be a correct representation of an actual state of the world (Golub and M. Jackson, 2012) but to achieve the success of the action, players should expect lower discrimination rather than a correct value. With each next person within the echo chamber joining the movement, others would observe higher y and through higher λ decrease their y_i^*, leading to more and more individuals meeting the condition of $y_i \geq y_i^*$ until majority of the echo chamber being participants.

2.3.2 Development

After the initial group of participants further development of an action would require for more connected member of the echo chamber to reach out to others with high P. Still the key to having an individual participate in an action is for them to observe high y, so the action would require the members of the echo chamber to have a many common neighbors outside of the group.

For the action to quickly develop, the highly connected members of the initial group have to be influential outside of it - have high $\mathbb{G}_{\text{נכ}}$ with individuals j and therefore impact them more than

others. Ideally for the development of an action, players in the echo chamber would have high in-degree from within the group and high out-degree towards individuals outside the clique.

The crucial step for the development of an action would be to "infect" another echo chamber. If that succeeds, a viral development can continue through interconnected social groups.

2.3.3 Centrality of participating players

In the proposed model, players' decision to participate in an action bases solely on their expected utility of doing so. After they make their choice to come out, they do not have more choices to make. Their decision, however, impacts other players' expectations in a way impacted by their out-degree in relation to in-degree of their neighbors.

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}:

$$I_{ij} = \frac{\mathbb{G}_{ij}}{\sum_{k \in \mathbb{I}} \mathbb{G}_{kj}}, \tag{12}$$

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. Total, direct effect of i's coming out on her neighbors is given by:

$$I_i = \sum_{j \in \mathbb{I}} \frac{\mathbb{G}_{ij}}{\sum_{k \in \mathbb{I}} \mathbb{G}_{kj}}. \tag{13}$$

Indirectly, player i's coming out may cause a domino effect by causing their neighbor to come out, which would cause players in their network to join the action and so forth. Therefore player's final impact on the success of the action can be measured using the eigenvalue centrality, but using the intensity matrix I instead of the traditional degrees.

2.3.4 Dying out

For an action to die out, it need not to meet the conditions from the sections above. If players that participate have very low out-degree and are not connected with people outside of their bubble, an

action would be limited to a closed environment. They wouldn't be able to continue a viral growth through neighboring groups.

If the levels of costs of concealing are low across the society, single initial group of participants may not have enough power to convince anyone else to join them.

2.4 Why certain events start public actions?

From Corollary 6 we see that in case of early actions, increase in variance of individual costs of hiding own identity or views would cause more people to join the movement. This may be an explanation to why some events start the revolution and some not. Being exposed to an event like the attacks on the Bialystok marches would increase the uncertainty about others' costs of concealing and the likelihood of more people bearing more costs of it. At the same time higher ϵ caused by such events would lead to number of people having very high costs of not coming out.

The more people are influenced by such events and the more they increase P_i the higher likelihood they will spark a movement acting in the other direction. As the variance of personal costs grows, increasing number of individuals would consider it more likely that the action would succeed and join it.

Another example may be the Black Lives Matter protests after the death of George Floyd. It was one of many events that were directed against black people in the USA, but got enough media coverage to cause more people to change their approach to speaking up about their views on discrimination.

3 Background

LGBTQ people in Poland

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 fraction of LGBTQ people who always avoid holding same-sex partner's hands in public for fear of being assaulted (58% compared to the EU-average of 30%). The fears are not irrational: LGBTQ individuals in Poland experience physical attacks for being LGBTQ more often than in any other EU country. According to the ILGA 2021 report, the status of LGBTQ rights in

Poland is the worst among European Union countries. Although same-sex sexual activity is legal, 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 (ILGA-Europe, 2021).

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, right-wing media, and catholic bishops launched a hate 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 "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 persecution of LGBTQ people, campaign leaders claimed they were against "LGBT ideology", not LGBT people.

At the same time, mobilization of the LGBTQ community and 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 20 July 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 similar incidents had not previously occurred at Pride marches, and the police had successfully protected the marches from counter-protests. A few days later, 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.

#jestemLGBT

Mass coming out action started on Twitter on the afternoon of 29 July 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 action and the hashtag #IamLGBT (*#jestemLGBT*) became the top trending hashtag in Poland. The tweets followed the pattern described in the initial tweet: in addition to coming out statement, users wrote about their jobs or career plans and that they were not ideology. Users often attached photos of themselves. Some users expressed their fear and helplesness, while some tweets had humorous elements in them. In addition to clear coming out tweets, some users did not explicitly come out (e.g., tweeting the hashtag only). Straight allies ("I'm not LGBT, but I support this action"), anti-LGBTQ users sending offensive replies, and large organizations and media outlets also joined in. Examples of tweets can be found in Appendix D.

The action was widely echoed in both Polish and foreign media. Major newspapers and TV channels tweeted about the action just a few hours its launch. The European Commission expressed its support for the action by tweeting a statement with the *#jestemLGBT* hashtag. After the two first days, the action began to die out, and we recorded last posts on 4 August 2019.

The second wave of the action took place on 27 May 2020 after a defamation case was dismissed in court against a right-wing activist who claimed on TV that "gays want to adopt children to rape them". The action was initiated by the same user as the first action and used the same hashtag. Although the second action was less successful than the first one in terms of media coverage, several thousand users participated, and the *#jestemLGBT* hashtag was one of the top trending hashtags in Poland. Importantly for our study, hundreds of users who did not join the first action decided to come out in the second action.

4 Data

We collected data from Twitter using Twitter API and additional libraries. First, we downloaded the list of all tweets that included the *#jestemLGBT* hashtag posted during two waves of the action. Second, we manually classified tweets into four groups: coming outs, tweets from straight allies, tweets from anti-LGBTQ users, and tweets of unknown type. We also manually assigned gender to users using either explicitly stated preferred pronouns or grammatical gender of tweets[1]. Then, we downloaded all tweets of the users who participated in the two waves of the action from the period 1 January 2019 - 30 October 2019. We describe data collection process in detail in Appendix A.

For each user, we need to have enough information to obtain meaningful measures of network intensity or tweets content. Thus, we restricted the sample based on users' Twitter activity to address these issues. First, users had to post at least 25 tweets and 10 replies during the pre-action period (1 January 2019 - 28 July 2019) to be included in the sample. The minimum pre-action network size was 5. In the robustness checks, we investigate sensitivity of our results to these sample restrictions. Our goal is to analyze the probability of coming out only for users who were active during the first action. Hence, we need to exclude users who did not join the action simply because they did not use Twitter during the action or registered on Twitter after the action. Unfortunately, we have no access to individual data about exact dates of logging-in on Twitter app. In order to exclude users who were not on Twitter during the action, we restrict our analysis to users who posted at least one reply tweet between 20-28 July 2019 (one week prior to the action). Figure 1 shows that the difference in the Twitter activity between the two groups was stable before the action. We see that this difference increases only on the day of the Twitter action. This is due to an increase in Twitter activity of participants and not due to a decrease in Twitter activity of those who did not join the action. The activity of users who did not participate in the action was stable at around 14 tweets per day both before and after the Twitter action started. We could have restricted our sample to users who posted at least one tweet during the action (29 July - 4 August). However, this selection criterion would be endogenous to the action. Nevertheless, in the robustness section, we present results for users who posted at least one tweet between 29 July and 4 August. Finally, our goal is to analyze the coming out of users who were not publicly known as LGBTQ prior to

[1] In Polish language, past tense is inflected for gender.

the action. Hence, we exclude LGBTQ activists from the sample.

We construct an hourly frequency panel dataset with outcome and network variables that vary over time. Our outcome variable, R_{it}, equals 1 if the user came out by hour t and 0 otherwise. Our main network variable is the fraction of the network that disclosed its LGBTQ identity by t weighted by the strength of pre-action ties between users i and j

$$NI_{it}^{LGBTQ} = \frac{\sum_{j \in \mathbb{I}} \mathbb{G}_{ji} R_{j,t-1}}{\sum_{j \in \mathbb{I}} \mathbb{G}_{ji}} \tag{14}$$

The connection strength is approximated by the number of replies sent from user i to user j. Hence, our measure captures not only the share of users in the network that came out by a given hour but also pre-action intensity of ties with these users. For each hour, the network variable is standardized with zero mean and standard deviation of one. Standardization is based on user not yet out at t to avoid simultaneity. We construct similar variables to control for the proportion of peers who joined the action as anti-LGBTQ users and straight allies.

5 Empirical strategy

We employ two approaches to assess the role of networks in individual coming out decisions during the first wave of the #IamLGBT action. Since coming out is an absorbing state, we cannot analyze changes in the probability of coming out over time for a single user using the panel fixed-effect estimator. Hence, in our baseline approach, we estimate a hazard model to test whether observing LGBTQ peers' coming outs affects the probability of coming out. In addition, we estimate linear probability models at 6-hour intervals to test how peer effects evolved over time.

In our baseline approach, we estimate the Cox proportional hazards model:

$$\lambda_{it}(\tau) = \lambda_0(\tau) exp(\kappa NI_{it-1}^{LGBTQ} + \beta X_{it}) \tag{15}$$

where $\lambda_{it}(\tau)$ is the hazard rate of coming out for user i at time t, conditional on not being out for τ hours since the beginning of the action. The hazard rate is a function of the baseline hazard $\lambda_0(\tau)$,

the fraction of the network that disclosed their LGBTQ identity before hour t, $NI_{i,t-1}^{LGBTQ}$, and additional characteristics X_{it}. Hence, in the construction of the network variable, we only include coming-outs preceding the individual decision to come out. For each hour, the network variable is standardized with zero mean and standard deviation of one. Standardization is based on user not yet out at t to avoid simultaneity. Positive values of κ mean that observing a peer coming out increases the instantaneous probability of coming out. In our sample, we include users who came out in the first action, and users who came out in the second wave of the action. Hence our sample includes users who decided not to come out in the studied period (first wave of the action). In a robustness test, we additionally estimate the Cox model on the sample restricted to users who came out in the first wave of the action.

We start our analysis from the fourth hour of the action since first anti-LGBTQ and ally users joined the action during the third hour of the action. In robustness tests, we present results including the second hour of the action without controlling for the participation of straight allies and anti-LGBTQ users (since we use lagged network variables, we do not analyze coming outs in the first hour of the action). We end our analysis after 54 hours of the action, because until that point, at least one participant joined every hour, except at night (see Figure B.2). After that, there were only a few coming outs with long gaps between them.

In addition, we estimate the following cross-sectional model of the probability of joining the action for selected hours t conditional on not participating in the coming out action by hour $t-6$:

$$R_{i,t} = \beta_t + \theta_t NI_{i,t-6}^{LGBTQ} + \beta_t X_{i,t} + \epsilon i, t \tag{16}$$

We use 6-hour intervals to have sufficient number of coming outs in analyzed intervals. The network variable measures the fraction of the network that disclosed their LGBTQ identity before hour $t-5$ (the starting point of the analyzed interval). Hence, the parameter θ_t describes the impact of the coming out of peers on an individual's decision to disclose LGBTQ identity during the 6-hour period ending at hour t. The estimated effects should be treated as a lower bound of the peer effects because our network variable does not include coming out of peers that occurred during the 6-hour

period. We also provide the results for the network variable defined at hour $t-1$ ($NI_{i,t-1}^{LGBTQ}$). Since this variable includes coming outs of peers that occurred both immediately before and after the individual's own coming out, these additional estimates provide an upper bound of the effects.

Two major challenges in identifying the effects of coming out by LGBTQ peers are simultaneity and omitted variable bias. The simultaneity arises from the reflection problem: looking at an individual's exposure to LGBTQ peer coming outs at the end of the action would conflate coming outs preceding the coming out of an individual, and coming outs that occurred after user's coming out, potentially caused by the user's coming out. The advantage of our data is that we can solve the simultaneity: we have information about exact time of coming out posts, and we use lagged network variables to explain the probability of coming out.

By contrast, we are unable to directly address the issue of unobserved heterogeneity. It is likely that exposure to LGBTQ coming outs is correlated with factors that increase the probability of coming out, e.g., having more LGBTQ peers, location, personality traits, education, age, attitudes of friends and family, etc. We are able to provide some insights into the potential role of unobserved characteristics by controlling for a rich set of covariates. The variation of the effects over time may also inform us about the role of unobservable factors. If exposure to LGBTQ coming outs 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 action. In robustness checks, we control for more distant lags of exposure to LGBTQ coming outs to alleviate the concern that our results are driven by strong pre-action ties with LGBTQ peers, and not by exposure to peer coming out. To further address this concern, we estimate the Cox model for users who participated in the first wave of the action where we control for their exposure to coming outs from the second wave of the action (coming outs of LGBTQ users who did not participate in the first wave).[2]

We cannot directly test whether estimated peer effects are driven by changing perception of cost of the coming out because we cannot measure these perceptions. Instead, we seek to rule out the most important alternative mechanism, which is the spread of the information about the existence of the action through networks. Recently, the information channel has been shown to be important for

[2]While coming out decisions during the second wave may be influenced by coming outs that occurred during the first wave, these effects should be similar regardless of the hour in which peers joined the first wave of the action.

the diffusion of collective action (Garcıa-Jimeno et al., 2022). To rule out this channel, we control for the participation of non-LGBTQ users (anti-LGBTQ and straight ally peers). This allows us to separate the effects of exposure to LGBTQ coming outs from the effects of observing posts with the *IamLGBT* hashtag. Information channel should lead to high peer effects at the early stage of the action when the information about the action is limited to peers of users who started the action. Hence, we analyze the variation of the effects over time to further address this concern. Finally, we can compare the size of the effects before and after the major Polish newspaper posted a tweet about the action. The newspaper published the tweet about the action 15 hours after it started, when the *IamLGBT* hashtag had already been the top trending hashtag for several hours. At this point, almost all of Polish Twitter should have known about the existence of the action.

6 Results

Coming outs in users' networks had a significant effect on individual coming out decision: they substantially increased the probability of participation in the Twitter action by LGBTQ users (see Table 1). One standard deviation increase in the intensity of LGBTQ coming outs in user's network increased the chance of coming out by 13%. The estimates of the effect remain stable after controlling for gender, measures of Twitter activity, and network variables. Hence, the effects are not driven by prior Twitter behavior, strong connections to media, politics or LGBTQ activists accounts. We observe that the effects of coming out of peers are somewhat stronger for peers of the same gender than for peers of a different gender (Figure 3). This is in line with the idea that individuals update their beliefs about the costs of coming out by observing peers' decision to disclose their identity because coming outs of peers of individual's own gender provide a better information about the level of discrimination an individual is going to face after the coming out. Finally, we find that peer effects are driven solely by mutual relationships (Figure 4).

Figure 5 shows that observing posts from anti-LGBTQ peers had no significant influence on the individual decision to come out. We find that posts from straight allies increased the instantaneous probability of coming out but this effect was much weaker than the effect of observing LGBTQ peers' coming out posts. This difference suggests that that the estimated effects are not simply due to spreading the information about the existence of the action through networks.

The share of LGBTQ peers in users' networks is the major unobservable characteristic that could be correlated with the exposure to LGBTQ coming outs. Since having more LGBTQ peers may increase the probability of coming out, our results would be upward biased. We control for more distant lags of the exposure to LGBTQ coming out to address this issue. Table 2 shows that more distant lags of the exposure have no significant impact on the probability of coming out, and our main estimates remain stable. Hence, our estimated peer effects are driven by coming outs of peers in the hour immediately preceding the coming out decision, and not by constantly higher exposure to coming outs due to the high share of LGBTQ peers in users' networks. In an alternative approach, we show that controlling for the fraction of the network that who came out during the second wave of the action does not change the estimated peer effects (Table C.7).

Robustness and heterogeneity

The effects remain virtually the same after including only individuals who did post at least one tweet during the action (Table C.6). The estimated effects are statistically significant but smaller when we include only users who came out during the first wave of the action (Table C.7). The results are not driven by the strong networks of public figures: effects do not change after excluding journalists, elected officials and political party members from the sample (Table C.8). We estimate the hazard model with log network variables to account for their skewed distribution, and the effects are equally strong (Table C.4). The estimation of parametric survival models yields similar results (Tables C.2-C.3). Finally, the estimated effects are not sensitive to the choice of the sample restriction thresholds (see Figure C.4).

We find no significant variation in the size of the peers LGBTQ coming out effect depending on gender (see Figure C.1). Peer effects are stronger for users who mostly post original tweets than for users who mostly engage in discussions under other users' posts (Figure C.3). Finally, the effects are similar for users with low-to-medium tweeting frequency but are small and statistically insignificant for most active users (Figure C.2).

Effects over time

We further study the variation in the size of the network effects by estimating linear probability models of joining the action at 6-hour intervals (conditional on not participating in the action until

hour $t-6$). Figure 6 shows that that there was considerable variation in the magnitude of peer effects over time. At the beginning of the action, the role of peer effects was limited. This is surprising and suggests that the spread of information about the existence of the action through networks was not the mechanism driving the action. We find large and significant effects during the second, and third days of the action. At this point, the knowledge about the action was widespread, as it was covered by major media outlets. We also see that the size of peer effects decreases during nighttime hours. This can be explained in two ways. First, nighttime hours were characterized by much lower activity (see Figure B.2). If the effect of observing a peer coming was decreasing over time, small peer effects during nighttime hours could be driven by the low number of coming out tweets during these hours. Second, it is possible that users who were encouraged by their peers' coming out wanted to see their peers' reaction to their own coming out, so they postponed this decision until hours when more users are on Twitter.

We show that the results are not affected by the choice of the interval length (see Figures C.5 and C.6). The upper bound estimates of the peers effects using the network variable that includes all peers coming outs that occurred during the 6-hour interval are not far from the baseline effects (Figure C.7). Importantly, these results show that the peer effects were small and statistically insignificant also during the first 6 hours of the action once again suggesting that the effects are not driven by the mechanism of spreading the awareness about the action through networks.

7 Conclusions

The surge in visibility of LGBTQ individuals in recent decades remains largely unexplained. Studying LGBTQ coming out decisions is extremely difficult due to the limited availability of suitable data. In our paper, we investigate the hypothesis that observing peer coming out increases the probability that an individual will make a decision to disclose their LGBTQ identity. We argue that this is because peer coming out actions reduce the perceived costs of LGBTQ identity disclosure. We collected data from two large spontaneous Twitter coming out actions in Poland. The data provided a unique opportunity to empirically test our hypothesis. We found significant peer effects and no significant variation in the size of the effect depending on gender or tweets frequency. We also showed that the size of the effects differ depending on the gender of the peer with individuals responding stronger to coming outs of peers of the same gender than peers of a

different gender. Thus, we believe that the estimated results can be attributed to the mechanisms of correcting beliefs about the costs of coming out.

We used information about the exact time of the coming out to avoid the reflection problem. We find that the peer effects were small during early phase of the action suggesting that our results are not driven by the mechanism of spreading information about the action through networks. Moreover, this also partly alleviate concerns about the unobserved characteristics, as we show that our results are not driven by users who had no doubts about the consequences and joined the action already at the beginning. Instead, the effects are driven by users who decided to disclose their LGBTQ identity more than 12 hours after the beginning of the action.

We believe that our results have important policy implications. Encouraging coming out is a good strategy for the LGBTQ movement because it has a snowball effect. Nevertheless, our findings suggest caution as to the effects of coming outs by public figures. Estimated effects are driven entirely by mutual relationships: we found no effects of coming outs among the idol-followers type of relationships. There are two explanations for this fact. First, it is possible that LGBTQ followers do not update their perceived costs of discrimination after the coming out of an idol because it is not a good signal for what they could expect when they come out. Second, it is possible that the effects of idol's coming out are not immediate, in contrast to coming outs of close peers.

We end with potential avenues for future research. Future work could assess the external validity of our results. We analyzed coming out actions on an online platform in the context of a society with high levels of anti-LGBTQ sentiment. It would be interesting to see whether peer effects are equally strong when disclosing LGBTQ identity to family members or colleagues at work, and whether peer decisions matter in a more liberal society. In this paper, we did not examine the effects of coming out. Although existing studies suggest that coming out improves individual wellbeing, rigorous causal evidence is scarce. The impact of experiencing coming out of a close peer on the behavior and attitudes of non-LGBTQ individuals is another venue for future research, and the work by Fernández et al. (2021) suggests that these effects may be substantial.

References

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.

Badgett, Lee, Christopher S Carpenter, and Dario Sansone (2021). "LGBTQ economics". In: *Journal of Economic Perspectives* 35.2, pp. 141–70.

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.

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.

Bursztyn, Leonardo, Georgy Egorov, Ruben Enikolopov, and Maria Petrova (2019). "Social media and xenophobia: evidence from Russia". In: *NBER 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.

Campbell, Travis, Lee Badgett, and Everest Brennan (2021). "Beyond the Gender Binary: Transgender Labor Force Status in the United States 2014-2017". In: *Working Paper*.

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.

Chen, Shuai and Jan C van Ours (2020). "Symbolism matters: The effect of same-sex marriage legalization on partnership stability". In: *Journal of Economic Behavior & Organization* 178, pp. 44–58. DOI: 10.1016/j.jebo.2020.07.021.

— (2021). "Mental health effects of same-sex marriage legalization". In: *Health Economics*. DOI: 10.1002/hec.4441.

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.

Daron Acemoglu Munther A. Dahleh, Ilan Lobel and Asuman Ozdaglar (2011). "Bayesian Learning in Social Networks". In: *Review of Economic Studies* 78 (4), pp. 1201–1236.

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 (2021). "Coming Out in America: Thirty Years of Cultural Change". In: *Working Paper*.

Frankel, David M., Stephen Morris, and Ady Pauzner (2003). "Equilibrium selection in global games with strategic complementarities". In: *Journal of Economic Theory* 108.1, pp. 1–44. ISSN: 0022-0531. DOI: https://doi.org/10.1016/S0022-0531(02)00018-2. URL: http://www.sciencedirect.com/science/article/pii/S0022053102000182.

Garcıa-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.

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.

Giuliano, Paola (2020). "Gender and culture". In: *Oxford Review of Economic Policy* 36.4, pp. 944–961. DOI: 10.1093/oxrep/graa044.

Golub, Benjamin and Matthew Jackson (2012). "How Homophily Affects the Speed of Learning and Best-Response Dynamics". In: *The Quarterly Journal of Economics* 127.3, pp. 1287–1338. URL: https://EconPapers.repec.org/RePEc:oup:qjecon:v:127:y:2012:i:3:p:1287-1338.

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.

Jackson, Matthew O. and Yves Zenou (2015). "Chapter 3 - Games on Networks". In: ed. by H. Peyton Young and Shmuel Zamir. Vol. 4. Handbook of Game Theory with Economic Applications. Elsevier, pp. 95–163. DOI: https://doi.org/10.1016/B978-0-444-53766-9.00003-3. URL: https://www.sciencedirect.com/science/article/pii/B9780444537669000033.

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.

Martí, Joan de and Yves Zenou (2015). "Network games with incomplete information". In: *Journal of Mathematical Economics* 61, pp. 221–240. ISSN: 0304-4068. DOI: https://doi.org/10.1016/j.jmateco.2015.10.002. URL: http://www.sciencedirect.com/science/article/pii/S0304406815001202.

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.

Oyama, Daisuke and Satoru Takahashi (2020). "Generalized Belief Operator and Robustness in Binary-Action Supermodular Games". In: *Econometrica* 88.2, pp. 693–726.

Plug, Erik, Dinand Webbink, and Nick Martin (2014). "Sexual orientation, prejudice, and segregation". In: *Journal of Labor Economics* 32.1, pp. 123–159. DOI: 10.1086/673315.

Sansone, Dario (2019). "Pink work: Same-sex marriage, employment and discrimination". In: *Journal of Public Economics* 180, p. 104086. DOI: 10.1016/j.jpubeco.2019.104086.

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

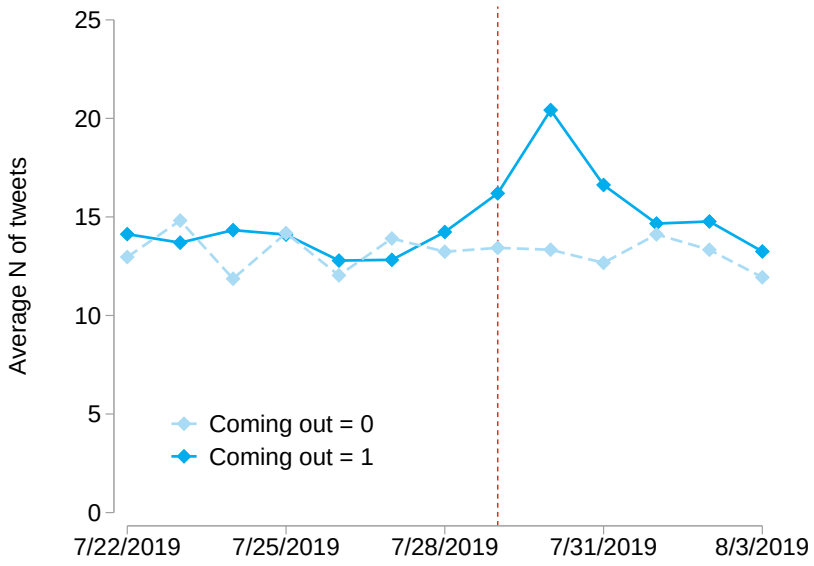


Figure 1: Twitter activity by coming out decision

Notes: Figure shows daily average number of tweets for two groups of users: those who joined the first action and those who did not join the action in the period from 22 July 2019 to 3 August 2019.

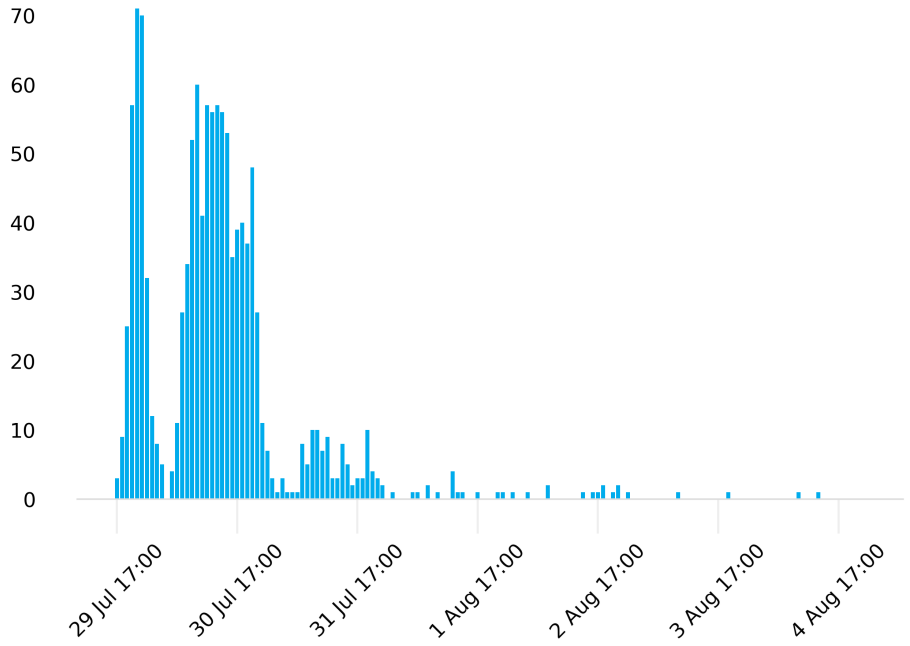


Figure 2: The development of the first LGBTQ coming out action

Notes: Figure displays the number of LGBTQ coming outs per hour from 29 July 17:00 to 5 August 00:00.

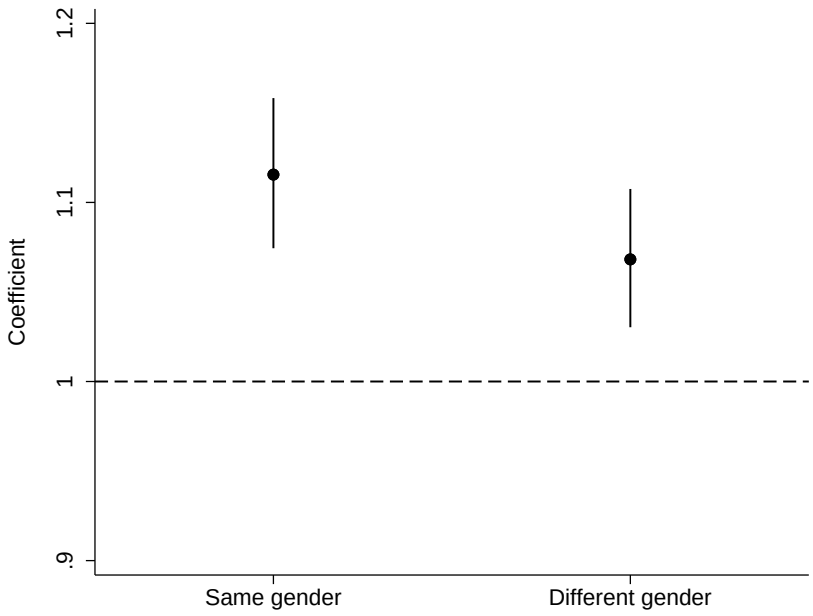


Figure 3: Peer effects and probability of coming out: LGBTQ peers' gender

Notes: Figure shows the estimates of a Cox proportional hazards model. The coefficients are reported in terms of proportional hazards (exponentiated coefficients). The 'Same gender' coefficient shows the effect of LGBTQ coming out of peers of the same gender as individual's own gender. The 'Different gender' coefficient shows the effect of LGBTQ coming out of peers of a different gender than individual's own gender. In the regression, we control for gender, Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

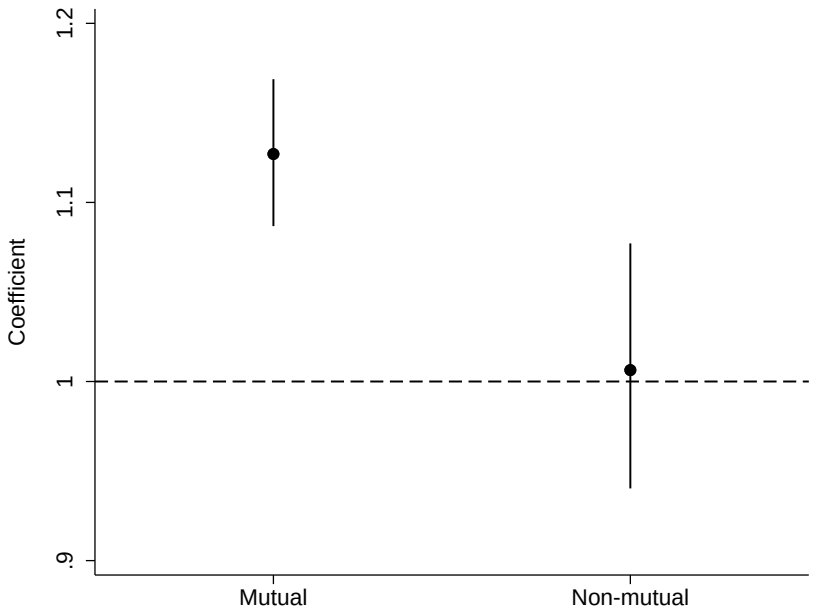


Figure 4: Peer effects and probability of coming out: mutual vs. non-mutual relationships

Notes: Figure shows exponentiated coefficients from a Cox proportional hazards model estimation of the effects of peers' LGBTQ coming out depending on the type of the peer relationship (mutual vs. non-mutual). The 'Mutual' coefficient shows the effects of LGBTQ coming out of peers who had a mutual relationship with the user (there was at least one reply tweet from the peer to the user). The 'Non-mutual' coefficient shows the effects of LGBTQ coming out of peers who had never posted a reply tweet to the user. In the regression, we control for gender, Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

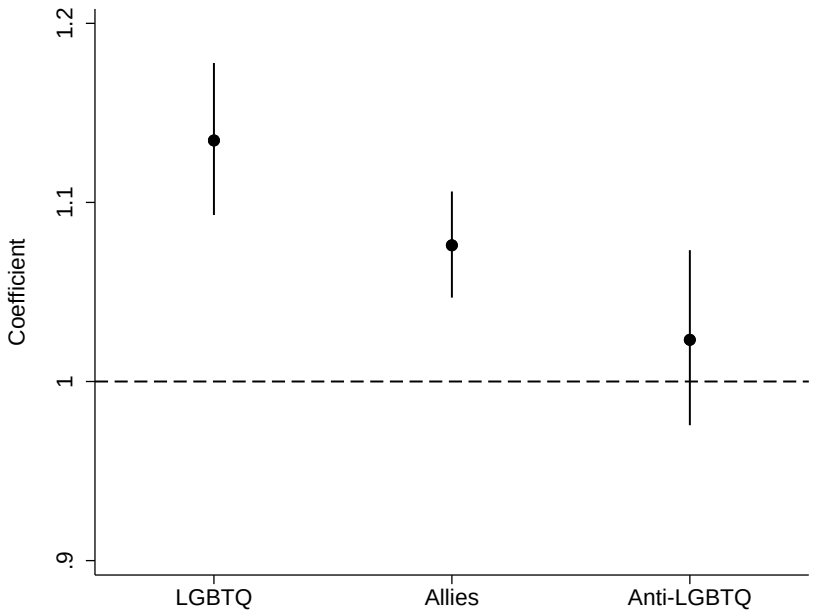


Figure 5: Peer effects and probability of coming out: other user types

Notes: Figure shows exponentiated coefficients from a Cox proportional hazards model estimation of the effects of peers' participation in the Twitter action on probability of coming out. LGBTQ network measures the fraction of the network that came out as an LGBTQ person by a given hour. Allies network is the relative intensity of replies to allies who joined the action by a given hour. Anti-LGBTQ network is the relative intensity of replies to anti-LGBTQ users who joined the action by a given hour. For each hour, network variables are standardized with zero mean and standard deviation of one. In the regression, we control for gender, Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

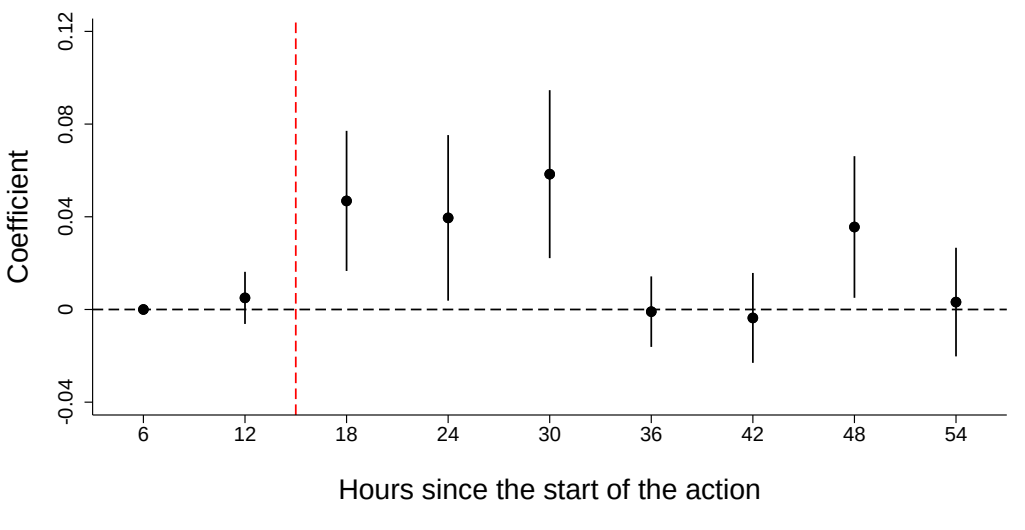


Figure 6: Peer effects and probability of coming out over time

Notes: Figure shows coefficients from a linear probability model estimation of the effects of LGBTQ peers' participation in the Twitter action on probability of coming out within 6 hours before given hour. We exclude users who joined the action by hour $t-6$. LGBTQ network is the relative intensity of replies to LGBTQ users who joined the action by hour $t-6$ (standardized for each hour with zero mean and standard deviation of one). We control for gender, Twitter activity, and network variables. 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 action.

Tables

Table 1: Peer effects and and probability of coming out

	(1)	(2)	(3)	(4)
Network: LGBTQ	1.132***	1.129***	1.132***	1.135***
coming outs	(0.022)	(0.022)	(0.021)	(0.022)
Gender	no	yes	yes	yes
Twitter activity	no	no	yes	yes
Network	no	no	no	yes
Observations	35104	35104	35104	35104

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 that came out as an LGBTQ person before a given hour. For each hour, network variables are standardized with zero mean and standard deviation of one. We control for gender (female, male, transgender / non-binary), measures of Twitter activity (log tweets count, average tweet length, hashtag use, emoji use, replies as % of all tweets), and network characteristics (log network size, replies to media, politics, LGBTQ activist accounts, intensity of straight allies, and anti-LGBTQ users in the network). We use standard errors clustered at the user level.
* $p < .10$; ** $p < .05$; *** $p < .01$

Table 2: Peer effects and and probability of coming out: controlling for lagged exposure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Network: LGBTQ	1.156***	1.233***	1.193***	1.154***	1.136***	1.132***	1.232***
coming outs	(0.027)	(0.063)	(0.053)	(0.048)	(0.043)	(0.040)	(0.063)
Network: LGBTQ		0.931					0.927
coming outs (t-2)		(0.047)					(0.074)
Network: LGBTQ			0.963				0.893
coming outs (t-3)			(0.042)				(0.088)
Network: LGBTQ				1.002			1.060
coming outs (t-4)				(0.039)			(0.081)
Network: LGBTQ					1.024		1.026
coming outs (t-5)					(0.037)		(0.065)
Network: LGBTQ						1.031	1.047
coming outs (t-6)						(0.035)	(0.051)
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
Observations	31157	31157	31157	31157	31157	31157	31157

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 that came out as an LGBTQ person before a given hour. For each hour, network variables are standardized with zero mean and standard deviation of one. Depending on specification, we control for more distant lags of the network variable. In all regressions, we control for gender (female, male, transgender / non-binary), measures of Twitter activity (log tweets count, average tweet length, hashtag use, emoji use, replies as % of all tweets), and network characteristics (log network size, replies to media, politics, LGBTQ activist accounts, intensity of straight allies, and anti-LGBTQ users in the network). Our sample includes observations from the 7th to the 54th hour of the action in order to be able to estimate more distant lags. We use standard errors clustered at the user level.
* $p < .10$; ** $p < .05$; *** $p < .01$

Appendix A Data collection

Lists of tweets with #IamLGBT hashtag

The list of user who participated in the first wave of the Twitter action was downloaded on 17 October 2019 (2 months after the action). It was downloaded using the *GetOldTweets3* library. The library scrapes the Twitter search results of a given term. We used the following term: "#jestemlgbt since:2019-07-29 until:2019-08-01". This query returns all tweets posted between 2019-07-29 00:00:00 - 2019-08-01 00:00:00 that included the hashtag. On 13 July 2021, we additionally scraped tweets with #jestemLGBT hashtag from 1 August - 4 August (using *snscrape* library because *GetOldTweets3* became obsolete). Initially, we skipped those 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 action (see Figure B.2). Nevertheless, we decided to add those users to make sure that we do not mistakenly classify them as "not yet out" individuals.

The list of users who participated in the second wave of the Twitter action was downloaded first on 28 May 2020 (1 day after the start of the action), and then extended for users that participated in the period from 28-29 May on 30 May 2020. We used the *GetOldTweets3* library and modified search filters to include the period of interest.

Classification of user types and genders

After downloading the list of tweets, we manually coded two variables: user type and gender. There were 4 types of users: LGBTQ coming outs, non-LGBTQ allies, anti-LGBTQ haters, and others. LGBTQ coming outs were detected based on two patterns. First, some users added an information about their sexual orientation or gender identity explicitly in the tweet (e.g., they mentioned they are bisexual or that they have a partner of the same gender). These direct statements were sometimes captioned on attached photos. Second, some users wrote "#IamLGBT and I am" finishing the sentence with a description of their occupation or personal characteristics.

Allies were detected based on the supportive statement that included an explicit declaration that they are not LGBTQ persons themself (most common pattern included "nie #jestemLGBT ale"

phrase which means "I am not LGBT but" followed by a supportive statement.

Anti-LGBTQ haters were detected in several ways. Some tweets included explicitly homophobic or transphobic hate speech. Some tweets included obscene photos or photos of totalitarian leaders. Some trolling tweets included photos of famous politicians and their fake coming out statements. Some posts criticized the action for sharing private information (eg., "there is nothing to be proud of"), accused foreign forces of staging the action, or made other negative remarks about the action.

The remaining tweets are classified as "others". This category includes tweets that had no additional text except the hashtag, tweets spreading the news about the action, tweets in which users did not disclose their LGBTQ or non-LGBTQ identity, and tweets on unrelated topics (using the popularity of the hashtag to get attention). We manually detected accounts of organizations (media, NGOs, European Commission) that posted tweets so we can account for that but we may have classified some post from LGBTQ individuals as "others". Nevertheless, we believe that the participation of LGBTQ individuals in the "#IamLGBT" action does not necessarily require them coming out. Hence, the tweets that only spread the news or express the support without disclosing the identity of the user are classified as "others", and not as LGBTQ coming outs.

Gender of users is classified based on their declarations and the text analysis of their tweets. First, we check whether users explicitly stated their gender in their coming out tweets. Second, we use the fact that verbs in Polish language are inflected for gender. Specifically, gender can be inferred from the ending of past tense verbs written from the first-person point of view (morphological endings also allow for distinguishing singular first-person verbs). The masculine verbs end with "łem" and feminine verbs end with "łam". The non-binary endings include "łxm" and "łom" although it is still quite rare, and non-binary individuals may also use one of the traditional forms. We compare the frequency of verbs with these endings to establish gender of individuals. In few cases where it is impossible to detect gender (users do not use first-person verb forms, they use them equally often), we classify their gender as "unknown".

Twitter activity data

Twitter activity data was downloaded using the *snscrape* library. For each user, we have downloaded up to 30,000 most recent tweets posted since 1 January 2019 until 1 November 2019. The user activity file includes following information for each posted tweet: unique tweet ID, time stamp, content, the number of like, the number of retweets, the number of replies, whether the tweet is a reply or an original tweet, screen name of the addressee of the reply, and unique account ID of the addressee of the reply (for reply tweets only). Twitter data was downloaded in August 2021. Initially, we have downloaded Twitter activity data for users that participated in the first action in November 2019, and for users who participated in the second action in June 2020. However, we have not downloaded the information about the account ID of the replies addressees (we downloaded only their screen names). Since screen names may be changed by users, our network variables could have underestimated the share of users who participated in the first action in 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 data downloaded once again for all users in August 2021 which includes information on the account ID that is consistent over time. The drawback of this approach is that we lose information about Twitter activity of users whose accounts were removed, suspended or made private. Twitter activity data may be also incomplete for users who removed some of their tweets from the studied period. We will document these differences using the data we downloaded initially. Importantly, the results of the analysis are virtually identical using the two data sources.

Appendix B Sources and descriptive statistics

Table B.1: Variable descriptions (i.)

Variable	Description	Source
R_i	a binary variable that equals 1 if the user came out during the first wave of the action, and 0 if the user came out during the second wave but not during the first wave	Twitter (hand coded)
Gender: female	a binary variable that equals 1 if the user uses female pronouns / verb forms and 0 otherwise	Twitter (hand coded)
Gender: male	a binary variable that equals 1 if the user uses male pronouns / verb forms and 0 otherwise	Twitter (hand coded)
Gender: transgender / non-binary	a binary variable that equals 1 if the user disclosed transgender or non-binary identity and 0 otherwise (hand coded)	
Tweets count	the count of tweets	Twitter
Average tweet length	average number of characters of tweets	Twitter
Hashtag use	the count of hashtags divided by the word count of all tweets	Twitter
Emoji use	the count of emojis divided by the word count of all tweets	Twitter
Replies (% of all tweets)	the count of replies divided by the count of all tweets	Twitter
Network size	the count of all users to which the user replied at least once (replies)	Twitter
Network: LGBTQ coming out	the number of replies to LGBTQ users who came out during the first wave of the action ($R_i = 1$) divided by the total count of replies	Twitter (LGBTQ user type dummy variable hand coded)
Network: allies	the number of replies to allies ('I am not LGBTQ but I support this action') who participated in the first wave of the action divided by the total count of replies	Twitter (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 action divided by the total count of replies	Twitter (anti-LGBTQ user type dummy variable hand coded)

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

Table B.2: Variable descriptions (ii.)

Variable	**Description**	**Source**
Network: media	the number of replies to journalists and news accounts (participants of the actions and those among the top 100 most popular accounts in the network) divided by the total count of replies	Twitter (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 actions and those among the top 100 most popular accounts in the network) divided by the total count of replies	Twitter (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 action (participants of the actions and those among the top 100 most popular accounts in the network) divided by the total count of replies	Twitter (LGBTQ activist account type dummy variable hand coded)

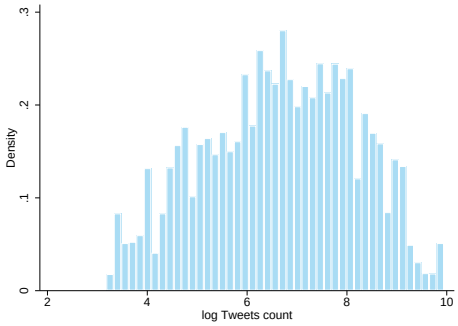
Notes: Description of variables used in the analysis. All variables except for R_i and 'Gender: transgender / non-binary' are based on Twitter activity before the first Twitter action (1 January 2019 - 28 July 2019). Network variables are based on replies during the pre-action period as well, while information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

Table B.3: Descriptive statistics

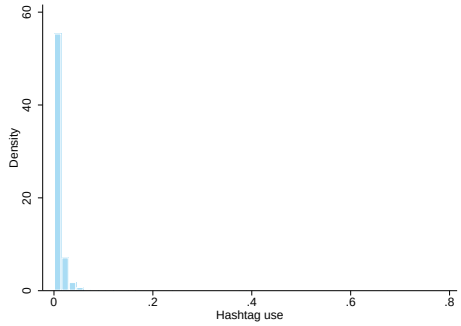
	Obs.	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Network: LGBT coming out	35104	-0.00	1.00	-0.67	22.67
Gender: female	35104	0.77	0.42	0.00	1.00
Gender: male	35104	0.20	0.40	0.00	1.00
Gender: transgender / non-binary	35104	0.03	0.16	0.00	1.00
log Tweets count	35104	6.65	1.52	3.18	9.95
Average tweet length	35104	10.36	4.44	2.03	37.04
Hashtag use	35104	0.01	0.03	0.00	0.67
Emoji use	35104	0.04	0.05	0.00	0.67
Replies (% of all tweets)	35104	0.46	0.20	0.02	1.00
Network: allies	35104	0.00	1.00	-0.61	30.96
Network: anti-LGBT	35104	0.00	1.00	-0.33	34.35
log Network size	35104	4.60	1.22	1.61	7.44
Network: media	35104	0.01	0.03	0.00	0.38
Network: politicians	35104	0.01	0.03	0.00	0.49
Network: LGBTQ activists	35104	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. All variables are measured in 2011 if not specified otherwise. The sources and description of the variables can be found in Table B.1. Network variables are standardized with zero mean and standard deviation of one.

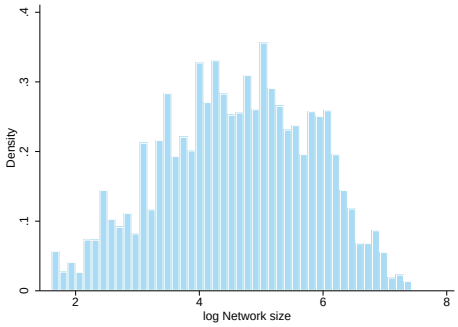
(a) log Tweets count



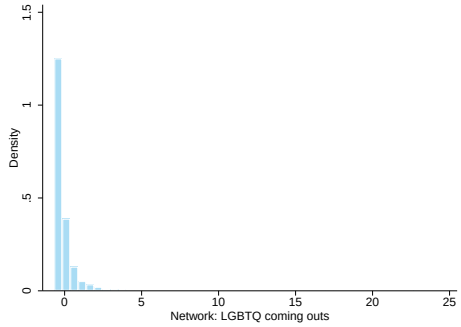
(b) Hashtag use



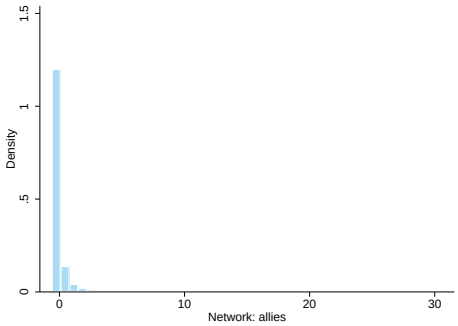
(c) log Network size



(d) Network: LGBTQ coming outs



(e) Network: allies



(f) Network: anti-LGBT

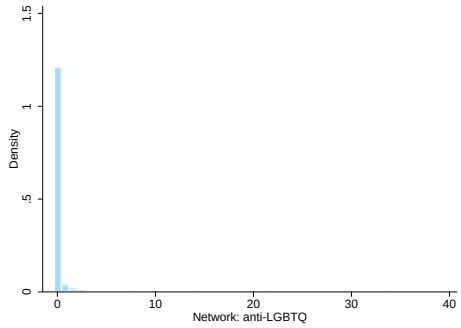


Figure B.1: Distribution of variables

Notes: Figure shows histograms of all variables.

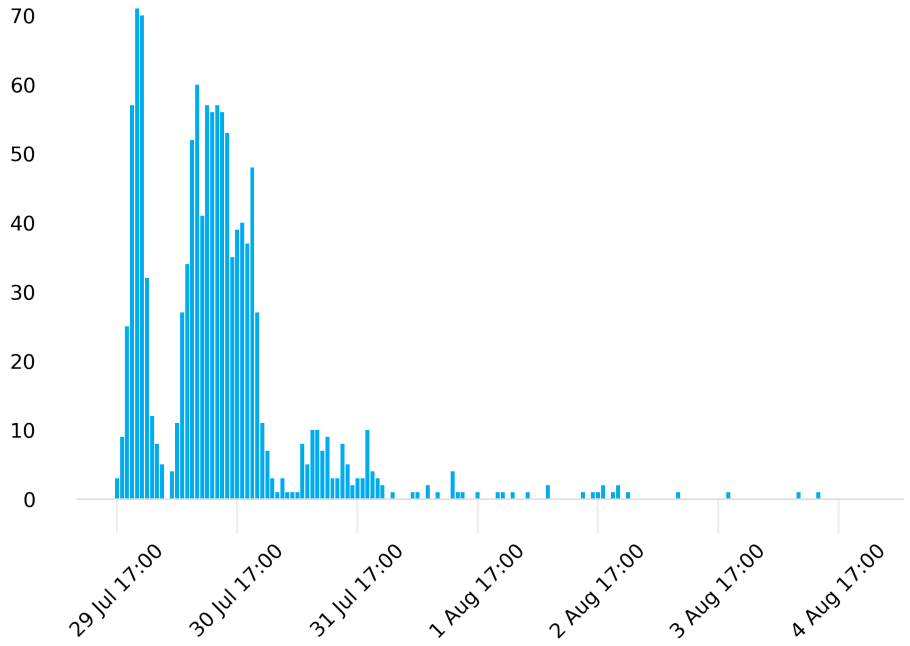


Figure B.2: The development of the first LGBTQ coming out action

Notes: Figure displays the number of LGBTQ coming outs per hour from 29 July 17:00 to 5 August 00:00.

Appendix C Additional results

Table C.1: Peer effects and probability of coming out: full results

	(1)	(2)	(3)	(4)
Network: LGBTQ	1.132***	1.129***	1.132***	1.135***
coming outs	(0.022)	(0.022)	(0.021)	(0.022)
Gender: female		0.888*	0.928	0.964
		(0.063)	(0.071)	(0.076)
Gender: transgender		0.866	0.829	0.822
/ non-binary		(0.167)	(0.155)	(0.165)
log Tweets count			1.099***	1.074
			(0.025)	(0.053)
Average tweet length			1.040***	1.036***
			(0.007)	(0.007)
Hashtag use			0.519	0.425
			(0.556)	(0.487)
Emoji use			1.044	1.205
			(0.789)	(0.900)
Replies (% of all			0.966	0.813
tweets)			(0.153)	(0.174)
Network: allies				1.076***
				(0.015)
Network: anti-LGBTQ				1.023
				(0.025)
log Network size				1.040
				(0.061)
Network: media				2.790
				(2.347)
Network: politicians				2.324
				(1.965)
Network: LGBTQ				1.456
activists				(2.800)
Observations	35104	35104	35104	35104

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 that came out as an LGBTQ person before a given hour. For each hour, network variables are standardized with zero mean and standard deviation of one. Control variables are described in the note of Table 1. All regressors are based on Twitter activity in the period between 1 January 2019 and 28 July 2019 (one day before the start of the action). We use standard errors clustered at the user level.

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

Table C.2: Peer effects and probability of coming out: parametric proportional hazards model

	(1)	(2)	(3)	(4)
Network: LGBTQ	1.126***	1.122***	1.124***	1.126***
coming outs	(0.024)	(0.025)	(0.023)	(0.024)
Gender	no	yes	yes	yes
Twitter activity	no	no	yes	yes
Network	no	no	no	yes
Observations	35104	35104	35104	35104

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 that came out as an LGBTQ person by a given hour. For each hour, network variables are standardized with zero mean and standard deviation of one. Control variables are described in the note of Table 1. We use standard errors clustered at the user level.

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

Table C.3: Peer effects and coming out time: parametric AFT model

	(1)	(2)	(3)	(4)
Network: LGBTQ	-0.102***	-0.100***	-0.098***	-0.106***
coming outs	(0.024)	(0.024)	(0.023)	(0.022)
Gender	no	yes	yes	yes
Twitter activity	no	no	yes	yes
Network	no	no	no	yes
Observations	35104	35104	35104	35104

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 that 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, network variables are standardized with zero mean and standard deviation of one. Control variables are described in the note of Table 1. We use standard errors clustered at the user level.

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

Table C.4: Peer effects and probability of coming out: log network variables

	(1)	(2)	(3)	(4)
Network: LGBTQ	1.186***	1.182***	1.178***	1.171***
coming outs	(0.031)	(0.031)	(0.031)	(0.033)
Gender	no	yes	yes	yes
Twitter activity	no	no	yes	yes
Network	no	no	no	yes
Observations	35104	35104	35104	35104

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 that came out as an LGBTQ person before a given hour. We take logs of the network variable after adding a small number (0.01). Then, for each hour, network variables are standardized with zero mean and standard deviation of one. Control variables are described in the note of Table 1. We use standard errors clustered at the user level.

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

Table C.5: Peer effects and probability of coming out: including second hour of the action

	(1)	(2)	(3)	(4)
Network: LGBTQ	1.133***	1.130***	1.133***	1.136***
coming outs	(0.021)	(0.021)	(0.020)	(0.021)
Gender	no	yes	yes	yes
Twitter activity	no	no	yes	yes
Network	no	no	no	yes
Observations	37915	37915	37915	37915

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 that came out as an LGBTQ person before a given hour. For each hour, network variables are standardized with zero mean and standard deviation of one. Compared to the baseline analysis, we include the second hour of the action. We control for variables described in the note of Table 1 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 action. All regressors are based on Twitter activity in the period between 1 January 2019 and 28 July 2019 (one day before the start of the action).
* $p < .10$; ** $p < .05$; *** $p < .01$

Table C.6: Peer effects and probability of coming out: only those active during the action

	(1)	(2)	(3)	(4)
Network: LGBTQ	1.136***	1.132***	1.132***	1.136***
coming outs	(0.020)	(0.021)	(0.020)	(0.021)
Gender	no	yes	yes	yes
Twitter activity	no	no	yes	yes
Network	no	no	no	yes
Observations	33042	33042	33042	33042

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 that came out as an LGBTQ person before a given hour. For each hour, network variables are standardized with zero mean and standard deviation of one. Control variables are described in the note of Table 1. The sample includes only users who were active during the first wave of the Twitter action (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 C.7: Peer effects and probability of coming out: excluding users who did not participate in the first action

	(1)	(2)	(3)	(4)	(5)
Network: LGBTQ	1.072***	1.072***	1.074***	1.072***	1.072***
coming outs	(0.023)	(0.023)	(0.023)	(0.024)	(0.024)
Gender	no	yes	yes	yes	yes
Twitter activity	no	no	yes	yes	yes
Network	no	no	no	yes	yes
Network: second wave	no	no	no	no	yes
Observations	18121	18121	18121	18121	18121

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

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

Table C.8: Peer effects and probability of coming out: excluding journalists, elected officials, and members of political parties

	(1)	(2)	(3)	(4)
Network: LGBTQ	1.134***	1.132***	1.134***	1.134***
coming outs	(0.022)	(0.022)	(0.021)	(0.022)
Gender	no	yes	yes	yes
Twitter activity	no	no	yes	yes
Network	no	no	no	yes
Observations	34547	34547	34547	34547

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 that came out as an LGBTQ person before a given hour. For each hour, network variables are standardized with zero mean and standard deviation of one. Control variables are described in the note of Table 1. All regressors are based on Twitter activity in the period between 1 January 2019 and 28 July 2019 (one day before the start of the action). We exclude journalists, elected officials, and political party members from the sample. We use standard errors clustered at the user level.

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

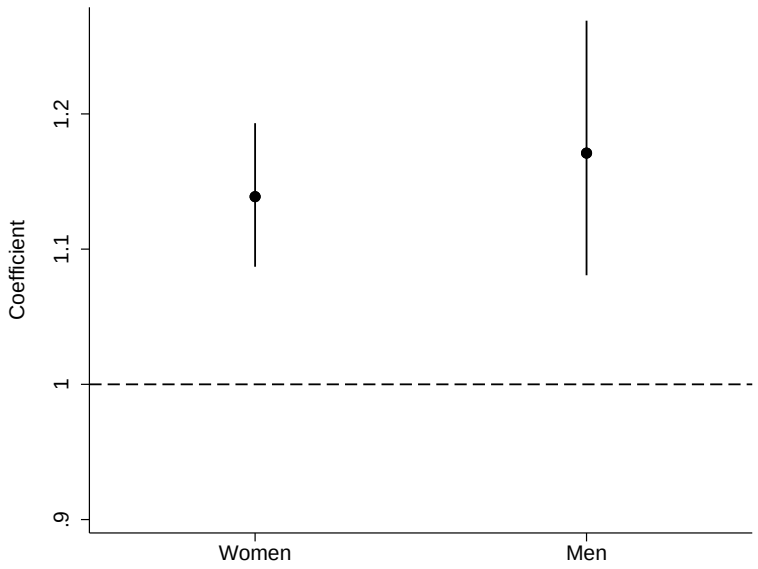


Figure C.1: Heterogeneity of the effect: gender

Notes: Figure shows exponentiated coefficients from a Cox proportional hazards model estimation of the effects of peers' LGBTQ coming out in the Twitter action 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 4 users who did not come out in the first wave in the sample). For each hour, network variables are standardized with zero mean and standard deviation of one. In all regressions, we control for gender, Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

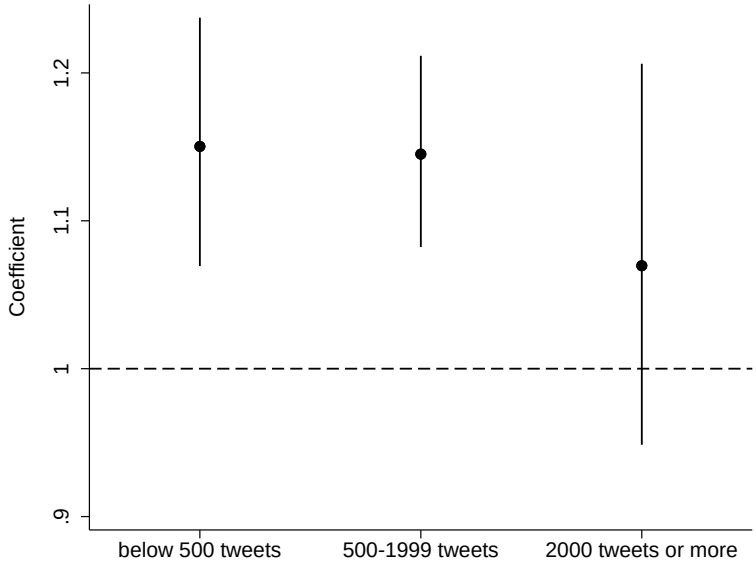


Figure C.2: Heterogeneity of the effect: Twitter activity level

Notes: Figure shows exponentiated coefficients from a Cox proportional hazards model estimation of the effects of peers' LGBTQ coming out in the Twitter action on the time of coming out for three groups of users based on the number of tweets they posted between 1 January 2019 and 28 July 2019. For each hour, network variables are standardized with zero mean and standard deviation of one. In all regressions, we control for gender, Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

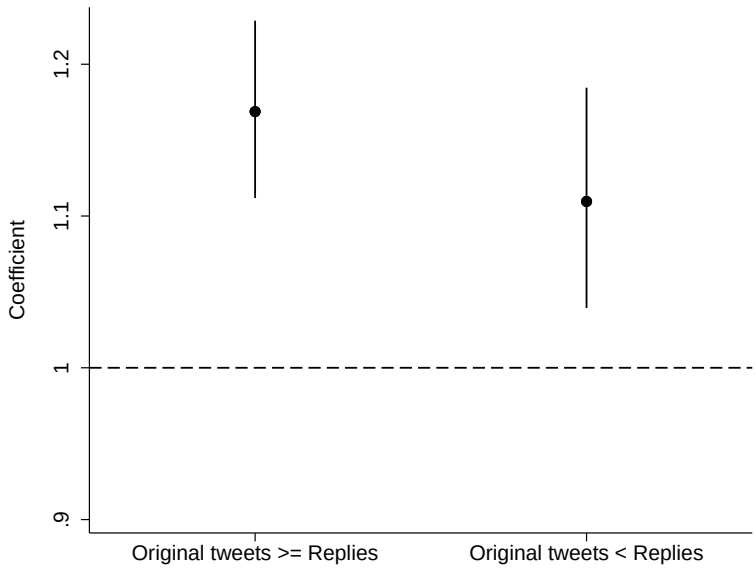
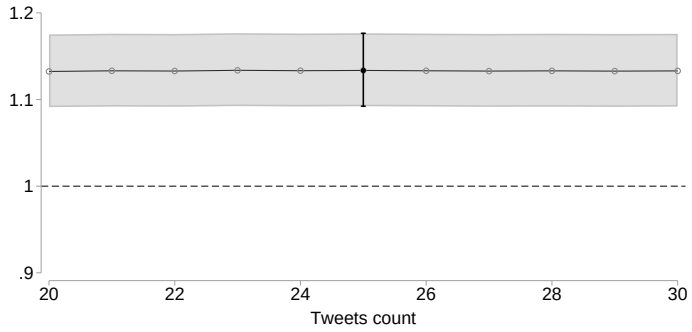


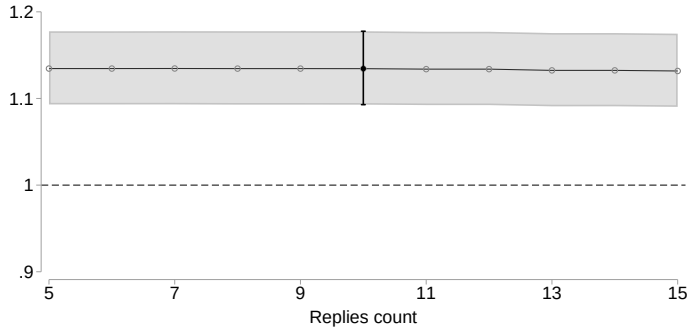
Figure C.3: Heterogeneity of the effect: reply behavior

Notes: Figure shows exponentiated coefficients from a Cox proportional hazards model estimation of the effects of peers' LGBTQ coming out in the Twitter action on the time of coming out for two groups: users with the reply share of tweets smaller or equal 0.5 and users with the reply share of tweets greater than 0.5. For each hour, network variables are standardized with zero mean and standard deviation of one. In all regressions, we control for gender, Twitter activity, and network variables. 95% confidence intervals are constructed based on standard errors clustered at the user level.

(a) Tweets count



(b) Replies count



(c) Network size

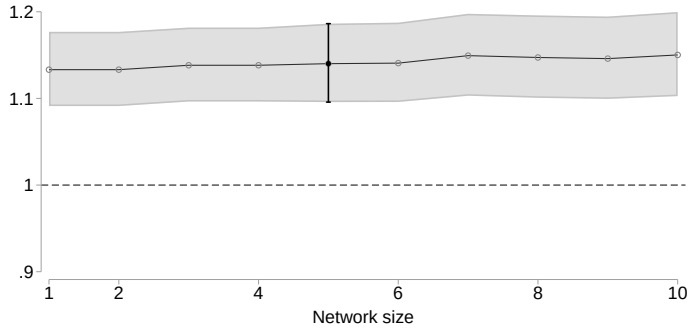


Figure C.4: Sensitivity of the results to changes in sample restrictions

Notes: Figure shows coefficients from Cox proportional hazards models of the effects of peers' LGBTQ coming out in the Twitter action on the hazard of coming out for varying sample restrictions. The coefficients are reported in terms of proportional hazards (exponentiated coefficients). In Figure C.4a, we show the results with varying cutoffs of the minimum number of tweets during the pre-treatment period (1 January 2019 - 28 July 2019). In Figure C.4b, we show the results with varying cutoffs of the minimum number of replies during the pre-treatment period. In Figure C.4c, we show the results with varying cutoffs of the minimum number of network size (the number of unique users to which the user replied at least once) during the pre-treatment period. For each hour, network variables are standardized with zero mean and standard deviation of one. In all regressions, we control for gender, measures of Twitter activity, and network characteristics. 95% confidence intervals are constructed based on standard errors clustered at the user level.

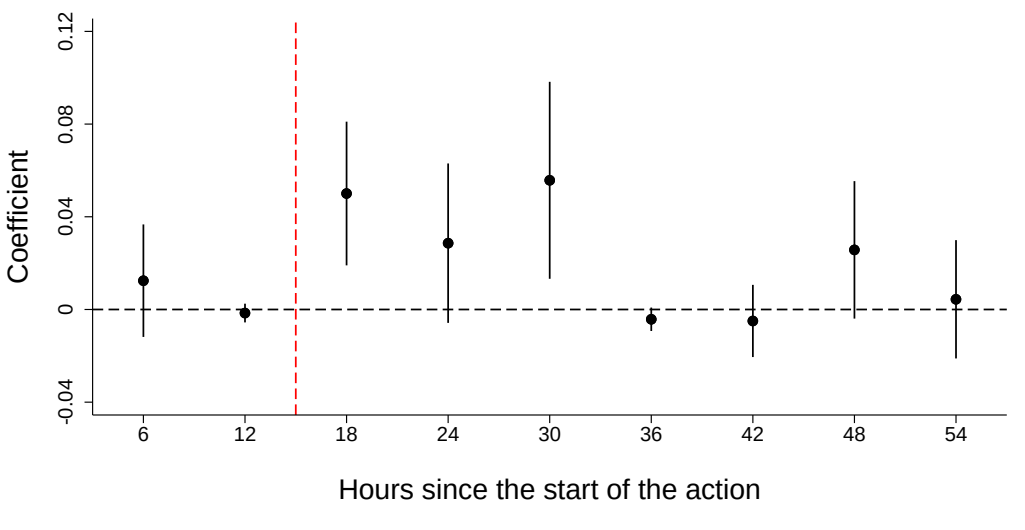


Figure C.5: Peer effects over time: 4-hour intervals

Notes: Figure shows coefficients from a linear probability model estimation of the effects of LGBTQ peers' participation in the Twitter action on probability of coming out within 4 hours before given hour. We exclude users who joined the action by hour $t-4$. LGBTQ network is the relative intensity of replies to LGBTQ users who joined the action by hour $t-4$ (standardized for each hour with zero mean and standard deviation of one). We control for gender, Twitter activity, and network variables. 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 action.

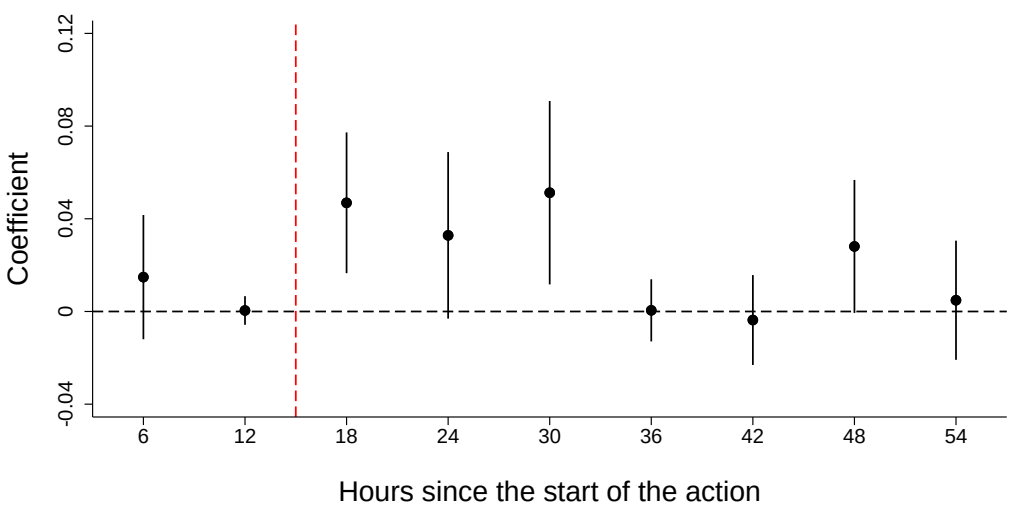


Figure C.6: Peer effects over time: 5-hour intervals

Notes: Figure shows coefficients from a linear probability model estimation of the effects of LGBTQ peers' participation in the Twitter action on probability of coming out within 5 hours before given hour. We exclude users who joined the action by hour $t-5$. LGBTQ network is the relative intensity of replies to LGBTQ users who joined the action by hour $t-5$ (standardized for each hour with zero mean and standard deviation of one). We control for gender, Twitter activity, and network variables. 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 action.

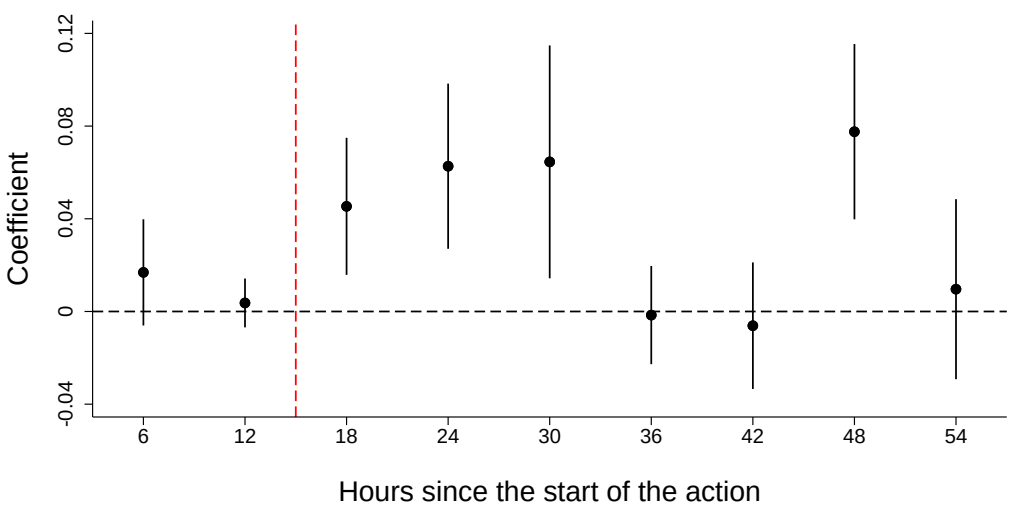


Figure C.7: Peer effects over time: network variable that includes coming outs of peers during the 6-hour interval

Notes: Figure shows coefficients from a linear probability model estimation of the effects of LGBTQ peers' participation in the Twitter action on probability of coming out within 6 hours before given hour. We exclude users who joined the action by hour $t - 6$. LGBTQ network is the relative intensity of replies to LGBTQ users who joined the action by hour t (standardized for each hour with zero mean and standard deviation of one). Hence, the network variable includes coming outs of peers who came out immediately after the user, and the estimated effects should be treated as an upper bound of the peers effects. We control for gender, Twitter activity, and network variables. 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 action.

Appendix D Examples of tweets

I'm LGBT and I'm pansexual, not that I'm attracted to the 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 how long. I'm not an ideologist. I love and I feel. And unfortunately, I also increasingly feel fear and helplessness.

my name is vincent and I am LGBT I'm no paedophile, I'm only 16 years old and 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've had enough.

I'm very supportive of the whole action 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, they won't 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.