The University of Chicago

Private Matters: The Effect of Data Privacy Laws on State-Sponsored Foreign Disinformation Campaigns

By Elisa Bayoumi

July 2021

A paper submitted in partial fulfillment of the requirements for the Master of Arts degree in the Master of Arts Program in the Committee on International Relations

Faculty Advisor: Rochelle Terman

Preceptor: Gentry Jenkins

Section I: Introduction

Since 2015, state-sponsored disinformation campaigns have plagued social media and targeted thousands of citizens around the world. These campaigns are highly coordinated efforts by foreign governments to spread partial truths or complete falsehoods about ongoing issues relevant to the politics of other states (Fallis 2015; Tucker et. al. 2018). These campaigns have two unique features: first, they are propagated over social media, and second, they are produced by an actor affiliated with a foreign government (Martin et. al. 2019; Nemr and Gangware 2019; Woolley and Howard 2018). The literature has thus far focused on discovering strategies actors use to spread disinformation (Bulger and Davidson 2018; Martin et. al. 2019; Allcott and Gentzkow 2017). One strategy for spreading disinformation is the microtargeting strategy. First outlined in the Cambridge Analytica scandal, the microtargeting strategy begins with a foreign actor buying data on foreign citizenry. Using this data, the state or a third party can craft a specific strategy to target the cognitive biases of individual social media users and voters. Voters are helpless to combat disinformation campaigns tailored to convince them of falsehoods, and ultimately believe the disinformation and share it with their followers. The tactic has been dominant on Facebook and Twitter because the platforms process personal data on the interests and habits of their users to allow for better targeting of advertisements. This practice was so persistent on Twitter that the platform had to ban all political advertisements from its site in the run-up to the 2020 U.S. election (Romm and Stanley-Becker 2019).

Is it possible to impede microtargeted foreign disinformation campaigns with data privacy laws? Data privacy laws restrict collecting, storing, and disseminating data about a user's online habits and interests. These laws would impact the source of the microtargeting strategy by

preventing the sale or transfer of data to foreign actors or companies. However, these laws remain controversial because they impede the regular use of data for trade and advertising.

In this paper, I argue that data privacy laws negatively affect foreign disinformation campaigns as they are prevented from using data to customize their campaigns. I test the effect of data privacy laws on disinformation campaigns by investigating the effect of the European Union's General Data Protection Regulation (GDPR) on Russian-linked actors targeting the United Kingdom (UK) with disinformation. This paper finds strong evidence that the implementation of GDPR coincides with a change in how frequently the Russian accounts spread disinformation relating to the UK on Twitter. Using Twitter's public archive of Russia's Internet Research Agency (IRA) accounts, I find a significant decrease in the number of tweets per day that mention the United Kingdom after the implementation of GDPR compared to the five months preceding GDPR. I hypothesize that this decrease in daily tweets indicates a temporary pause on the disinformation campaign targeted at British citizens as the data required for the microtargeting campaign became harder to get. This finding is consistent despite the confounding variable of a large news story that caught the attention of the Russian trolls and caused an increase in the total number of tweets per day. Critically, I demonstrate that the implementation deadline for GDPR correlates with an interruption in the behavior of Russian accounts spreading disinformation targeted at the UK, suggesting that data privacy laws may indeed impact foreign disinformation campaigns.

Understanding how foreign actors spread disinformation is particularly relevant given both the increasingly common incident of disinformation and its threat to democratic countries. Experts estimate that 28 countries experienced some form of disinformation campaign in 2017; 48 countries in 2018; and at least 70 countries in 2019 (Bradshaw and Howard 2019). While only

7 countries are estimated to have sophisticated enough tools to conduct a disinformation campaign targeting another state, the number of countries subject to such campaigns has been rising (Bradshaw and Howard 2019). Meanwhile, these campaigns have serious legitimacy costs for their target state. First, foreign actors manipulating foreign citizens infringes on the right of governmental sovereignty. Without significant efforts to curb and punish foreign disinformation campaigns, states will see opportunities to meddle in other countries and affect citizens they have no control over. Second, disinformation campaigns are incredibly problematic for the legitimacy of democratic governments and elections. Because democratic governments are given the right to rule by an informed citizenry, the prospect of wrongly informed or even manipulated populations electing officials questions the legitimacy of the entire government and elections system. Finally, foreign disinformation can destroy the trust of the citizenry in having access to reliable information. This destruction of trust in institutions and truthfulness harms the population overall by undercutting the trust of essential institutions such as journalism or government. As one witness in an investigation into disinformation efforts said, “It doesn’t really matter whether [the news] is fake or not; what matters is that people think it’s fake. That really damages their trust in [democratic] institutions.” (Defence Committee, 2019a). While many have struggled to identify the exact incentive for foreign disinformation campaigns, they pose a fundamental threat to many essential Western institutions, calling for deeper analysis into what limits their spread or effectiveness.

The rest of the paper precedes as follows. Section II reviews the definitions and literature for this paper, as well as identifying two explanations of how foreign actors spread disinformation: the microtargeting explanation and the manipulation explanation. While these are not the only explanations for how states conduct foreign disinformation campaigns, they are

the only ones relevant for this paper's argument and case study. Section III explains the main argument of this paper and describes the hypotheses. Section IV lays out case study selection and describes the source selection. Section V discusses the data selection and operationalization of the event study. Section VI discusses the results of the event study, and Section VI discusses these results and addresses their implications and possibilities for future research.

Section II: Literature Review

Given the recent work of online foreign disinformation campaigns, one portion of the literature on disinformation concerns the exact definition of what constitutes disinformation. This philosophical debate is critical for expressing the nuances of different levels of truthfulness in misinformation and disinformation, but it is not pertinent to the question of whether data privacy laws affect disinformation campaigns. Therefore, for this paper, disinformation is defined as political propaganda that is either partial truth or complete falsehood about ongoing issues relevant to the politics of targeted democratic states (Fallis 2015; Tucker et. al. 2018). This definition is sufficiently broad to include misinformation – partial truths which are presented in a false or misleading context – and completely false claims (See Wardle and Derakhshan 2018). The data in this case study has been selected to circumvent substantial debates regarding whether the data is disinformation or not, and whether it is part of a foreign disinformation campaign or not.[1]

Another part of the literature asks why online users would believe foreign disinformation campaigns in the first place. This literature describes the social and cognitive conditions that make social media uniquely positioned to encourage the spread of disinformation. Foreign disinformation campaigns involve foreign interference in creating or distributing the

[1] For further discussion of defining disinformation, see Tandoc Jr. et. al. 2017, Gelfert 2019, and Karlova and Lee 2012.

misinformation, but only work because the social media platform encourages users to believe this shared disinformation. I therefore must address the research on why online users believe and spread disinformation before addressing strategies behind foreign disinformation campaigns.

One critical element of explaining why online users believe fake and real news online is the enormous amount of information on the internet. This plethora of news articles and providers creates an environment that inhibits users from critically accurately assessing real or fake news, making users more prone to believing disinformation for two reasons (Allcott and Gentzkow 2017).

The first reason that the increase in news leads to more people believing fake news is that it overwhelms online users with news reports and notifications. Users feel bombarded with news and are too exhausted to consistently use their critical thinking skills to evaluate truthfulness (La Garde and Hudgins, 2018). Known as "fake news fatigue", online users may not think clearly about a piece of news, or use the cognitive shortcuts highlighted below in place of critical thinking (Bulger and Davidson 2018). This results in users sharing fake news articles simply because they are too tired to check if the information is completely way true, particularly because reposting the information is as easy as hitting a simple "share" button (Pennycook and Rand 2019).

A second consequence of having more news outlets online is the appearance of news reports representing all sides of the political spectrum, enabling confirmation bias. Confirmation bias in this context means users are more likely to believe the fake news that supports their political stance once they encounter it on the internet (Nickerson 1998). Some analysis suggests this bias comes from an online user's utility from confirming their beliefs about the world rather than challenging them. Others suggest it is because fake news may be presented as an opinion

rather than a factual issue, thereby allowing the reader to believe that the facts are a matter of opinion or up for debate (Prior et. al. 2016). Overall, the increase in news from all sides of the political spectrum has decreased the objectivity of the user when reading the news, making it more likely that the readers believe the fake news, even if it's from foreign governments.

There are several other important cognitive shortcuts at play when online users believe fake news. Pennycook and Rand investigate several of these in their 2017 article. Recruiting volunteers to evaluate the veracity of randomly generated sentences, they found that their subjects tended to be overly accepting of weak claims, particularly when given limited information. For example, the presence or absence of a source's headline did not affect the abilities of those in the study to determine the veracity of the headline, suggesting they evaluated the headline itself instead of the argument. Pennycook and Rand also discovered several other cognitive shortcuts. They first found that the more the subjects were exposed to a false claim, the more likely they were to believe it. The researchers also found evidence that the more their readers claimed false familiarity or expertise in certain concepts, the more likely they were to misjudge statement veracity, Finally, they found substantial evidence that readers who believed in other false information, such as vaccines causing autism, were significantly more likely to believe false headlines (Pennycook and Rand 2017). The result of their research is a thorough and disheartening confirmation of the numerous cognitive fallacies that online users employ to believe false news, making it increasingly likely that users would believe and share fake news on social media.

These cognitive shortcuts manipulated on social media are only part of the reason users believe fake news. The other commonly cited reason online users believe fake news is that the structure of social media produces echo chambers and exacerbates cognitive shortcuts. Echo

chambers are online spaces where one is exposed to opinions that only agree with and reflect their own (Garimella et. al. 2018). On social media, creating an echo chamber is particularly easy due to the user's ability to follow – or unfollow – anyone. Confirmation bias suggests that most people would select to follow other users with whom they agree. Another element producing echo chambers is the social media algorithms to recommend the same preferences and viewpoints to their users. The algorithm particularly customizes to a user's political preferences, recommending and promoting content associated with the user's partisan beliefs and not that of the other political perspective. While social media companies defend this practice by arguing it provides a more user-friendly experience, this practice causes users to only receive news from one perspective and remain completely oblivious about alternative interpretations or arguments.

These echo chambers exacerbate many of the cognitive shortcuts previously mentioned, including the confirmation bias, repetition bias of seeing news reiterated multiple times, and false familiarity with a subject. For those who already believe in theories of questionable merit, the algorithm will recommend further fake news conspiracies. Research also demonstrates that others caught in the same echo chamber are more likely to trust the links from friends and family without analyzing the actual veracity of the news, spreading fake news messages like a contagion (Tornberg et. al. 2018). Garimella et. al. find that fighting this echo chamber phenomenon is largely hopeless, as users who attempt to hear from both sides of the spectrum pay a "price of bipartisanship" and are significantly less important in the social network as well as promoted less than those who play the partisan game (Garimella et. al. 2018). This dangerous combination of cognitive shortcuts and echo chamber has already been used to explain why a fake tweet calling the Toronto shooting suspect "angry" and "middle Eastern" went viral while the tweet calling the suspect "white" by the same account remained under the environment (Meserole 2018). When

stuck in an echo chamber, online users do not need foreign meddling to believe fake news as it enters their feed (Menczer 2016).

While online users might be more likely to believe disinformation due to their cognitive shortcuts and echo-chambers, this does not amount to a coordinated disinformation campaign on social media. A disinformation campaign is a centrally organized group that manipulates the cognitive shortcuts and echo chambers on social media to achieve a goal. State-sponsored disinformation campaigns are particularly sophisticated, as they have the support and resources of an entire country to assist them (Lazer et. al. 2017). A state-sponsored foreign disinformation campaign uses the features of social media to enhance its strategy to spread disinformation, but they develop a more sophisticated strategy to spread their disinformation.

If social media is so attuned to spreading disinformation, why would foreign governments need to go to such lengths to buy data or use bots and trolls to spread disinformation?[2] It is because these sophisticated strategies make the campaign more effective in its messaging and finding its target audience. Many disinformation campaigns are not about sharing *any* disinformation, but *specific* disinformation which benefits the foreign policy goals of the state sponsoring the campaign. Therefore, foreign actors must actively create and spread their disinformation rather than watching unaffiliated trolls spread misinformation. Second, it is still difficult to convince most citizens of disinformation, despite the cognitive shortcuts or are in echo chambers. It therefore makes much more sense to tailor specific disinformation campaigns to those most likely to believe the campaign rather than producing an ineffective campaign that is largely ignored. For example, foreign actors want to hide their foreign associations both from

[2] A "bot" is defined as an autonomous program that independently executes certain tasks on the internet. A "troll" is a person who deliberately misrepresents their viewpoints online, typically in an offensive manner. In the context of social media and disinformation campaigns, both bots are programed on behalf of the state conducting the disinformation campaign, and trolls are hired by the same state. These are just two examples of accounts ways that foreign disinformation campaigns conduct online accounts

would-be believers of disinformation and from the international community. Citizens are particularly suspicious of information from a foreign source, especially if that source is not a known ally or a state that does not have a good reputation for truthfulness. Other actors, like social media companies and governments, might also become suspicious of the increased sharing of political content from a foreign government if seemingly everyday users like bots did not share the information. Foreign disinformation campaigns therefore require more advanced tactics for sharing disinformation, two of which are addressed in this paper.

II.A: the Microtargeting Strategy

The microtargeting strategy is where foreign actors get personal data on the citizens of the target state and using such data to identify vulnerable citizens that they should target with their disinformation campaign. This strategy was first identified in the Cambridge Analytica scandal, a company revealed that in 2018 the Russian government had been interested in Facebook data on U.S. citizens to learn about their political beliefs and sway the 2016 U.S. Presidential election. Christopher Wylie's book outlines Cambridge Analytica's work collecting information and tailoring products with the express intention of influencing a foreign citizen's political decisions. Leveraging data from social media as well as other sources, Wylie suggests his former company conducted "psychological warfare" to influence the users in the way that other actors desired. Based on these accounts, the overall process of how the microtargeting explanation spreads disinformation to the target state's users is simple. First, foreign actors access increasingly personal data on their target state's citizens to learn more about decision-making patterns and beliefs. These users are then sorted into categories based on their likes and their likelihood of engaging with and believing the selected disinformation. Finally, the foreign actor buys ads that targeted these specific groups or accounts with strategies designed to exploit

their thought patterns and make the user believe the disinformation. As social companies have started cracking down on foreign governments buying political ads, evidence suggests bots also act on this microtargeting explanation by connecting with users who are most likely to believe the disinformation and bring them into the echo chamber of disinformation (Tornberg 2018). Foreign actors are therefore able to view electoral politics as marketing opportunities, and craft narratives and advertisements perfectly tailored to get individuals to believe and share the news.

This tactic of spreading fake news is only possible because personal data is so affordable and unregulated. The European Data Protection Supervisor pointed out that fake news is a symptom of "concentrated, unaccountable digital markets, and constant tracking and reckless handling of personal data" (Claesson 2019). Efforts to stop or interrupt the microtargeting strategy have so far been mixed. One way to combat microtargeting is to target the actual vectors used to spread the disinformation: either foreign governments buying ads or foreign bots and trolls (Daskal 2019). Another tactic has been educating the public about online threats and how to spot campaigns like the ones by Cambridge Analytica. Authors have argued that the actual content of the educational campaign has not been effective, but the publicizing of the online disinformation has made the public suspicious of ads and bots online. This argument suggests that no further policies are necessary, as the window of opportunity for a successful microtargeted campaign has passed (Lankosa 2019). Further, few countries have prevented the beginning of the microtargeting strategy by preventing foreign governments from buying personal data on their citizens.

However, the issue with these solutions is that they have led to a cat and mouse game over developing bots and hiding troll accounts to fool social media companies and citizens (Lazer et. al. 2018). In recent years, bots have developed much faster than most detection

software, education programs, or even public suspicion has been able to detect. Fake accounts have also disguised many of the classic elements people look for when assessing a fake account (Kessler 2014). While some of these strategies have limited certain foreign disinformation campaigns, the microtargeted campaigns have continued as the software for spreading disinformation has gotten more sophisticated.

II.B: Manipulating the Algorithm

The other explanation of how foreign actors spread disinformation to their target state's users is that foreign agents manipulate the social media algorithm to call attention to their disinformation stories. While there may be other ways to manipulate the algorithm, the most obvious and frequently cited examples are harnessing several bot and troll accounts to post a lot of content and manipulate conversation patterns. Because social media algorithms promote trending topics with a lot of engagement in a short amount of time, the social media platform begins to mention the disinformation topic or hashtag to more and more account holders. The critical difference between the manipulation and the non-manipulation explanation is that the manipulation explanation argues that accounts created by foreign states spread trending issues, whereas the latter only argues that non-manipulated users from the target state start the conversation on their own. The two most common examples of manipulating the algorithm are astroturfing and hashtag hijacking.

In an astroturfing campaign, a foreign state will use numerous fake accounts to create the perception of a grassroots movement online. These accounts may repeat the same hashtag, word, or even full comment to create the impression of multiple users sharing the same sentiment. Beyond creating unique posts, fake accounts may repost or pad the number of interactions with a user's content to make it seem "viral". To complete the perception of a "popular" account, bots

will finally follow and retweet the accounts that promote what the target state desires (Keller et. al. 2019a).

Slightly different from astroturfing, hashtag hijacking is a tactic where bots and trolls will respond to a trending hashtag and coopt it with disinformation messages. While this is much more common in activist circles or autocracies responding to domestic political issues, it is occasionally used in foreign disinformation campaigns as well (Jackson and Welles 2015). The Internet Research Agency (IRA) backed by the Russian government, for example, used bots and trolls to infiltrate the popular #blacklivesmatter, #alllivesmatter, and #oscarssowhite conversations in the United States (Keller et. al. 2019b). Further research will be to determine how much of a role these foreign-linked accounts play in hashtag hijacking compared to domestic accounts.

There are some substantial weaknesses to the use of bots and trolls to share disinformation. First, the crux of this strategy involves a relatively simple complication of having a massive number of bots; however social media companies and others are actively trying to find and ban bothersome bots from the websites (See Ratkiewicz et. al. 2011). Astroturfing and hashtag hijacking therefore becomes a dangerous game when the overall number of bot and troll accounts must seem real enough in their messages but produce enough content to have an actual impact. This numbers-game often leads to obvious mistakes; for example, Keller et. al. find that some IRA-backed accounts tweeted out the same messages, though they were supposedly from accounts on the opposite sides of the political spectrum (Keller et. al. 2019). The manipulation tactic of spreading misinformation is therefore a difficult struggle of developing sufficiently complex algorithms for several bots to seem “normal” without supervision at a large scale, all while the technology to discover bots becomes more complex.

Another question plaguing the argument that disinformation is spread through manipulating social media algorithms is the difficulty of creating viral conversations. Keller et. al. find that the incidences of observed astroturfing are increasing over time, with 20% of observed foreign disinformation campaigns involving some form of hashtag hijacking (Keller et. al. 2019a: 20). Despite their increase in popularity, there is evidence that the bots and trolls have a hard time making their topics go viral (Zannettou et. al. 2019). Yet on other occasions, including for the U.S. in 2016, bots achieved retweet networks for 4 million tweets and had a measurable influence (Woolley and Guilbeault 2018). Individually, bots and trolls also struggle to influence online conversations. Howard and Kollanyi's study of bots in the Brexit campaign suggests that bots have a small overall impact online. Even though technology has developed to where most online users sincerely struggle with identifying bots compared to people, their research found that bots are typically reposting other user's content, generating a massive amount of the posts on a polarizing issue like Brexit without adding much original content or disinformation (Howard and Kollanyi 2016). When a bot or troll account is generating its own content, research shows that a massive amount of personal authenticity and cultural competence is required before sharing disinformation is relatively effective (Xia et. al. 2018). Furthermore, there is limited evidence that the information trolls share reaches a substantial number of Twitter users, suggesting trolls are overall bad at spreading disinformation to other users (Zannettou et. al. 2019).

Indeed, successful astroturfing and hashtag hijacking attempts are complicated, and it is impossible to claim that this manipulation tactic is always effective. A more likely explanation is that users in the target state need to be better tailored rather than releasing too much undirected

disinformation into the online space. This involves careful crafting of disinformation content, or the use of microtargeting data to tailor messages towards their intended audience.

Section III: The Argument

The literature thus far has focused on the discovery and analysis of tactics to spread disinformation, particularly microtargeting. This paper builds on the work of Arnaudo, Kornbluh, and Claesson to argue that data privacy laws are essential to combat foreign disinformation campaigns because they stop the targeted spread of misinformation and therefore limit the campaign's effectiveness. The passage of a data privacy law might not completely stop the targeted campaign online; however, it will interrupt the existing tactics of spreading disinformation and cause a notable change in the strategy of spreading disinformation. Most likely, the interruption will lead to a change in the online behavior for accounts sharing disinformation. The accounts may either decrease their posts or increase their posts, only doing the latter if they switch from a microtargeting strategy to a more generalized manipulation strategy.

While the term data privacy laws typically refer to any sort of law concerned with the handling of data on individuals, it here refers to a subset of laws concerned with the handling of data regarding the online activities and habits of individuals. This narrowed definition ensures that the data privacy laws will apply to the microtargeting technique for spreading disinformation, as the technique requires data about an individual's online activities and behaviors. For example, a common example of a data privacy law in the U.S. is the Health Insurance Portability and Accessibility Act (HIPAA); however, the medical data the law protects would probably not help for targeting online users with misinformation and is therefore not a relevant data privacy law to this paper. The exact way that this data about online habits and

activities are regulated is not as important to the question at hand. Data privacy laws relevant to this paper's argument can restrict the acquisition of the data, processing of the data, selling or transferring of the data, or storage of the data. If the law overall restricts the number of people encountering the data on people's online activities, habits, and interests, the data privacy law applies to my argument.

With this argument in mind, there are three possible outcomes to the passage of the data privacy law. In the first case, there is no change in behavior caused by the implementation of a data privacy law. This could happen for two main reasons. First, the foreign disinformation campaign might not use the microtargeting strategy that involves buying data and instead tailor their online campaign with more traditional marketing or analog tactics. Data privacy laws would therefore not affect the non-microtargeted campaign, and the campaign itself would remain unchanged. Second, the data privacy law might not affect how the foreign state gets the data on users abroad. The term "data privacy" covers such a broad number of regulations and laws, many of which are not specific to combatting foreign disinformation campaigns. For example, a data privacy law that restricts the storage of personal data for a certain amount of time might give the foreign state sufficient time to know which citizens to target, even though they delete the data later. It is also not clear beyond the Cambridge Analytica case how foreign actors get the personal data to target their disinformation campaign, so a data privacy law might target the wrong part of the data collection process. For example, a law restricting social media companies from selling to foreign states overlooks the third-party companies that sell personal data on citizens or ignore that certain states may collect the private data they can access from public social media accounts. Ultimately, there are several reasons that data privacy laws might not

affect the tactics of foreign disinformation campaigns, but the most likely explanation is that the specific type of data privacy law does not work against that campaign.

Hypothesis 1: Data privacy laws will not affect foreign disinformation campaigns. This is either because the foreign disinformation campaign does not use microtargeting, or because the data privacy law does not impact the supply chain of data for the microtargeting campaign.

The second and third scenarios build on my argument that the data privacy law affect foreign disinformation campaigns, particularly as they have continued to use microtargeting techniques to spread disinformation. The two scenarios differ in how the foreign state chooses to adapt their campaign to the new world of limited data for microtargeting disinformation; therefore, the passage of the data privacy law would coincide with different online behavior changes in the foreign disinformation campaign. In the first case, the state in charge of the campaign abandons its efforts to influence foreign citizens until an alternative data source is discovered. Instead of adapting to the new situation, the state decides to not expend its resources continuing an ineffective campaign based on manipulating a social media algorithm, so it puts the campaign on pause until it can continue the microtargeting strategy. The microtargeting strategy would only be continued if a state were able to find an efficient and effective alternative data source on foreign citizens. Alternative sources of data could be extracting information from less regulated sources, like metadata and other big data. Ultimately, this second scenario has the foreign disinformation campaign reemerge after pausing their operations, meaning that the disinformation campaign will involve fewer overall online interactions until such activities can be better targeted.

Hypothesis 2: Data privacy laws weaken foreign disinformation campaigns in the short term, forcing the source state to pause the campaign while they search for alternative data sources.

In the third scenario, a state attempts to adapt the disinformation campaign to the new world of data privacy. In doing so, they abandon their previous tactic of microtargeting for an alternative tactic of spread, likely manipulation tactics. This change likely involves the reactivation of bot and troll accounts or changes in their sharing and posting behavior to meet the new requirements for manipulating the algorithm. This adopting an untargeted disinformation campaign could be combined with the second scenario, meaning that the manipulation tactic is only adopted until another data source is found, upon which the campaign returns to their microtargeted campaign. However, this third scenario is distinct because of the change in tactics of spreading disinformation, compared to the pause in the disinformation campaign in the second scenario.

Hypothesis 3: Data privacy laws affect foreign disinformation campaigns, forcing the source state to adopt an alternative tactic of spreading disinformation.

Section IV: Case selection

Having presented the three hypotheses for this paper, I will summarize the case selection process for this study.

The case selection process is limited by the lack of modern data privacy laws related both to the internet and online social media activity, as well as the selling of data to a third party or even foreign states. States with particularly stringent data privacy laws have had such laws in place for a long time, making it even harder to test for a behavior change. Other states have not

passed data privacy laws that protect information in the digital age. The one obvious exception to this rule is the European Union's GDPR.

GDPR is a set of regulations passed by the European Union in 2015 to protect the data privacy of European citizens. The bill concerned all electronic processing of personal data which emerges or is linked to the European Union's trade or market. Personal data is defined as any data that can directly or indirectly identify a specific person from which it is sourced, which is data particularly pertinent to microtargeting (European Commission, GDPR). The regulations attempt to govern the processing of data by six main principles of privacy. Though it does not ban either the storage or processing of data, it has clear requirements for the collection, storage, and dissemination of personal data. Because the regulations emerged out of growing concern of Silicon Valley's disregard for data privacy, the bill also focuses on international transfers of personal data using third parties, building off the European Court of Justice's rulings in *Google v. Spain* and *Facebook v. Ireland* (Albrecht 2016; GDPR Article 45). With serious fines and sanctions for companies or states violating the terms of the regulation, many European states and companies rushed to implement the rules by its implementation on May 24th, 2018. The passage of the regulation remained incredibly controversial, as corporations raised concerns that such measures would interrupt their business practices and privacy experts worried that the regulations did not go far enough (See Richards 2020; Zarsky 2017).

Though not initially created to curb disinformation campaigns, several regulations of GDPR could interrupt the collection of data for microtargeted disinformation campaigns. The first regulation to interrupt microtargeted disinformation campaigns could be the requirement of user consent and control for the collection and processing of personal data online. GDPR empowers users to limit the collection and spread of their online data and could have empowered

enough users to restrict the collection and sharing of their data that only a small number of users can be microtargeted for disinformation. However, the dark patterns on the web mean that very few users successfully restrict the collection and processing of their personal data, making it unlikely that the consent controls from GDPR had much impact.[3]

Alternatively, the data anonymity requirements could be restricting the ability of foreign disinformation campaigns to discover and target specific users prone to believing their campaigns. As personal data has such stringent requirements under GDPR, many companies have taken to "anonymizing" their data or removing personal indicators from the processed data. As anonymized data is not covered by GDPR, this manipulation of the data is more freely available for processing and transfer than personal data. While a debate rages about what truly constitutes "anonymized" data and whether data can truly be anonymized, the general implication of this principle is simple: the foreign disinformation campaign only has access to anonymized or generalized data. While there might be a lot of personal data to be retrieved, even with the removal of personal identifiers, the identification of persons and users online is much more complex and can take a long time compared to un-anonymized data. Therefore, foreign disinformation campaigns may struggle to identify specific people or groups of people from this anonymized data, slowing down the overall campaign until the data can be sufficiently processed to find a workaround.

Finally, GDPR is the most stringent towards third-party processing facilities, like those used in the Cambridge Analytica scandal. Companies that store, process, and sell data must now

[3] Dark patterns, first defined by Grey et. al., are manipulations of a website or interface to get the user to make a decision that is less preferred by said user but more preferred by the website shareholders. Many dark patterns involve visual changes to design elements, such as shrinking the font of an "unsubscribe" button at the end of an email or making the color of the button to close a pop-up tab too light to notice at first. While not illegal, these manipulative tactics often work against the spirit of GDPR's consent requirements and allow websites to continue collecting data when the user is unaware of another option. Nouwen et. al.'s work suggests that these patterns have increased despite GDPR, and that regulation of such patterns remains elusive.

seek approval from the EU that its goals are "necessary for the purposes of the legitimate interests" of the company that collected the data, and does not violate the fundamental rights of privacy guaranteed to all EU citizens. This regulation results in a lot of bureaucracy related to involving third parties in processing most personal data, and particularly affects the sale of data relating to EU citizens to foreign actors. Many of the companies that might have previously helped foreign disinformation campaigns collect information for microtargeting are subject to increased scrutiny and struggle to produce the same data they did previously.

While these are the most obvious regulations relating to foreign disinformation campaigns, many other mechanisms might be impacted by the largest data privacy law passed by a state with few previous online data regulations. As any one of these GDPR regulations could impact foreign disinformation campaigns and affect them in the ways outlined in the hypothesis, GDPR is the obvious case to be examined for this question.

Given that GDPR is the most obvious data privacy law for this paper, the question remains of which foreign disinformation campaign should be selected – that is, which campaign from which state targeting which other state. I have selected the Russian disinformation campaign influencing the UK. Of the states perpetuating foreign disinformation campaigns, one of the most well-documented campaigns against the EU has come from Russia. After disclosing Russia's involvement with the 2016 US presidential election, several states in Europe revealed that they too had been the subject of Russia's online disinformation efforts. Russia particularly targeted Western states experiencing political turmoil, and the most obvious campaign was against the United Kingdom. Though only revealed in late 2017, the British government disclosed that the Russians had been involved in influencing discussions related to the EU membership referendum in 2016, commonly referred to as "Brexit", and possibly as early as the

2014 Referendum on Scottish Independence (Booth Et. Al. 2017, Richards 2020). These campaigns continued as the debate about Brexit and other political issues continued through the implementation of GDPR as political strife – as indicated by three general elections in the years since the referendum – continue to plague national actors.[4] Several international data mining companies – including Cambridge Analytica and a Canadian company called Aggregate AQI – created targeted messages to encourage traditionally reluctant voters to vote for more politically risky endeavors (Richards 2020). While there is no credible evidence that these companies were employed or aligned with Russian actors, their work suggests how prevalent microtargeting tactics were towards the British public, and how easily the Russians could have adopted such tactics. Given that both the referendum and the elections were the result of a popular vote, Russia had increased incentives to target a disinformation campaign at UK citizens and influence their preferences to ascertain their preferred outcome from the votes.

The decision to studying a Russian disinformation campaign has several practical benefits. Firstly, it is a relatively well-documented, simple campaign that provides some of the best data we have on foreign disinformation campaigns against EU states. Most states do not wish to publicize the specifics of a foreign disinformation campaign for fear that it would teach other states how to conduct campaigns. However, information has been released on how Russia conducted its early disinformation campaigns and the actual content of said disinformation campaign. This is a combined result of sufficient safeguards being put in place since the campaign was conducted, and the increased demand for information and data for reports and

[4] For further evidence that the UK has been subject to a disinformation campaign from Russia well beyond the initial Brexit election, see Martin et. al. 2019,, Bradshaw and Howard 2018, Bradshaw and Howard 2019, and Vériter et. al.2020.

research into foreign disinformation campaigns.[5] The case selection of a disinformation campaign against the UK is also a good choice as it is not only one of the best-documented IRA campaigns, but also solves a practical issue of understanding the content of the tweet, as it is in English.

Much of Russia's online disinformation campaign emerges from its Internet Research Agency, or IRA. Reports about the IRA vary but have several key features. First, agents at the IRA oversee the making and producing online content to share disinformation with larger and larger followings. These IRA-controlled accounts can be either troll or bot accounts that tweet frequently and unusually compared to a real person's social media account. Many IRA accounts try to disguise their true intentions by talking about banal topics or attempting to increase regular social interaction. However, they erratically veer into political discussions designed to help the Kremlin achieve its goals. These political posts need not be consistent or logically coherent and often involve a variety of posts about a diverse set of political topics, some of which conflict with previously stated political beliefs (Dawson and Innes 2019). The most successful IRA accounts reach hundreds of thousands of followers and are even featured on popular humor websites before their suspicious politics and erratic behaviors come to light (see Popken 2017).

The final issue regarding case selection is what online activity from which social media platform should be used to measure the effects of the data privacy law on the disinformation campaign. For the case of the Russian disinformation campaign against the UK, which was mostly conducted on social media websites like Facebook, Reddit, and YouTube, I follow the literature in selecting Twitter. Twitter is the most used social media platform for studying

[5] This observation about increased safeguards does not mean that foreign disinformation campaigns will never take place again. Instead, they will be more complex in certain elements – such as having a more complex bot or having more complex data acquisition tactics – to work around these protections. This is an element of the cat-and-mouse game against disinformation campaigns.

disinformation campaigns due to its practical and simple data on online behaviors and interactions. Twitter has the most straightforward way to interact with other users, which is limited to tweets, retweets, responses, and direct messaging. Most of the tweets, retweets, and responses from an account are public or can become public and collects lots of data including the time of the post, content of the post, and users mentioned in the post. Twitter was also one of the most popular websites for foreign disinformation campaigns. This is likely due to its relatively relaxed account-creation process and the platform's popularity as one of the most used social media platforms in the world (Smith and Anderson 2018). Finally, GDPR has substantially changed Twitter's handling of data and advertising practices. Twitter changed its management of European data in its Irish processing facility following GDPR and created stricter regulations for advertisers buying microtargeted advertisements (GDPR Twitter for Business FAQ). While Twitter has not had perfect GDPR compliance, its violation is from delays in notifying users of a data breach rather than issues with restricting access to personal data (Porter 2020).

Section V: Data and Methods

Now that the specific case selection for this paper has been outlined, I will address the data and methodology of estimating the effect of GDPR on the Russian disinformation campaign against the UK. Specifically, I will address the independent, dependent, and control variables in my model, and explain the specifications behind the time-series analysis of the implementation of GDPR.

Firstly, the data for this paper has been sourced from Twitter, which releases data about accounts that have been confirmed to be involved in foreign disinformation campaigns. The accounts in this dataset were all shut down throughout 2019 for affiliation with the IRA, and several accounts continued their activities through July of 2019 (Twitter 2019). This shutdown

process means that many, though not all 416 accounts, were active through the implementation of GDPR on May 24th, 2018. This is sufficient to meet the time requirements for this study, though an ideal dataset would have many more accounts and tweets. Most importantly, the dataset codes for the user ID of the Twitter account, the text of the tweet, and the time of the posting of the tweet.

This data was then narrowed according to the independent variable, dependent variable, and control variables for this model. The independent variable for this project is the implementation of GDPR. The official deadline for implementing GDPR was May 24th, 2018. Though several websites – including Twitter – updated their policies before the May 24th deadline, they did so only two weeks before the deadline at the earliest. This varied implementation of data privacy regulations is unlikely to have a large effect, particularly as though the period examined has been extended to account for this two-week variance. Therefore, the independent variable for this project is May 24th, or the day that GDPR was officially implemented in the EU.

The dependent variable for this paper is the change in the average amount of tweets by IRA accounts relating to the UK per day. Following the procedure outlined by Howard and Kollanyi in their 2016 paper, the average amount of tweets per day simplifies the natural variance in the rhythm of tweets as the day progresses. I collected the total number of tweets per day that discussed the UK, England, or Brexit.[6] This resulted in 109 accounts and 5837 total tweets being selected.[7] These accounts were active between 2012 and October of 2018, though

[6] While additional nuance could have been added to find which topics targeted UK citizens, the most obvious comments were those mentioning the UK, England, or Brexit. Several accounts tweeting about these topics appeared in the tweet to pretend to be British citizens, or U.S. citizens remarking on the U.K. Ultimately, this rudimentary strategy proved effective in discovering relevant tweets to British audiences and therefore worked for this analysis.
[7] Though this seems small relative to other twitter datasets, all of these tweets emerged from accounts affiliated with the IRA. As the focus of this study is only accounts affiliated with foreign disinformation campaigns, not the

many were shut down before October. While this restriction to the tweets mentioning the UK casts a narrow time scope for the study, it is sufficient to analyze the effect of the independent variable by comparing online activity before and after the implementation of GDPR on May 24th, 2018.

The period of observation selected for this study should be long enough to demonstrate a substantive change in online behavior, while not confounding the data as accounts get shut down and overall activity consequently decreases. This paper observes all tweets from January 1st, 2018, to the date that most accounts have been suspended, or July 20th, 2018. The trend in tweets per day drops off significantly after July 20th, as several accounts never tweeted again after that date. The data has therefore been narrowed, as needed, by the selected dates.

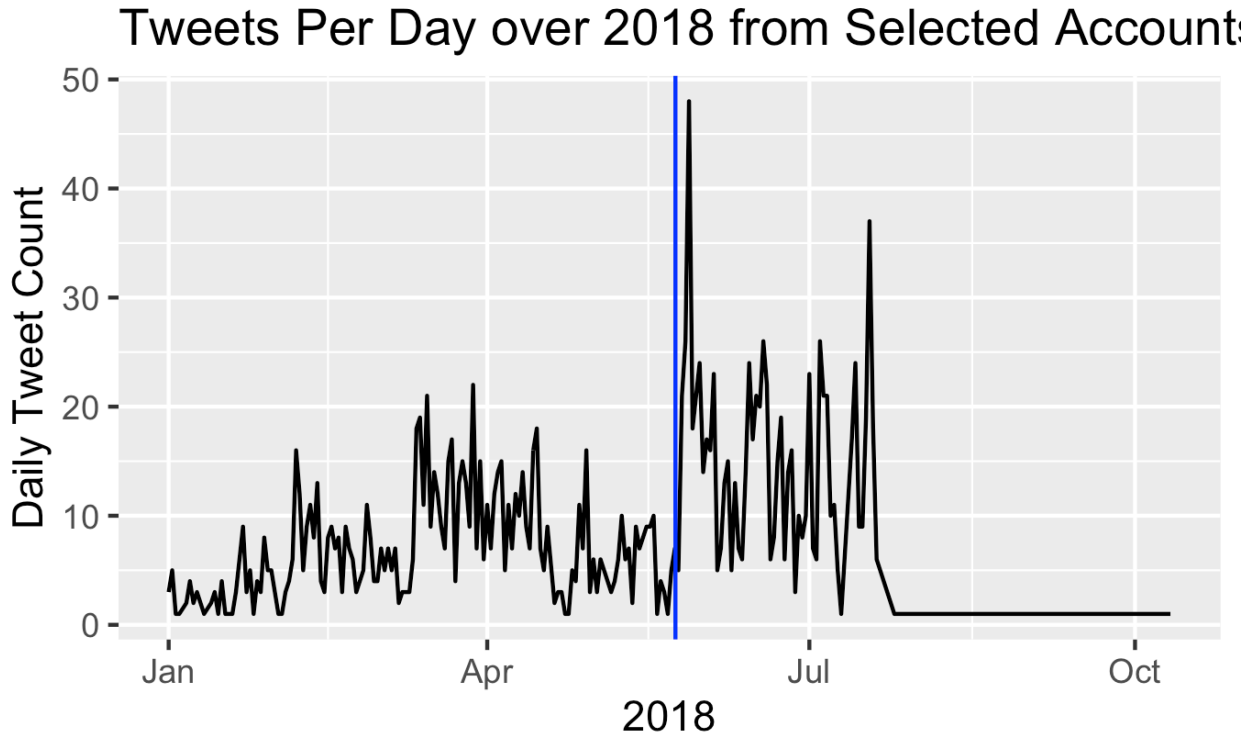

Image 1: The daily number of tweets per day that mentioned the UK, England, or Brexit. The day that GDPR was implemented, or May 24th, 2018, is visually represented as the blue line. The drop off in activity following July 20th, 2018, is attributed to Twitter shutting down the accounts, rather than a plan from the IRA

---

accounts of foreign citizens or the discussions that they influence, this number can be expected to be much smaller. In an ideal world, there would be much more data available from the time period of interest.

The control variable for this model is time, which is treated as its own variable in this study which naturally increases over time as tweets and tweet interactions naturally vary and increase. Some accounts may increase their activity throughout the study; for example, an account that might be getting more and more popular over time might naturally tweet more over time. Time therefore must be used as a control variable to resolve whether an increase in the number of tweets over time is simply natural or not.

Finally, there is the possibility of a confounding variable in an idiosyncratic event that may interrupt the study of the event. In this case, two confounding variables must be considered: the passage of other laws to combat disinformation campaigns, and possible large disinformation news stories that would draw specific attention from disinformation campaigns. First, there could be another law that affects the disinformation campaign which is not the data privacy law in question. Given the increased concern about disinformation campaigns, there is a risk that another law specifically addressing disinformation campaigns may interrupt the behavior of the campaign instead of the data privacy law. Second, a large news story that is particularly relevant to a disinformation campaign may change the behavior of a disinformation campaign. This behavior change – typically increasing the number of posts above the status quo – could coincide with the implementation of the data privacy regulations in two ways. First, the time before implementing the law could overlap with the end of a news cycle of one of these big stories, meaning the news story would increase online interactions before the law is implemented, and the natural passage of the story would decrease interactions afterward. Alternatively, the breaking of an important news story could coincide with the implementation of the data privacy law, increasing the online presence of the data privacy campaign after the passage of the law. These changes from a big news story or the implementation of another disinformation law could

affect the online interactions of accounts spreading disinformation, and they must be considered when evaluating the hypotheses.

In accounting for these confounding variables, the case selection of the Russian campaign against the EU solves the first problem because no other major anti-disinformation or data privacy laws were passed during the period of interest. The UK government did not pass any major laws that might have affected the disinformation campaign between the dates of interest to this paper. The UK has struggled to pass data protection laws as it balances pressures from technology companies to keep an open data economy with the security and privacy concerns of its citizens (Richards 2020). Meanwhile, the actions against disinformation remained focus on investigations to determine what foreign disinformation campaigns had done, rather than taking actions against them. The only actionable item – the creation of the National Security Communications Unit to combat fake news –happened in very early January of 2018 and had most likely already affected the disinformation campaigns before the May 25th implementation of GDPR (Flunke and Flamini 2018). Likewise, the European Commission was still investigating disinformation efforts rather than taking the legislative action it would officially pass in 2019 (European Commission 2018).

Addressing the second confounding variable – a large news story of interest to the disinformation campaign – is slightly more complicated. When examining the data, I ran a text analysis of the tweet content to see the most popular conversation topics and hashtags. Many of the tweets in the dataset related to Tommy Robinson, a British far-right anti-Islam political activist and former leader and founder of the English Defense League. On May 25th, 2018, Robinson was arrested for disturbing the peace in attempting to live-stream an ongoing court trial and sentenced to 13 months in jail, only one day after GDPR was implemented. The judge in the

case issued a reporting restriction on Robinson's arrest as it pertained to an ongoing trial, which captured the attention of the far-right media across the world. His arrest and the moratorium on reporting about his case trended on Twitter (MacGuill 2018). This story of such a divisive figure in British captured the attention of the IRA, where accounts began commenting on the ordeal.

Image 2: Word Cloud of Topics in Tweets mentioning the UK, England, or Brexit. Topics related to the Tommy Robinson case dominated the conversation, as highlighted in the image above

A right-wing news story like Robinson's impacts the number of tweets per day, and this confounding increase in online activity must be considered. This story is particularly important given the increase in conversations about Robinson that occurred only one day after the deadline for implementing GDPR's protocol. As a result, the number of tweets about the Robinson story for the time of interest was separated and treated as a final control variable, though the tweets discussing the case were not removed from the actual dataset.

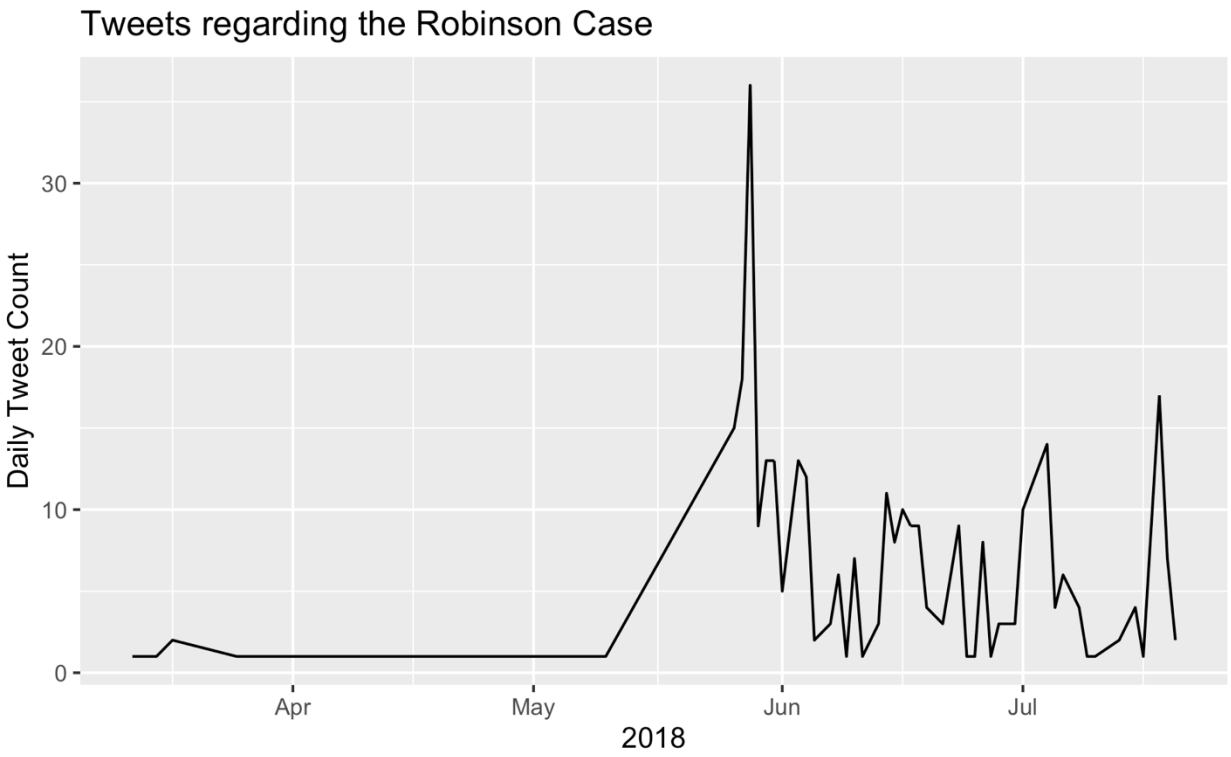

Image 3: Tweets about the Tommy Robinson case over time

Given that the independent variable for this study is an event, the most logical approach to operationalize the question of whether the implementation of GDPR affected the dependent variable is an event study analysis.

Event studies are used to measure the effect of a specific event on the value of a variable. The event study not only determines whether there is a change in the variable by the interrupting event, but whether overall that event increases or decreases. The model can formally be described below in equation (1):

$$R = \alpha+\beta t +c(\gamma_i+\lambda_i*(t-i)). \quad (1)$$

In this model, R is the real value of the tweets per day, and i is the event date. Critically, the dummy variable i remains 0 before the day of implementing GDPR and is 1 after the date of implementing GDPR. With c as a control variable, this equation approximates the new linear model of expected tweets after GDPR has been implemented. Variable t, or time since the

implementation of GDPR which remains 0 before the day of implementing GDPR, 1 the day of implementing GDPR, and increases by 1 day for every day after GDPR, has one subtracted from it because it would otherwise interfere with γ, or the change in the constant from implementing GDPR.

The variables in the event study are thus. First, the natural fluctuation of tweets over time is accounted for by a time trend, which starts at count 1 on January 1st, 2018, and increases by 1 for each day observed in the period of interest. Second, the dummy variable γ addresses the significance of the event. The coding for γ makes all days before the implementation of GDPR 0, and all the days after the implementation of GDPR 1, including the day that GDPR was implemented. This variable γ therefore estimates if there was any change in the number of tweets per day for the period preceding the implementation of GDPR, and the period after implementing GDPR. A third variable λ measures the significance of the time since passing GDPR. The λ variable is coded as 0 for all days before implementing GDPR, 1 for the day GDPR is implemented, and increases by 1 for every day since GDPR has been implemented. It serves as an additional test to see if the time trend after the event is significantly different from the time trend preceding the event. A final variable accounts for the significance of the Robinson story by counting the total number of tweets about the Robinson story per day. The total number of IRA tweets, or all tweets mentioning the UK, England, or affiliated terms per day for the period between January 1st, 2018, and July 7th, 2018, are regressed on all four of these variables.

For hypothesis 1, both γ and λ would be very small, practically negligible, as there would be no change from the initial linear model of $\alpha+\beta t$ necessary for calculating the number of tweets per day. For hypothesis 2, γ would be negative to demonstrate a drop-off in social media activity after the implementation of GDPR where the foreign disinformation campaign

temporarily pauses its activity. An increase in λ would indicate that the foreign disinformation campaign had resumed. For hypothesis 3, γ would be positive, indicating an increase in social media activity after implementing GDPR as the foreign disinformation campaign attempts to substitute its microtargeting activities with algorithm manipulation. Finally, λ would demonstrate the effectiveness of the algorithm manipulation, as it indicates whether the foreign disinformation campaign continued its increased online activities or slowly abandoned them over time.

Section VI: Results

Two regression tables below detail the results of the linear regressions. In Table 1, I regressed the number of tweets per day about the UK on all the variables described above – the time trend since January 1st, 2018, the binary event having occurred variable, the time trend after the event, and the number of tweets from the Robinson story. The results suggest that the time trend and the Robinson story are highly significant with $p<0.01$, and the binary event indicator is significant with $p<0.05$. While the effect of the Robinson story is positive, the overall effect of GDPR on the number of tweets is negative at -2.917, meaning that the total number of tweets per day after implementing GDPR reduced by approximately three tweets. The percent reduction in the constant is approximately 33%. This is because the constant 4.966+0.027 multiplied by the total number of days between January 1st and May 24th (140), or 8.74, by 2.917 after the event. This finding supports Hypothesis 2 that the IRA put a pause on their online activity following GDPR.

Regression Results (Table 1)

	Dependent variable:
	Number of Tweets
	Results
Time Trend (Total)	0.027^{***}
	(0.010)
γ (Event Significance)	-2.917^{*}
	(1.657)
λ (Time Trend After Event)	-0.008
	(0.036)
Significance of Robinson story	1.338^{***}
	(0.100)
Constant	4.966^{***}
	(0.776)
Observations	200
R^2	0.615
Adjusted R^2	0.607
Residual Std. Error	4.531 (df = 195)
F Statistic	77.758^{***} (df = 4; 195)
Note:	*p<0.1; **p<0.05; ***p<0.01

The time trend after the event is not significant at the p<0.1 level, meaning GDPR therefore did not significantly affect the overall growth in the number of tweets per day; however, the actual coefficient is negative, meaning that the number of tweets after the implementation of GDPR is decreasing by about 0.008. This estimated impact is rather large compared to the total time trend of 0.027, meaning that the growth in the number of tweets per day over time could have slowed down by approximately 0.296, or ~29.6%. This reduction coefficient is similar to the 33% reduction of the constant on May 24th. The reduction of both the constant number of tweets and the rate at which tweets are increasing over time are very similar

at about 0.3 reductions. However, unlike the leveling effect of the binary event operator, these findings are inconclusive as they are not significant.

Given the insignificance of the time trend after the event, and that it might have interfered with the binary event significance γ, I took λ out of the regression, as observed in Table 2. For this regression, the time trend remained significant of increasing the total number of tweets per day by 0.026 at $p<0.01$, making time a small but significant variable. The binary before-or-after event variable γ increased to -3.14, with $p <0.05$, and a t-test significance nearing 0.01 significance level. This effect is particularly sizable given the overall constant of 5.006, meaning that the implementation of GDPR correlated with a ~36% drop in the overall constant number of tweets. Finally, the Robinson story remained significant at 0< 0.01.

Regression Results (Table 2)

	Dependent variable:
	Number of Tweets Model Results
Time Trend (Total)	0.026^{***}
	(0.009)
γ (Event Significance)	-3.143^{**}
	(1.299)
Significance of Robinson story	1.347^{***}
	(0.092)
Constant	5.006^{***}
	(0.753)
Observations	200
R^2	0.615
Adjusted R^2	0.609
Residual Std. Error	4.520 (df = 196)
F Statistic	104.166^{***} (df = 3; 196)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Section VII: Discussion and Conclusion:

Based on the evidence from the regressions on the total number of tweets, it appears that the day that GDPR was implemented likely had a big and significant effect on the number of tweets emerging from the IRA's disinformation accounts. This observed change means we can reject the null hypothesis, or Hypothesis 1, that the implementation of GDPR did not effect on the overall number of tweets with $p<0.1$, or a 90% confidence interval for the first regression table, and $p<0.5$, or a 95% confidence interval in the second regression. As the regression suggests the event date reduced the constant level of tweets for both regressions, the evidence supports Hypothesis 2 that the IRA temporarily paused some of their activities following GDPR. This result is only conclusive when considering the Robinson news story, which increased the number of tweets from May 25th, 2018, onwards. Finally, the effect that the implementation of GDPR had on the overall time trend is inconclusive, despite having a large negative coefficient compared to the overall time trend. Interestingly, both the constant and the time trend variables were reduced by about the same amount – approximately 0.3 – which would suggest a permanent change in the number of tweets by approximately 30%.

These findings have several implications, beyond supporting Hypothesis 2 that GDPR reduced the total number of online interactions from foreign disinformation campaigns. First, it suggests that the Russians continued to employ microtargeting techniques to influence UK citizens well after the Cambridge Analytica scandal broke. Further, it suggests that the supply chain for the data Russians use to micro-target individuals on Twitter may have been affected by GDPR. This baseline gives us further insight into the mechanisms the IRA used for their 2018 disinformation, as well as several mechanisms to further study to interrupt data harvesting for

foreign disinformation campaigns. Finally, there is evidence that data privacy laws may address the security risks of disinformation. This study suggests that data privacy laws should play a role in combatting disinformation, particularly restrictions on personal data.

The contributions of this paper are the following. First, I organize and label the strategies of spreading disinformation targeted at other states identified by the growing literature and discuss two strategies – the microtargeting and manipulation explanations – in detail. Second, I emphasize the need for research into the interruptions into strategies of spreading disinformation, rather than simply the strategies themselves. Foreign disinformation campaigns are often unintentionally interrupted by an event or a law that can unknowingly affect the method of spreading disinformation. Further study should investigate the behavior after these "interruption" events to find out if foreign disinformation campaigns abandon the project until the problem can be overcome or adopt another strategy to spread disinformation entirely. Foreign actors can switch and even combine different strategies to spread disinformation in the same campaign, and the literature has thus far ignored this overlap and evolution from interruption events. Finally, I lend evidence to the assertion that the controversial data privacy laws may have security benefits in minimizing or interrupting the microtargeting tactic to spread foreign disinformation.

The most obvious area for further study based on this finding is exactly what elements of GDPR might have interrupted the supply chain of data to IRA, and therefore reduced the online activity of actors spreading disinformation. This paper suggested three different elements of GDPR that may have affected the supply chain of data for microtargeting: the consent requirement, the anonymization requirement, and the scrutiny of third-party sellers and international actors. As the EU and the UK implemented all these elements at the same time, it is impossible to observe from the reduction in the total number of tweets which, if any, of these

elements impacted the supply chain. This causal link should be investigated further. Possible paths of study could be a deeper investigation of the overall impact of the three GDPR elements in disrupting the status quo. Another possibility is contrasting GDPR with other data privacy laws that only employ some of these elements, such as the consent requirement but no scrutiny of third-party data brokers, in a natural experiment. Several of these bills are being considered in legislatures all over the world and will likely have differing effects on disinformation campaigns given their different data privacy regulations.

This study also made several assumptions, many of which should be examined further given sufficient data. I assumed that many of the accounts were shut down in the middle of September of 2018. However, the dataset did not reveal exactly when accounts were taken offline. Given this information, several other variables should be considered. For example, this study could not account for the popularity of accounts based on the number of followers, as several popular accounts appeared to be shut down by Twitter in late 2017. Confirmation of when Twitter took down such accounts would greatly assist in checking if these accounts were shut down by the platform, or in a more organized manner. Data about when the account was removed by Twitter would also assist in controlling for the total number of accounts active at any given time over the time trend. Other data, such as the number of followers at the time of posting the tweet rather than at the time of shutting down the account, would further help account for the reach of a disinformation campaign and possible increases in online activity from popular disinformation accounts.

One final area of study is the treatment of large news stories before and after the passage of GDPR. The Tommy Robinson story was likely so important to the disinformation campaign that it increased the total number of tweets per day, but the question remains how much the story

affected the total number of tweets if it occurred before GDPR was implemented. There is an opportunity to investigate the different approaches to large, trending news stories before and after GDPR, and seeing if there is a difference indicative of a more generalized or less microtargeted campaign to spread disinformation.

Works Cited

Albrecht, J. 2016. "How the GDPR Will Change the World."

Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives* 31(2): 211–36.

Arnaudo, Daniel. 2019. "How Building Data Protection Regimes Can Counter Disinformation." *Defusing Disinfo*. https://defusingdis.info/2019/03/11/how-building-data-protection-regimes-can-counter-disinformation/ (November 8, 2020).

Bakir, Vian, and Andrew McStay. 2018. "Fake News and The Economy of Emotions." *Digital Journalism* 6(2): 154–75.

Bastos, Marco T., and Dan Mercea. 2019. "The Brexit Botnet and User-Generated Hyperpartisan News." *Social Science Computer Review* 37(1): 38–54.

Benkler, Yochai, Robert Faris, and Hal Roberts. Network Propaganda *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190923624.001.0001/oso-9780190923624 (November 8, 2020).

Bennett, Lance, and Steven Livingston. 2018. "The Disinformation Order: Disruptive Communication and the Decline of Democratic Institutions." *European Journal of Communication* 33: 122–39.

Booth, Robert et al. 2017. "Russia Used Hundreds of Fake Accounts to Tweet about Brexit, Data Shows." *the Guardian*. http://www.theguardian.com/world/2017/nov/14/how-400-russia-run-fake-accounts-posted-bogus-brexit-tweets (July 1, 2021).

Bradshaw, Samantha, and Philip Howard. 2019. *The Global Disinformation Order 2019 Global Inventory of Organised Social Media Manipulation*. University of Oxford. https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf.

Bradshaw, Samantha, and Philip N. Howard. 2018. *Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation*. University of Oxford. https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/07/ct2018.pdf.

"Brexit: The False, Misleading and Suspicious Claims CrossCheck Has Uncovered so Far." 2019. *First Draft*. https://firstdraftnews.org:443/latest/brexit-the-false-misleading-and-suspicious-claims-crosscheck-has-uncovered/ (November 2, 2020).

Bulger, Monica, and Patrick Davison. 2018. "The Promises, Challenges, and Futures of Media Literacy." *Journal of Media Literacy Education* 10(1): 1–21.

Chadwick, A., and C. Vaccari. 2019. "News Sharing on UK Social Media: Misinformation, Disinformation, and Correction."

Chamberlain, P. R. 2010. "Twitter as a Vector for Disinformation." *Journal of Information Warfare* 9(1): 11–17.

Claesson, Annina. 2019. "Coming Together to Fight Fake News: Lessons from the European Approach to Disinformation." https://www.csis.org/coming-together-fight-fake-news-lessons-european-approach-disinformation (December 6, 2020).

Conversation, Filippo Menczer, The. "Fake Online News Spreads Through Social Echo Chambers." *Scientific American*. https://www.scientificamerican.com/article/fake-online-news-spreads-through-social-echo-chambers/ (March 10, 2021).

Daskal, Jennifer. 2019. "Facebook's Ban on Foreign Political Ads Means the Site Is Segregating Speech." *Washington Post*. https://www.washingtonpost.com/outlook/2019/12/16/facebooks-ban-foreign-political-ads-means-site-is-segregating-speech/ (March 19, 2021).

Dawson, Andrew, and Martin Innes. 2019. "How Russia's Internet Research Agency Built Its Disinformation Campaign." *The Political Quarterly* 90(2): 245–56.

European Commission. 2015a. "Art. 4 GDPR – Definitions." *General Data Protection Regulation (GDPR)*. https://gdpr-info.eu/art-4-gdpr/ (June 30, 2021).

———. 2015b. *General Data Protection Regulation*. https://gdpr-info.eu/art-4-gdpr/.

———. 2018. *A Multi-Dimensional Approach to Disinformation*. Brussels: European Commission. https://digital-strategy.ec.europa.eu/en/library/final-report-high-level-expert-group-fake-news-and-online-disinformation.

Fallis, Don. 2015. "What Is Disinformation?" *Library Trends* 63(3): 401–26.

"Final Report of the High Level Expert Group on Fake News and Online Disinformation | Shaping Europe's Digital Future." https://digital-strategy.ec.europa.eu/en/library/final-report-high-level-expert-group-fake-news-and-online-disinformation (July 17, 2021).

Fletcher, Richard, Alessio Cornia, Lucas Graves, and Rasmus Kleis Nielsen. *Measuring the Reach of "Fake News" and Online Disinformation in Europe*. University of Oxford: Reuters Institute. https://reutersinstitute.politics.ox.ac.uk/our-research/measuring-reach-fake-news-and-online-disinformation-europe.

Funke, Daniel, and Daniela Flamini. 2018. "A Guide to Anti-Misinformation Actions around the World." *Poynter*. https://www.poynter.org/ifcn/anti-misinformation-actions/ (July 1, 2021).

Garimella, Kiran, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. 2018. "Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship." In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, 913–22. https://doi.org/10.1145/3178876.3186139 (March 10, 2021).

"GDPR Twitter for Business FAQ." *Twitter GDPR*. https://gdpr.twitter.com/en/faq.html (July 17, 2021).

Gelfert, Axel. 2018. "Fake News: A Definition." *Informal Logic* 38(1): 84–117.

Gerrits, André W. M. 2018. "Disinformation in International Relations: How Important Is It?" *Security and Human Rights* 29(1–4): 3–23.

Gray, Colin M. et al. 2018. "The Dark (Patterns) Side of UX Design." In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: Association for Computing Machinery, 1–14. https://doi.org/10.1145/3173574.3174108 (July 17, 2021).

Guess, Andrew, Brendan Nyhan, and Jason Reifler. 2018. "Selective Exposure to Misinformation: Evidence from the Consumption of Fake News during the 2016 U.S. Presidential Campaign." *European Research Council*. http://www.ask-force.org/web/Fundamentalists/Guess-Selective-Exposure-to-Misinformation-Evidence-Presidential-Campaign-2018.pdf.

Howard, Philip N., and Bence Kollanyi. 2016. "Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum." *arXiv:1606.06356 [physics]*. http://arxiv.org/abs/1606.06356 (March 17, 2021).

Humprecht, Edda, Frank Esser, and Peter Van Aelst. 2020. "Resilience to Online Disinformation: A Framework for Cross-National Comparative Research." *The International Journal of Press/Politics* 25(3): 493–516.

Jackson, Sarah J., and Brooke Foucault Welles. 2015. "Hijacking #myNYPD: Social Media Dissent and Networked Counterpublics." *Journal of Communication* 65(6): 932–52.

Jr, Edson C. Tandoc, Zheng Wei Lim, and Richard Ling. 2018. "Defining 'Fake News.'" *Digital Journalism* 6(2): 137–53.

Karlova, Natascha A., and Jin Ha Lee. 2011. "Notes from the Underground City of Disinformation: A Conceptual Investigation." *Proceedings of the American Society for Information Science and Technology* 48(1): 1–9.

Keller, Franziska B., David Schoch, Sebastian Stier, and JungHwan Yang. 2020. "Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign." *Political Communication* 37(2): 256–80.

Kessler, Sarah. 2014. "How Twitter Bots Fool You Into Thinking They Are Real People." *Fast Company*. https://www.fastcompany.com/3031500/how-twitter-bots-fool-you-into-thinking-they-are-real-people (March 1, 2021).

Kornbluh, Karen. "Could Europe's New Data Protection Regulation Curb Online Disinformation?" *Council on Foreign Relations*. https://www.cfr.org/blog/could-europes-new-data-protection-regulation-curb-online-disinformation (November 8, 2020).

Lanoszka, Alexander. 2019. "Disinformation in International Politics." *European Journal of International Security* 4(2): 227–48.

Lazer, David et al. 2017. "Combating Fake News: An Agenda for Research and Action." *Harvard Kennedy School Shorenstein Center*. https://shorensteincenter.org/combating-fake-news-agenda-for-research/ (November 14, 2020).

Lazer, David M. J. et al. 2018. "The Science of Fake News." *Science* 359(6380): 1094–96.

MacGuill, Dan. "Was a Far-Right Activist Jailed for Breaching a Court Order Designed to 'Protect Muslim Pedophiles'?" *Snopes.com*. https://www.snopes.com/fact-check/tommy-robinson-arrest/ (July 2, 2021).

Martin, Diego A., Jacob N. Shapiro, and Michelle Nedashkovskaya. 2019. "Recent Trends in Online Foreign Influence Efforts." *Journal of Information Warfare* 18(3): 15–48.

Meserole, Chris. 2018. "How Misinformation Spreads on Social Media—And What to Do about It." *Brookings*. https://www.brookings.edu/blog/order-from-chaos/2018/05/09/how-misinformation-spreads-on-social-media-and-what-to-do-about-it/ (March 17, 2021).

Nickerson, Raymond S. 1998. "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises:" *Review of General Psychology*. https://journals.sagepub.com/doi/10.1037/1089-2680.2.2.175 (November 8, 2020).

Nouwens, Midas et al. 2020. "Dark Patterns after the GDPR: Scraping Consent Pop-Ups and Demonstrating Their Influence." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: Association for Computing Machinery, 1–13. https://doi.org/10.1145/3313831.3376321 (July 17, 2021).

Pennycook, Gordon, and David G. Rand. 2017. *Who Falls for Fake News? The Roles of Bullshit Receptivity, Overclaiming, Familiarity, and Analytic Thinking*. Rochester, NY: Social Science Research Network. SSRN Scholarly Paper. https://papers.ssrn.com/abstract=3023545 (March 10, 2021).

———. 2019. "Lazy, Not Biased: Susceptibility to Partisan Fake News Is Better Explained by Lack of Reasoning than by Motivated Reasoning." *Cognition* 188: 39–50.

Popken, Ben. 2017. "Russian Trolls Duped Global Media Thousands of Times." *NBC News*. https://www.nbcnews.com/tech/social-media/trump-other-politicians-celebs-shared-boosted-russian-troll-tweets-n817036 (July 1, 2021).

Porter, Jon. 2020. "Twitter Hit with €450,000 GDPR Fine Nearly Two Years after Disclosing Data Breach." *The Verge*. https://www.theverge.com/2020/12/15/22176008/twitter-gdpr-fine-protected-tweets-ireland-data-protection-commission (July 17, 2021).

Prior, Markus, Gaurav Sood, and Kabir Khanna. 2015. "You Cannot Be Serious: The Impact of Accuracy Incentives on Partisan Bias in Reports of Economic Perceptions." *Quarterly Journal of Political Science* 10(4): 489–518.

Richards, Julian. 2021. "Fake News, Disinformation and the Democratic State:: A Case Study of the UK Government's Narrative." *Icono14* 19(1): 95–122.

Romm, Tony, and Isaac Stanley-Becker. 2019. "Twitter to Ban All Political Ads amid 2020 Election Uproar." *Washington Post*. https://www.washingtonpost.com/technology/2019/10/30/twitter-ban-all-political-ads-amid-election-uproar/ (July 17, 2021).

RSI Security. 2020. "Data Protection & Social Media: How GDPR Influences Today's Social Media Platforms." *RSI Security*. https://blog.rsisecurity.com/data-protection-and-social-media/ (December 6, 2020).

Satariano, Adam, and Amie Tsang. 2019. "Who's Spreading Disinformation in U.K. Election? You Might Be Surprised." *The New York Times*. https://www.nytimes.com/2019/12/10/world/europe/elections-disinformation-social-media.html (November 2, 2020).

Shin, Jieun, Lian Jian, Kevin Driscoll, and François Bar. 2016. "Political Rumoring on Twitter during the 2012 US Presidential Election: Rumor Diffusion and Correction:" *New Media & Society*. https://journals.sagepub.com/doi/10.1177/1461444816634054 (November 8, 2020).

Smith, Aaron, and Monica Anderson. 2018. "Social Media Use 2018: Demographics and Statistics." *Pew Research Center*. https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018/ (July 1, 2021).

Somerville, Alistair, and Jonas Heerin. 2020. "The Disinformation Shift: From Foreign to Domestic." *Georgetown Journal of International Affairs*. https://gjia.georgetown.edu/2020/11/28/the-disinformation-shift-from-foreign-to-domestic/ (March 4, 2021).

Törnberg, Petter. 2018. "Echo Chambers and Viral Misinformation: Modeling Fake News as Complex Contagion." *PLoS ONE* 13(9). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6147442/ (March 10, 2021).

Tucker, Joshua A. et al. 2018. *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature*. Rochester, NY: Social Science Research Network. SSRN Scholarly Paper. https://papers.ssrn.com/abstract=3144139 (November 2, 2020).

Twitter. 2019. *Twitter Elections Integrity Datasets: Russia (January 2019) Dataset*. Twitter. https://transparency.twitter.com/en/reports/information-operations.html.

Urman, Aleksandra, and Mykola Makhortykh. 2019. "Webs of Deception: Detecting and Measuring the Diffusion of Online Disinformation During the Elections in Ukraine." Presented at the Digital Societies 2019, Konstanz, Germany. *Urman, Aleksandra; Makhortykh, Mykola (26 September 2019). Webs of Deception: Detecting and Measuring the Diffusion of Online Disinformation During the Elections in Ukraine (Unpublished). In: Digital Societies 2019. Konstanz, Germany. 25-27 September.* https://digitalsocieties2019.net/program-events/full-program/ (November 2, 2020).

Vériter, Sophie L., Corneliu Bjola, and Joachim A. Koops. 2020. "Tackling COVID-19 Disinformation: Internal and External Challenges for the European Union." *The Hague Journal of Diplomacy* 15(4): 569–82.

Wardle, Claire, and Hossein Derakhshan. 2018. "Thinking about 'Information Disorder': Formats of Misinformation, Disinformation and Mal-Information." In *Journalism, "Fake News" and Disinformation: A Handbook for Journalism Education and Training*, UNESCO, 43–53. https://en.unesco.org/sites/default/files/f._jfnd_handbook_module_2.pdf.

Wilson, Tom, and Kate Starbird. 2020. "Cross-Platform Disinformation Campaigns: Lessons Learned and Next Steps." *Harvard Kennedy School Misinformation Review* 1(1). https://par.nsf.gov/biblio/10171226-cross-platform-disinformation-campaigns-lessons-learned-next-steps (March 18, 2021).

Woolley, Samuel C., and Douglas Guilbeault. 2018. Computational Propaganda *United States: Manufacturing Consensus Online*. Oxford University Press. https://oxford-universitypressscholarship-com.proxy.uchicago.edu/view/10.1093/oso/9780190931407.001.0001/oso-9780190931407-chapter-9 (March 18, 2021).

Woolley, Samuel C., and Philip N. Howard. 2018. Computational Propaganda *Conclusion: Political Parties, Politicians, and Computational Propaganda*. Oxford University Press. https://oxford-universitypressscholarship-com.proxy.uchicago.edu/view/10.1093/oso/9780190931407.001.0001/oso-9780190931407-chapter-11 (March 18, 2021).

Xia, Yiping et al. 2019. "Disinformation, Performed: Self-Presentation of a Russian IRA Account on Twitter." *Information, Communication & Society* 22(11): 1646–64.

Zannettou, Savvas et al. 2019. "Disinformation Warfare: Understanding State-Sponsored Trolls on Twitter and Their Influence on the Web." *arXiv:1801.09288 [cs]*. http://arxiv.org/abs/1801.09288 (March 18, 2021).

Zarsky, Tal. 2017. *Incompatible: The GDPR in the Age of Big Data*. Rochester, NY: Social Science Research Network. SSRN Scholarly Paper. https://papers.ssrn.com/abstract=3022646 (July 17, 2021).