

Oxford Internet Institute, University of Oxford

Assignment Cover Sheet

Candidate Number	1055782
Assignment	Dissertation
Term	Trinity Term 2021
Title/Question	Leveraging the Left: Examining Russian Left-Leaning Provocation on Instagram in 2020
Word Count	14,998

By placing a tick in this box ☑ I hereby certify as follows:

(a) This thesis or coursework is entirely my own work, except where acknowledgments of other sources are given. I also confirm that this coursework has not been submitted, wholly or substantially, to another examination at this or any other University or educational institution;

(b) I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at `https://www.ox.ac.uk/students/academic/guidance/skills?wssl=1`.

(c) I agree that my work may be checked for plagiarism using Turnitin software and have read the Notice to Candidates which can be seen at: `http://www.admin.ox.ac.uk/proctors/turnitin2w.shtml`, and that I agree to my work being screened and used as explained in that Notice;

(d) I have clearly indicated (with appropriate references) the presence of all material I have paraphrased, quoted or used from other sources, including any diagrams, charts, tables or graphs.

(e) I have acknowledged appropriately any assistance I have received in addition to that provided by my [tutor/supervisor/adviser].

(f) I have not sought assistance from a professional agency;

(g) I understand that any false claims for this work will be reported to the Proctors and may be penalized in accordance with the University regulations.

Please remember:

- To attach a second relevant cover sheet if you have a disability such as dyslexia or dyspraxia. These are available from the Higher Degrees Office, but the Disability Advisory Service will be able to guide you.

University of Oxford



Leveraging the Left: Examining Russian Left-Leaning Provocation on Instagram in 2020

1055782

Thesis submitted in partial fulfilment of the requirement for the degree of MSc in Social Data Science at the Oxford Internet Institute at the University of Oxford

Trinity Term 2021

14,998 words

1 Acknowledgements

I would like to dedicate this thesis to my late grandparent and namesake. Their memory has inspired me both personally and intellectually, especially to strive for the kind of opportunities in higher education that were never open to them. Though they certainly left big shoes for me to fill, I would like to believe that they would be proud that a dissertation from Oxford has their name on it.

I am also deeply grateful to my advisor, — , for their advice, guidance and support during the process of writing my dissertation. In addition, I would like to thank my parents and my siblings, J., A., P., and N., for their endless love and support throughout this year, even from thousands of miles away. Lastly, I would like to thank my friends at Oxford and from home, especially A. and C., for their encouragement and unwavering kindness.

2 Abstract

Pursuing a strategy of provocation, the Russian government has leveraged proxy media accounts on social media to disseminate strong, often-fringe political stances within a broader left-leaning ecosystem. While these proxy voices, termed "the Fringenet," author such narratives, their posts are spread by "fellow travellers," accounts that wittingly or unwittingly amplify Russian provocation efforts. Beyond their symbolic importance, it remains unclear whether reposts have any significance, such as generating follower growth for Fringenet accounts.

This thesis seeks to tackle this question with a series of theoretical and methodological contributions to the literature. By drawing together a novel dataset of Instagram reposts and posts by Russian provocation accounts, social network theory, and time-series modelling, I examine how three Fringenet media channels built their amplification network of followers. To determine whether reposts influence follower growth, I use vector autoregression models paired with Granger causality tests and impulse response functions to explore the time-ordered relationship between reposts and follower growth. Further, to examine who is reposting Russian provocation content and on what topics, I conduct an intensive manual coding process of account types (N = 1,870) and posts and reposts (N = 4,050).

My results reveal both a top-down broadcast model – where influential users like private companies and influencers generated follower growth for Fringenet accounts by signalling to their followers that these accounts were authentic and credible – and a bottom-up social contagion model – where the sheer number of users reposting Fringenet accounts led to follower growth by building a critical mass. Furthermore, my descriptive analyses highlight that reposts were frequently shared by fellow independent and alternative news accounts as well as authentic private companies and activists and civil society groups. Finally, I find reposts and posts converging around topics such as Black Lives Matter, climate change, and COVID-19.

Given recent decisions to label Russian media accounts as state-run, the substantive and methodological contributions of this thesis are relevant for the academic community and for policymakers. For the academic community, this thesis demonstrates the need for a multiplayer analysis when studying Russian information operations. For policymakers, this thesis highlights the drawbacks of a focus solely on countering disinformation through fact-checking, which overlooks the way Russian accounts are not merely passive purveyors of false information but rather active participants fanning the flames of ongoing political discourse.

Contents

1 Acknowledgements **1**

2 Abstract **2**

3 Introduction **6**

4 Russia's Provocation Strategy: The Fringenet, "Fellow Travellers," and Followers **10**

4.1 Russian Information Operations: The Fringenet's Strategy of Provocation & Propaganda . . . 10

4.2 Fellow Travellers: Networks of Amplification . . . 15

4.3 Following: Assembling Networks and Audiences of Information Operations . . . 17

5 Theorizing the Role of Followings & "Fellow Travellers:" From Ego Networks & Audiences to Network Effects **20**

5.1 The top-down broadcast model . . . 20

5.2 The bottom-up social contagion model . . . 22

6 Methodology & Data **24**

6.1 Data . . . 24

6.1.1 Account Selection: A Case Study on Maffick Media . . . 24

6.1.2 Data Collection . . . 29

6.2 Methods . . . 30

6.2.1 Time Series Modelling . . . 30

6.2.2 Coding of Content and Account Type . . . 33

6.3 Measures . . . 38

6.3.1 Follower Count . . . 38

6.3.2 Number of Reposts . . . 40

6.3.3 Size of Repost Reach . . . 40

6.3.4 Reposting by Influential Actors . . . 41

7 Results **43**

7.1 Hypothesis Testing: Time Series Modelling . . . 43

7.2 Account Types & Reposted Content . . . 48

8 Discussion **54**

8.1 Substantive and Theoretical Significance . . . 54

8.2 Methodological Discussion 57
8.3 Practical Implications 59

9 Conclusion **60**

References **63**

10 Appendix **71**
10.1 Figures and Tables 71
10.2 Code Details 76

List of Figures

1 Schematic Depiction of Broadcast vs. Viral Diffusion Models 21
2 Sample Soapbox Stand Instagram Content 26
3 Sample Redfish Stream Instagram Content 27
4 Sample Wasted.Ed Instagram Content 28
5 Locations Tagged in Russian Provocation Posts 36
6 Maffick Media Cumulative Follower Gain on Instagram, 2020-2021 . . 40
7 Schematic Overview of Data Collection and Analysis Pipeline 42
8 Audience Reached through Repost Amplification for Redfish Stream, Wasted.Ed, and Soapbox Stand on Instagram 45
9 IRFs for Redfish Stream 47
10 IRFs for Soapbox Stand 47
11 IRFs for Wasted-Ed 48
12 Number of Reposts of Maffick Media by Account Type and Average Engagement, 2020-2021 49
13 Log of Count of Posts vs. Reposts of Russian Provocation Content by Topic 50
14 Top Words per Topic within Russian Provocation Posts and Reposts 52
15 Correlation Matrix of Repost User Type by Repost Content Theme . 53
16 LSTM Classification Matrix for Instagram Topic Reposts 71
17 Breakdown of Languages in Reposts of Maffick Media Content 74
18 Breakdown of Number of Reposts by Users that Reposted Maffick Media Content 75

List of Tables

S1 Categories for Repost Account Type 35
S2 Categories for Repost and Post Topic Theme 38
S3 Summary Statistics: Mean and Standard Deviation of Variables by Account 43
S4 Breakdown of Reposts, Engagement, and Amplification 44
S5 Granger Causality Tests for Maffick Media Accounts 45
S6 Breakdown of Reposts by Topic, Counts, and Sample Reposted Content 51
S7 Model for Repost Account Type by Govlochenko et al. (2018) 72
S8 Model for Repost and Post Type by Howard et al. (2019) 73

3 Introduction

As issues like police violence and climate change took centre stage in the United States in 2020, social media feeds flooded with social justice activist posts about how to be an ally and where to donate. Within the sea of posts, content from Russian-directed media collectives seamlessly blended into a broader left-leaning[1] ecosystem. Despite attempts to study and expose Russian efforts, Russian provocation accounts continue to maintain their hold within left-leaning, millennial social media ecosystems, amassing hundreds of thousands of followers on Instagram.

As political communication has moved online – a trend that has only been further heightened in a year of lockdowns and isolation – information operations have become more prevalent and sophisticated, ranging from faking grassroots support to artificially boosting visibility on social media (Metaxas & Mustafaraj, 2012; Ratkiewicz et al., 2011). Within the corpus of academic work on this phenomenon, Russian disinformation and misinformation has garnered much academic attention with a variety of in-depth analyses around the 2016 US Presidential Russian disinformation campaign (Howard, Ganesh, Liotsiou, Kelly, & François, 2019). However, while most research on Russian information operations has assessed the tactics and tropes of the Internet Research Agency (IRA), a Russian troll agency that runs major Russian proxy sites (Badawy, Ferrara, & Lerman, 2018; Badawy, Addawood, Lerman, & Ferrara, 2019; Bail et al., 2020; Howard et al., 2019), the fundamental question of how these accounts build the followings upon which their influence is based remains unanswered. Examining how several highly successful IRA Twitter accounts transformed themselves from average Twitter users to microcelebrities, Zhang et al. (2021) highlight the vast number of American Twitter users, including prominent politicians, celebrities, and strategists (O'Brien, 2017), as well as

[1]Given that what constitutes "left-leaning" politics is culturally specific, this thesis takes care to define such terminology. By "left-leaning," this thesis means left-wing politics which support social equality and egalitarianism and critique historical power structures (Clark, 2016). Left-wing politics is often centred around issues where there has been a historical, societal, or structural system of inequality that should be addressed.

members of the news media, mainstream and fringe alike (Lukito et al., 2020), who engaged with these accounts and shared their content. Snegovaya and Watanabe (2021) have described these users that wittingly or unwittingly amplify Russian provocation and disinformation efforts as "useful idiots" or "fellow travellers." However, beyond catching the press' attention, it remains unclear whether reposts have any significance in generating follower growth. Especially in the context of a social network such as Instagram, literature is sparse due to data access limitations through the Instagram API. The question thus rises to the fore:

RQ1: What role – if any – do reposts by "fellow travellers" play in generating follower growth for left-leaning Russian provocation accounts on Instagram?

Building on a novel dataset of Instagram reposts, social network theory, and time-series modelling, I address this gap through a case study of Maffick Media, a state-supported Russian newsgroup that operates channels producing content on issues ranging from climate change to LGBTQ+ activism on Instagram.[2] Drawing on two dominant network dynamic models, namely a bottom-up social contagion model and a top-down broadcast model, I examine how reposts of Maffick Media content by various users, ranging from average Instagram accounts to high-profile corporate accounts, drove follower growth to Maffick Media accounts in 2020-2021. Given these accounts' rapid growth paired with a highly active network of reposting accounts, examining which users are reposting their content and on what topics is of particular interest. To shed light on this question, I complement the analytical research question with a descriptive question that seeks to quantify which users reshare content and on what topics:

RQ2: What kind of content and which accounts reposted Russian left-leaning provocation efforts on Instagram?

The question of whether reposts drove follower growth to these accounts is an

[2]All code is available on Github: [URL removed for submission]. For further detail on the Github repository, see: Appendix 10.2 Code Details

inherently causal one. Epistemologically, causal effects are difficult to measure, particularly within social media literature where experimental manipulation teeters between infeasible to unethical (Grimmelmann, 2015).

By aiming to tackle these challenges, the second main contribution of this thesis is methodological. Building on approaches by Wells et al. (2020) and Zhang et al. (2021), I use time series modelling and a computational quantitative analysis to test a variety of hypotheses about the effects of several variables, from the number of reposts to the kind of user that reposted content, on follower growth. To examine who is reposting Russian provocation content and on what topics, I conduct an intensive manual coding process of account types (N = 1,870) and posts and reposts (N = 4,050). Moreover, the very dataset at the core of this thesis represents the surmounting of a key methodological hurdle in social media research. Given how few studies have analyzed Instagram data within information studies, this thesis thus speaks to not only theoretical field questions but also methodological questions of data access, collection, and management.

The research questions and methodological contributions of this thesis are of relevance for both the academic community and policymakers. From an academic perspective, understanding how Russian provocation accounts build their following through a multiplayer analysis will provide future research with important substantive knowledge on the dissemination of provocation content. For policymakers, a clearer understanding of reposts and their significance beyond the symbolic is a crucial component for calibrating social media regulations; while 2020 marked the first time that social media companies began marking accounts as "state-media" (Bonifacic, 2020), this thesis speaks to the importance of not just highlighting the user but also their content as it spreads from network to network. Moreover, this thesis makes evident the perils of solely relying on fact-checking, which overlooks the ways in which Russian provocation efforts leverage factually accurate but incendiary information with a covert agenda to exacerbate divisions.

This thesis proceeds as follows: Chapter 4 derives the framing of Russian influence operations as a strategy of provocation and polarization by situating current efforts within a history of Russian *aktivnye meropriyatiya* (active measures). I review previous literature on the Russian Fringenet, proxy voices that proliferate state narratives without direct attribution to the country, and motivate the analysis of left-leaning provocation content on Instagram. Chapter 5 conceptualizes two different potential network dynamic models for the development of social media followings: 1) The bottom-up social contagion model driven by a homophilous network; and 2) The top-down broadcast model driven by influential users. After outlining the data collection process, Chapter 6 details the two-pronged approach of leveraging time-series modelling to estimate the effect of reposts of Russian provocation content on follower growth and manual coding of repost accounts (N = 1,870) and posts and reposts (N = 4,050) to analyze which users were disseminating Russian provocation content and what topics were reposted. Chapter 7 and Chapter 8 present the empirical results on the analytical and descriptive questions posed and situate these findings within a broader Russian provocation and polarization literature. Finally, Chapter 9 concludes with a critical reflection on limitations and perspectives for future research.

4 Russia's Provocation Strategy: The Fringenet, "Fellow Travellers," and Followers

To situate my motivation for examining reposts of left-leaning Russian provocation accounts on Instagram, this section is divided into three main sections: the Fringenet, "fellow travellers," and followers on Instagram. First, I derive the framing of Russian influence operations as a strategy of provocation by contextualizing present-day efforts within a broader history of Russian *aktivnye meropriyatiya*. With the aim of provocation and polarization, I describe how Russia has leveraged the Russian Fringenet to build followings and disseminate content to various audiences, mixing "pop" news content with political commentary. Second, I highlight the role of "fellow travellers," users who repost and reshare Russian provocation content to their followers. Third, I explain why the accumulation of a social media following is key to the influence and reach of an account, building on previous work on Russia's information operations on Twitter and Facebook.

4.1 Russian Information Operations: The Fringenet's Strategy of Provocation & Propaganda

Though often framed in the discourse as unprecedented, Russian information operations should be contextualized within the USSR's historical strategy of *aktivnye meropriyatiya* (active measures). During the Cold War, active measures were covert offensive instruments of Soviet foreign policy that aimed to systematically disrupt relations between other nations, discredit Soviet opponents, and influence policies of foreign governments in favour of Soviet plans and policies (Walton, 2016). Such active measures included: media manipulation, front groups, forgery, and influence operations to discredit targets. For example, the USSR exploited racial tensions in the United States by sending hundreds of hate letters to African missions allegedly penned by average US citizens and white supremacists. The letters, which were filled

with racist language, were publicized as examples of how racism was still rampant in the US and received significant press attention by US and foreign press corporations (Selvage, 2019; Abrams, 2016).

By contextualizing Russian influence operations within the history of *aktivnye meropriyatiya*, one can inform present theories around Russian influence operations on social media platforms. With rising concern over the scope, speed, and scale (Silva, Giovanini, Fernandes, Oliveira, & Silva, 2020) of Russia's influence operations on social media, a variety of political science theories have been advanced. For instance, Paul and Matthews (2016, 1) argue that Russia's influence strategy is based on a "firehose of falsehood" propaganda model characterized by "high numbers of channels and messages and a shameless willingness to disseminate partial truths or outright fictions." Meanwhile, Walker and Ludwig (2020) contend that Russia's behaviour is best explained as "sharp power," a contrast to Nye's "soft power"[3] where a country seeks to "target audiences by manipulating or poisoning the information that reaches them" (Cardenal, Kucharczyk, Mesežnikov, & Pleschová, 2017, 13).

However, both the "firehose of falsehood" model and "sharp power" model leave unexplained Russian influence operations that are not composed of disinformation. By disinformation, I mean false context, impostor, fabricated, or manipulated content with the intent to misinform and to cause some kind of political, financial, social, or psychological harm (Wardle & Derakhshan, 2017; De Cock Buning, 2018). As Keating and Kaczmarska (2018) highlight, while Russia has used disinformation as a foreign policy tool, information operations do not always fall within the simplified binary of "true" and "false." Rather, Russia has also followed a strategy of provocation, one in which factually accurate but provocative information is disseminated to specific, niche audiences with the intent of sowing dissent or exacerbating tensions. This strategy is best understood by contextualizing it within earlier *aktivnye meropriyatiya* efforts: for example, Arif et al.'s (2018) insightful study of

[3]According to Nye, "soft power" is the ability of a country to influence others without force or coercion. For further detail see: Nye, J. S. (1990). Soft power. Foreign policy, (80), 153-171.

the IRA's penetration into conversations around the #BlackLivesMatter movement in 2016 finds fertile conceptual ground in drawing a parallel to earlier efforts to exacerbate racial tensions in the US by sending hate letters to African missions. Therefore, while there certainly may be elements of a "firehose of falsehood" model or "sharp power" model to current Russian efforts, this thesis will frame the discussion of Russian influence operations less as a conversation of disinformation and misinformation but rather as a broader strategy of provocation and polarization.

Within its strategy of provocation, Russia has leveraged the "Fringenet," proxy media voices that proliferate pro-Russian messages online (Global Engagement Center, 2020). Specifically, the Fringenet is a collection of nominally independent news sites, ranging from conspiracy sites to far-right and far-left "think tanks" to "data dump websites" (Weisburd, Watts, & Berger, 2016), that proliferate pro-Russian messages. Consistently, a number of Russian Fringenet sites present themselves as think tanks or independent news organizations, omitting any mention of their Russian funding or affiliation. By presenting as "objective" news organizations, these sites are posed to mix light news and clickbait content with underlying pro-Russian political commentary (Snegovaya & Watanabe, 2021). While certain content pushed by the Russian Fringenet easily falls under my definition of disinformation as fabricated or manipulated information, other outlets publish factually accurate content under a false premise or mix accurate and biased reporting. For example, on the far-right, the Newsroom for American and European Based Citizens, an IRA-affiliated site, focused primarily on pushing articles on politically sensitive issues within US politics with a conservative bent. The site embedded itself within a broader conservative ecosystem and relied on a network of accounts posing as journalists to share its articles on right-wing social media sites (Stubbs, 2020). In contrast, Peace Data, a Russia-linked network targeting progressive and left-wing audiences in the US and the UK, presented itself as a news site that focused largely on the environment and corporate and political corruption. While some of Peace Data's freelance

journalists were real reporters, others were personas masquerading as left-wing journalists and editors (Collier & Dilanian, 2020; Nimmo, 2020). Peace Data's network disseminated articles highlighting racial tensions in the US, accusations of foreign interference and war crimes committed by the US, and capitalism-induced suffering (Nimmo, 2020).

For a strategy of provocation, the Russian Fringenet has three key advantages that derive from their sheer number, nominal independence, and media multiplier effect (1055782, 2021). First, the sheer number of Fringenet outlets allows for wide dissemination of alternative variations of a narrative. Each outlet can cater and fine-tune its narrative to a specific audience, thereby allowing for a kind of micro-targeting (Hindman & Barash, 2018). Second, as nominally independent outlets, the Russian Fringenet sites can shield the state from direct attribution while disseminating blatant or dangerous disinformation. Similarly, nominal independence allows the Russian Fringenet to quickly respond to developing news stories and to take advantage of the fast-paced media cycle as there is no state approval process required (Dawson & Innes, 2019). Third, given the expansive network of sites, the Russian Fringenet can easily create a media multiplier effect by amplifying each other's articles and posts. While this can lead to a "firehose of falsehood" model where accounts create "disinformation storms" by resharing each other's content (Global Engagement Center, 2020), these actors can also create an illusion of consensus around provocative but subjectively accurate narratives and monopolize the information environment.

Though both far-right and far-left Russian Fringenet sites use similar strategies, such as relying on a network of journalist accounts to disseminate their content and mixing pop content with political commentary, the former has been more closely studied than the latter. Indeed, while previous research has demonstrated that a right-wing media ecosystem is particularly susceptible to disinformation (Benkler, Faris, & Roberts, 2018; Hjorth & Adler-Nissen, 2019), research into the IRA's ef-

forts in 2016 has also brought to the fore content targeting marginalized or minority groups in the US, including BIPOC and LGBTQ+ communities, on left-leaning issues, such as gay marriage or transgender rights (Howard et al., 2019). Illustrating this focus, the most high-engaging IRA efforts on Facebook and Instagram from 2014-2017 specifically targeted Black American communities around issues of historical oppression and present-day discrimination (Howard et al., 2019). Similarly, LGBTQ+ voters were extensively messaged (Howard et al., 2019) with content that portrayed certain religious groups or segments of conservative voters as homophobic or transphobic.

To date, research on information operations has often been limited to Facebook and Twitter studies. In part, this trend may be attributed to the public release of 2,752 Russian IRA Twitter accounts by the US Congress as part of its official investigation into the US Presidential 2016 election (Select Committee on Intelligence, 2019). Additionally, the open nature of the Twitter and Facebook APIs allow for a larger and more intensive sample collection, unlike Instagram data, which is not as readily accessible. In 2020, Instagram restricted access to its API, concluding a two-year gradual process of restricting platform data for third-party actors (McCrow-Young, 2021). As a consequence, research across a variety of fields on the platform has been notably scarce (Maares & Hanusch, 2020). The limited number of Instagram studies inevitably creates a gap in the literature. Such a gap is particularly significant given not only the strong reach of Instagram among a young adult audience (Pew Research Center, 2021) but also evidence suggesting that Instagram was perhaps the most effective platform for the IRA in 2016 (DiResta, Shaffer, Ruppel, Sullivan, & Matney, 2019). To the former, a majority of 18- to 29-year-olds in the US self-report that they use Instagram (71%), while just 42% use Twitter (Pew Research Center, 2021). To the latter, Instagram was a significant front in the IRA's influence operation (DiResta et al., 2019) that reached over 187 million engagements in comparison to 126 million engagements on Facebook. In short, given the popu-

larity of the platform combined with the success of past Russian efforts, Instagram has thus been understudied proportional to its centrality in Russian operations.

4.2 Fellow Travellers: Networks of Amplification

Though the Russian Fringenet authors provocative content on social media platforms, the dissemination of such content requires a much more complicated, multiplayer analysis. Indeed, the reach of Russian Twitter accounts is highly supported by followers who propagate their narratives: "these users can then be exploited by the Kremlin operations as 'useful idiots' or 'fellow travellers' and (unwittingly or not) help amplify narratives that benefit Russia" (Snegovaya & Watanabe, 2021, 35). Shedding light on how the most successful IRA Twitter accounts built their followings from 2014-2017, Zhang et al. (2021, 306) found that large Twitter followings "facilitated the direct dissemination of strategic messages by serving as their audiences," "functioned as a set of individuals likely to amplify their messages through retweeting and other means," and "provided visible engagement metrics that looked to other observers, including journalists, like evidence of authority and authenticity."

A wide range of agents on a spectrum from inauthentic, coordinated accounts to authentic, inadvertent followers are involved in the dissemination of Russian provocation efforts (Tucker et al., 2018). On the far end of the spectrum, literature has demonstrated how Russian provocation and disinformation is spread through networks of inauthentic accounts posing as journalists, editors, or outlets. Mining mentions and retweets to chart how Russian disinformation narratives spread through social media ecosystems, DiResta and Grossman (2019) identify numerous suspect accounts, which appear to freelance consistently with a web of "alternative news" sites. According to DiResta and Grossman (2019), these accounts typically claim to be independent freelance journalists or academics, have a limited social media presence with just one or two photos, and actively publish articles on a number

of "alternative news" sites.

On the other side of the spectrum, hyper-partisan media sites which interact and share Russian Fringenet pieces play a prominent role in dissemination. Disinformation and misinformation stories (Faris et al., 2017) from hyper-partisan platforms are much more widely shared on Facebook than articles from purely disinformation news websites. Similarly, established news media, including mainstream outlets, can act as megaphones for these narratives (Arceneaux & Johnson, 2015); by recirculating and thereby validating Russian provocation narratives, mainstream outlets can inadvertently incorporate foreign information operations into their news coverage (Tucker et al., 2018).

Beyond institutions and organizations, individual authentic actors on social media sites can often serve as conduits for Russian provocation narratives. Not only has Russia increasingly recruited Americans on both sides of the political spectrum to write articles and posts for these websites that indirectly align with Russia's agenda (Snegovaya & Watanabe, 2021), but also thousands of US Twitter users, including prominent politicians and public personalities, interacted with and shared IRA efforts in 2015 (O'Brien, 2017). While much of the limelight has been centred on Russian sites, bots, and fake social media accounts, average users have become empowered in information ecosystems through connective actions based on personalized content sharing (Bennett & Segerberg, 2012). As a result, these days, building an audience no longer requires access to traditional media. While such empowerment is key for online activism and coalition building (Mundt, Ross, & Burnett, 2018), "decentralized networks of people" (Boyd, 2017) can also be used for ill, as in the case of Russia's strategy of harnessing average users to disseminate Russian provocation narratives.

4.3 Following: Assembling Networks and Audiences of Information Operations

Having situated the Fringenet and "fellow travellers" within a broader strategy of provocation, I explore why the accumulation of a social media following is key to the influence and reach of an account. Informed by Zhang et al.'s (2021) model, I conceive of the social media following of Russian Fringenet accounts as both an ego network and an audience. As an ego network, each account (ego) maintains ties (edges) with other accounts (alters). When one alter reposts the ego's content, the alter thereby disseminates the original ego's content to its own network.

On Instagram, a repost takes the form of a tag, where one user reposts content shared by another account and provides a source by tagging the original user. By tagging a user, accounts not only share the original content, thereby broadening the reach of a message, but also may bring more followers to the original ego through transitivity. Under principles of network theory, transitivity is the degree to which, if A follows B and B follows C, then A follows C. On Instagram, user B reposting user A's content may lead user B's followers (i.e. user C) to follow user A. In theory, reposted content provides social proof (Hutto, Yardi, & Gilbert, 2013) to user C that user A may be worth following, thus leading to triadic closure. In sum, an ego's followers allow for access to their own networks of social ties, easily generating a multiplier effect that can further increase the ego's influence and reach (Susarla, Oh, & Tan, 2011).

Beyond acting as an ego network, an account's social media following is also its audience. Examining former President Donald Trump's Twitter activity around the 2016 US Presidential election, Wells et al. (2020) develop the idea of an "attentive public." According to Wells et al. (2020), attentive publics are inherent to the contemporary media moment; attentive publics are products of a media system that combine "the attention-seeking products of numerous media structures, networking platforms that enable entities to assemble relatively stable followings, the exercise of

personal preferences by thousands of individuals, and the strategic actions of those who seek their own attentive audiences" (Wells et al., 2020, 185). An attentive public has a number of benefits for the account, including allowing direct and stable access to an audience without the traditional gatekeeping of mainstream media (Gainous & Wagner, 2013; Graham, Broersma, Hazelhoff, & van 't Haar, 2013; Parmelee & Bichard, 2011).

Moreover, beyond the direct audience, the visibility of attentive publics, usually displayed in following, retweets, or a "verified status," can be used to gauge a user's popularity, authority, and connectivity (Fu, 2016), which in turn translate into influence and power (Wells et al., 2020). A large following increases perceived credibility and trustworthiness (Jansen, Zhang, Sobel, & Chowdury, 2009) and signals social status and influence (Marwick & Lewis, 2019). The visibility of a verified check, for example, acts as a visual cue that an account – and its content – is worthy of further attention and can easily garner further attention through reposts that amplify its message (Wells et al., 2020; Tufekci, 2013). In this way, influence begets influence; the ability to capture an audience's attention translates into further communicative power in an attention economy (Webster, 2014).

Finally, the accumulation of an attentive public on platforms like Instagram has the further advantage that social media platforms' algorithms bias toward popular content. In essence, being able to command attention is rewarded with more traffic by popularity-driven algorithms (Gillespie, 2011). The market logic of a social platform like Instagram creates an economy of influence that self-perpetuates (Zhang et al., 2021); beyond popularity-driven algorithms, gaining more followers only makes more users more likely to follow such an account given social status signalling, consequently forming a positive feedback loop (Zhang et al., 2021).

Therefore, by treating an account's social media following as both an ego network and an audience, we can understand why a Russian Fringenet's social media following comes hand in hand with information dissemination and influence. However,

the question remains how accounts build such followings.

5 Theorizing the Role of Followings & "Fellow Travellers:" From Ego Networks & Audiences to Network Effects

In this section, I consider two different potential network dynamic models for tie formation in the development of social media followings: 1) The bottom-up social contagion model driven by a homophilous network; and 2) The top-down broadcast model driven by influential users (Zhang et al., 2021; Goel, Anderson, Hofman, & Watts, 2016). Figure 1 schematically illustrates these two modes of distribution (broadcast vs. social contagion); the former is characterized by a significant number of adoptions from a single parent node while the latter is characterized by a multi-generational branching process, in which any one node directly translates into only a few other adoptions (Goel et al., 2016). Both models have previously been applied to the dissemination of information and resulting follower gain on social media (Zhang et al., 2021; Wells et al., 2020). Critically, the two are not incompatible, though each has its own implications for how Russian provocation content spreads and generates potential follower growth. Therefore, after exploring the theoretical applicability of both models, I develop three hypotheses based on the two models, which I later test in Section 6.2.1 Time Series Modelling.

5.1 The top-down broadcast model

Turning to the first model, the top-down model is characterized by "opinion leaders" or "influential" users who act as broadcasters to their networks. Katz and Lazardsfeld (2017, 3) originally defined opinion leaders as "the individuals who were likely to influence other persons in their immediate environment." Though Katz and Lazardsfeld's (2017) definition originates from a pre-social media age where the large-scale distribution of information relied on broadcast media, opinion leaders still translate empirically onto social media networks. For example, Goel et al. (2016) found that popular tweets usually spread through a broadcast model in a top-down manner.

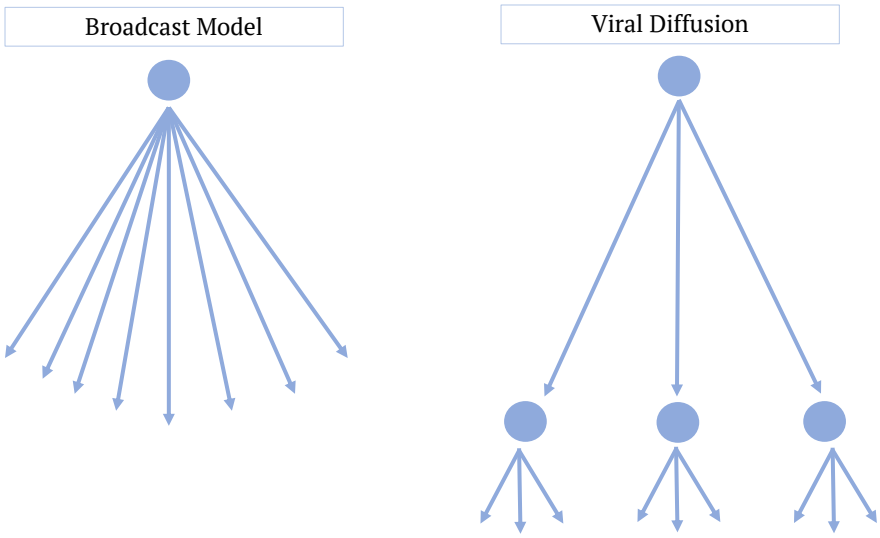


Figure 1: Schematic Depiction of Broadcast vs. Viral Diffusion Models

In the context of Russian left-leaning provocation accounts, the top-down broadcast model driven by influential users suggests that followers will flow from gaining the attraction of high-profile users with sizeable followings to repost their content. As Russian left-leaning provocation accounts turn influencers, politicians, and the like into amplifiers, more Instagram users will both be exposed to their messages and become more likely to follow due to the follow-worthiness signaled by the uptake by an established account. Russian accounts have often garnered prominent followers and reposts; high-profile Trump supporters like Donald Trump Jr. as well as far-right media figures like Ann Coulter and Dinesh D'Souza repeatedly retweeted one of the most prominent IRA accounts (@TEN_GOP) during the run-up to the 2016 US Presidential Election (O'Brien, 2017). Through this top-down network, such influential actors bridged the tie between an IRA account and future followers by signalling to their own followers that @TEN_GOP was an authentic and credible source. Thus, based on this model, I hypothesize that influential actors reposting Russian left-leaning provocation accounts will increase the accounts' followings:

H1: Influential accounts reposting Russian left-leaning accounts' content will lead to an increase in Russian accounts' followings.

5.2 The bottom-up social contagion model

Turning to the second model, the bottom-up social contagion model is defined by a highly homophilous or homogenous network. Examining the diffusion of scientific information and conspiracy theories, Del Vicario et al. (2016) find that homogeneity is a driver for the diffusion of information. Similarly, Aral et al. (2009) argue that homophily can explain over 50% of the perceived behavioural contagion in a network. Building from decision-making literature across a variety of fields including economics, social psychology, and marketing, Watts and Dodds (2007) argue that binary decisions exhibit positive externalities, meaning that the probability that an individual will choose a certain alternative increases with the relative number of others choosing that alternative. Watts and Dodds' (2007) theory has a number of explanations, ranging from network effects (Liebowitz & Margolis, 1994) to learning from others (Bikhchandani, Hirshleifer, & Welch, 1992; Salganik, Dodds, & Watts, 2006) to conformity pressures (Bernheim, 1994). Under Watts and Dodds' (2007) theory, a certain threshold of average users on a social media platform may thus be more generative than a few highly influential users. Within Russian influence operation studies, the role of average users has been spotlighted by previous studies. For example, examining content on Twitter around the Malaysian Airlines passenger flight (MH17) downing, Golovchenko et al. (2018) find that that average citizen accounts were the most active drivers of both disinformation and counter-disinformation.

In the context of Russian left-leaning provocation accounts, the bottom-up social contagion model suggests that followers will flow from gaining like-minded, average users to repost their content. In other words, if a Russian left-leaning alternative news site was looking to build a following, it may be most efficient and effective

to build an attentive public of average users, who "make up the critical mass via which local cascades become global" (Watts & Dodds, 2007, 446). As Russian left-leaning provocation accounts turn average reposting users into amplifiers of their provocative and polarizing content, the accounts gain more followers in a bottom-up fashion. Based on this model, I hypothesize that the more reposts Russian left-leaning accounts gain, the more followers they accumulate (H2); and that the greater the size of the reposting accounts' follower networks, the more followers Russian provocation accounts gain (H3):

H2: The more reposts Russian left-leaning accounts gain, the more followers they accumulate.

H3: The greater the size of the reposting reach of accounts, the more followers Russian left-leaning accounts will gain.

6 Methodology & Data

This section outlines how I leverage time series modelling to examine the relationship between reposts of Russian provocations efforts on Instagram and follower growth. I first describe this thesis' case study of three Russian provocation accounts run by Maffick Media. I then present the research design for collecting a year's worth of data from Instagram using two unofficial Instagram APIs.

I next lay out my two-pronged method for RQ1 and RQ2. To probe at RQ1, I describe how I replicate Wells et al.'s (2020) and Zhang et al.'s (2021) Twitter time-series modelling, specifically using a vector autoregression model to determine the time-ordered relationship between a variety of independent variables and the dependent variable (followers gained by each account). I use Wald two-way Granger causality tests and impulse response functions (IRFs) from the VAR models to provide further information regarding the longer-term effects. I review how I operationalize each potential independent variable including: follower count, number of reposts, size of repost reach, and reposts by influential actors. To probe at RQ2, I describe how I combine manual coding of the repost accounts (N = 1,870) with a random sample of 50% of the reposts and posts (N = 4,050) to analyze which users were disseminating Russian provocation content and what topics were reposted. The full data collection and analysis pipeline is summarized in an illustrative flow chart in Figure 7 at the end of this section.

6.1 Data

6.1.1 Account Selection: A Case Study on Maffick Media

To examine left-leaning Russian provocation accounts, I focus on a case study of Maffick Media, one of the most successful left-leaning, Russian state-backed media companies on Instagram. Incorporated in 2018, Maffick Media runs a network of nominally "independent" media productions. Though Maffick Media's accounts

present as independent, left-leaning pages, Maffick Media's history and finances reveal close ties to the Russian state-run media outlet RT. Maffick Media's company registration documents show that the majority shareholder is Ruptly, an RT subsidiary in Berlin. Ruptly's goals include "to build and extend the core strengths and values of [its] parent company RT" (Hanlon & Morley, 2019); according to RT's Editor-in-Chief, such strengths include serving as an "information weapon" for Russia (DFR Lab, 2018).

Maffick Media runs a number of channels on Instagram, producing content on issues ranging from climate change to LGBTQ+ activism. Its channels are aimed at a millennial, left-leaning audience. Prior analyses of Maffick Media by the press (O'Sullivan, Griffin, Devine, & Shubert, 2019) have described its content as "generally critical of US foreign policy and the mainstream American media, while largely avoiding criticism of the Russian government" and taking "strong, often-fringe political stances." Since the media called attention to Maffick Media's ties in 2019, Maffick Media's accounts have been labelled as Russian state-backed entities by Facebook, Twitter, and YouTube. However, the media company has nevertheless continued to grow its social media following; on Facebook alone, Maffick Media's accounts have over 6.4 million followers combined.

To select three of Maffick Media's channels, I set certain data collection as well as attribution criteria. In terms of data collection criteria, accounts must be mid-sized, public Instagram accounts. All accounts must leave public any tagged content, meaning I can collect all reposts of these accounts. In terms of attribution criteria, all three accounts must have a clear connection to Maffick Media, either run by the media conglomerate or closely tied to its operations (e.g. same building, same IP address, etc.). Additionally, all accounts must be marked by Instagram as "Russian state-controlled media." To provide further detail on the sample of accounts that met these criteria, I include brief overviews of each of the three channels.

Soapbox Stand

Soapbox Stand is a general media channel run by Maffick Media. Soapbox Stand's content often takes strong, left-leaning political stances on contemporary social and political issues; the channel's content is "packaged as meme-able satire and no-nonsense takes on history, environmental issues, and sensitive global politics" (Hanlon & Morley, 2019). While Soapbox Stand presents as an independent media channel with no connection to the Russian government, reporting by the German press (T-Online, 2019) as well as the Alliance for Securing Democracy (2019) reveal close ties to Russian-controlled media in the channel's history and financial filings. On Instagram, Soapbox Stand has garnered 53,100 followers with sample content such as:

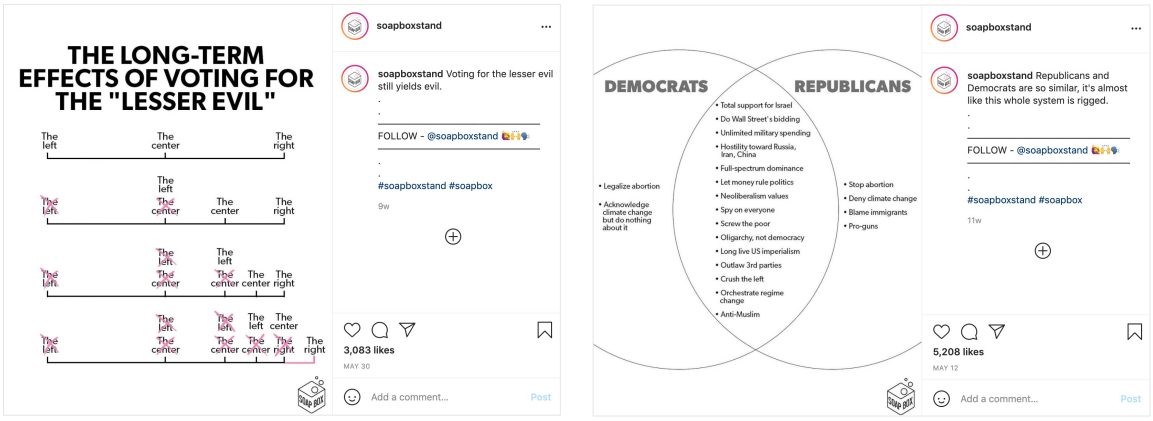


Figure 2: Sample Soapbox Stand Instagram Content

Redfish Stream

Redfish Stream is a media channel that has amassed nearly 500,000 followers on Instagram alone with its far-left political content. Describing itself (Redfish Stream, 2021) as a "new kind of media company...aimed to inspire social and economic change," the outlet received recognition from Vice for its reporting on the 2017 Grenfell Tower Fire in England, a catastrophe that killed over 70 people in a public housing development and catalyzed public outcry on social inequality (Redfish Stream, 2021). According to Vice, Redfish Stream was a "fantastic example of amateur community-produced media" (Davis, 2018). While Redfish Stream presents as an independent outlet, Redfish's registration documents show that the company's sole shareholder is Ruptly TV. Moreover, Redfish Stream's staff includes several former RT reporters (Davis, 2018), and the company is headquartered in the same office as Maffick Media and Ruptly TV in Berlin. On Instagram, Redfish Stream has 421,000 followers with sample content, including:



Figure 3: Sample Redfish Stream Instagram Content

Get Wasted.Ed

Get Wasted.Ed is a sustainability-focused media channel run by Maffick Media. The account covers a variety of issues under sustainability, including critiquing capitalism for its consumerist culture, highlighting the sociopolitical inequalities of climate change, and advocating for green reform, such as a plastic ban. Similar to RT, Get Wasted.Ed archetypically mixes "light news" clickbait videos with anti-US and pro-socialist political content (Dilanian & Ramgopal, 2020). Get.Wasted.Ed's account has 404,000 followers; its content is English-based with a sustainability focus, such as:

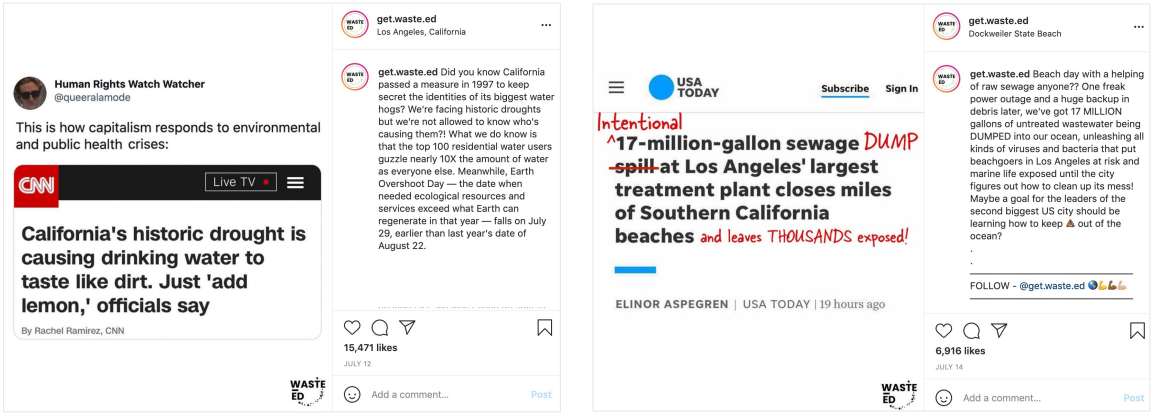


Figure 4: Sample Wasted.Ed Instagram Content

6.1.2 Data Collection

Due to the limitations of the Instagram API itself, I use two unofficial Instagram APIs[4] to build the dataset. While literature reflecting on accessing, archiving and anonymizing Instagram data often suggests unofficial APIs as a viable alternative (McCrow-Young, 2021), such literature has also highlighted the paywalls to most third-party sites that widen access inequalities. The unofficial Instagram API selection came out of reviewing dozens of repositories on Github, as well as literature adjacent to Russian information studies that have leveraged these unofficial APIs on previous research (Vargas, Rivera-Rogel, & Carrión-Salinas, 2020; Thejas, Kumar, Iyengar, Badrinath, & Sunitha, 2019; Fluegel & King, 2021). The selection of unofficial Instagram APIs for this thesis is based on two criteria: 1) Past work that demonstrated the viability of these unofficial Instagram APIs; and 2) No paywall to demonstrate the possibility of and encourage further equitable and accessible research. By using these two unofficial Instagram APIs, I collect all posts and reposts of three Maffick Media Instagram accounts from January 1st, 2020 – December 31st, 2020 as well as a daily count of followers gained.[5]

However, I am limited by the data available via unofficial Instagram APIs, which introduces a variety of potential errors discussed in 7.2 Methodological Discussion. Due to data collection limits of both unofficial Instagram APIs, I was limited by how much data I could scrape from Instagram at a time without being blocked by Instagram. As a result, I limit the time frame of the analysis to twelve months. Furthermore, because the APIs only include data on certain accounts and for certain dates, I could not examine all of Maffick Media's accounts nor could I collect the same date range for all of the accounts. In the context of this thesis, while I was able to collect all post and repost data for Soapbox Stand, I could only access follower count data from December 14, 2020 to the present. As a result, the date range for

[4]For further detail on the unofficial Instagram APIs used to collect post and repost data as well as follower count: https://hub.phantombuster.com/docs/api & https://socialblade.com

[5]For further detail on the code used to collect this dataset, see Appendix: Code Details.

Soapbox Stand spans from December 2020 – June 2021, while the date range for Wasted.Ed and Redfish Stream span from January 2020 – December 2020. While the date ranges are comparative – not only overlapping but also only extending six months further –, this data collection limitation reduces the sample size for Soapbox Stand as the dataset is comprised of six months of data as opposed to twelve months.

I collect all details on posts and reposts for these three accounts, including: post URL, description, number of comments and likes, tagged location, date, type (e.g. video or photo), image URL, and tagged accounts. Due to the nature of the dataset and the platform, not all reposts can be collected. Specifically, I can only collect public tags within the Instagram feed, which gives rise to two major limitations: 1) Private tags; and 2) Story tags. First, if an account privately tagged any of the three accounts studied or an account chose to hide certain tags, this repost would not be included in the data set. Second, if an account reposted content to their story rather than their feed, this tag would not be included in the data set. While the former limitation is more negligible as private accounts tend to have a smaller following, the latter limitation is more critical as Instagram stories is a main feature on the platform. I discuss the possible consequences of this nonresponse or missing data error in 7.2 Methodological Discussion.

In total, the final dataset includes 8,098 Instagram posts – 4,566 reposts and 3,532 posts by Maffick Media accounts – and 1,870 accounts that reposted Maffick Media accounts.

6.2 Methods

6.2.1 Time Series Modelling

To test whether reposts by "fellow travellers" generate follower growth for left-leaning Russian provocation accounts, I replicate Wells et al.'s (2020) and Zhang et al.'s (2021) vector autoregression (VAR) model analysis in R. Examining former President Trump's Twitter account, Wells et al.'s (2020) pair VAR models with

Granger causality tests to demonstrate how Trump's attentive public developed from social media amplification, media exposure, and significant election events. Building off Wells et al.'s (2020) work, Zhang et al. (2021) examine how four IRA accounts built their followings on Twitter by pairing VAR models with Granger causality tests and impulse response functions (IRFs).

An extension of univariate autoregression, VAR accounts for multivariate data by using vectors. In a VAR model, each endogenous variable is assigned to its own linear equation, which includes lagged values of the variable itself, lagged values of all other variables in the model, and an error term. The general VAR equation with one variable and n time lags can be written as (Chiovaro, Windsor, & Paxton, 2021):

$$y_t \;=\; c \;+\; \Phi_1 y_{t-1} \;+\; .. \;+\; \Phi_n y_{t-n} \;+\; e_t \tag{1}$$

In the above equation, y is the variable, c is a constant, Φ are the weights of the terms, and e is an error term. A first-order VAR with two variables can be expressed as (Chiovaro et al., 2021):

$$y_{1,\,t} \;=\; c \;+\; \Phi_{1,1} y_{1,\,t-1} \;+\; .. \;+\; \Phi_{1,2} y_{2,t-1} \;+\; e_{1,t} \tag{2}$$

$$y_{2,\,t} \;=\; c \;+\; \Phi_{2,1} y_{1,\,t-1} \;+\; .. \;+\; \Phi_{2,2} y_{2,t-1} \;+\; e_{2,t} \tag{3}$$

VAR models are extensively used for applied modelling and forecasting in which each variable is explained by its own past values and the past values of all other variables in the system (Groshek, 2011). While VAR models are frequently used in network studies (Wells et al., 2020), VAR models have a number of limitations, particularly due to the large number of parameters to be estimated. As the number of variables modeled by the VAR increases, the number of parameters to be estimated grows quadratically; for example, if a model contained 6 variables and a lag

length of 8, there would be 48 coefficients on the lagged variables plus a constant term to be estimated (Sims, 1980). Issues of multicollinearity between the different lagged variables can also lead to imprecise coefficient estimates; a large number of parameters can result in a good within-sample fit but poor forecasting accuracy outside-sample (Litterman, 1986).

To use a VAR model, I treat each independent variable as endogenous to capture the interdependencies between variables and account for all potential relationships. Given concerns of multicollinearity with VAR models, I limit the number of variables in the model (6.3 Measures). Before modelling the data, however, I first difference the time series to find any non-stationary time series. The primary assumption of VAR models is that any time series is stationary, meaning that the underlying joint probability distribution, mean, and variance remain the same over time (Chiovaro et al., 2021). Non-stationary time series or integrated components can result in inaccurate estimates (Box-Steffensmeier, Freeman, Hitt, & Pevehouse, 2014). Beyond stationary time series, variables in VAR models must also be sampled consistently at a regular interval and be of the same scale.

To assess the stationarity of the time series, I use an Augmented Dickey-Fuller (ADF) test and difference non-stationary time series. Having differenced the time series, I use Granger causality tests to estimate whether lags of one variable can be used to forecast another. In Granger causality tests, vector error correction (VEC) models are assessed to determine the number of lag terms needed to find the best fitting model for the unrestricted regression. To select the number of lag differenced terms, I select the equation with the minimum Bayes information criterion (BIC), where $\hat{\sum}_u$ is the estimate of the k x k covariance matrix of the VAR errors and det (.) is the determinant:

$$BIC\,(p) \;=\; \log\left[\det\;\left(\hat{\sum}_u\right)\right] + k(kp\;+\;1)\;\frac{\log\,(T)}{T} \tag{4}$$

Taking the minimum BIC as the appropriate lag lengths, I use Wald two-way

Granger causality tests, a bivariate test of causality in which the restricted and unrestricted regressions are estimated to determine the F statistic and respective p-values. A rejection of the null hypothesis would mean that the first series (for example, number of reposts of Russian provocation content) Granger-causes the second series (for example, number of followers gained).

To provide further detail regarding the longer-term effects of one variable on the dependent variable, I also use impulse response functions (IRFs) from the VAR models (Swanson & Granger, 1997); the IRFs test the impact of a change in, for example, the number of influential users reposting content on follower growth. Controlling for other variables in the model, the IRFs capture the relationship between one variable and a standard deviation increase in another variable. I present both the results of the Wald two-way Granger causality tests and the relevant IRFs in 7. Results.

6.2.2 Coding of Content and Account Type

Turning to RQ2, I examine what kind of accounts reposted Russian left-leaning provocation efforts as well as what kind of content was reposted. To determine which accounts were reposting Russian left-leaning provocation as well as which content was reposted, I manually code all accounts by type and a random sample of 50% of the reposts and posts by topic. Before manually classifying the data, I first examined the possibility of leveraging a statistical learning model, such as a Long Short Term Memory model, to classify all of the Instagram posts, reposts, and accounts; however, I found that standard classifiers had insufficient precision and recall to be useful.[6] In part, this is likely due to the significance of both the

[6]To test whether a statistical learning model could be employed, I piloted a Long Short-Term Memory (LSTM) model on the repost data. Having tokenized the corpus using nltk's tokenizer package, I trained Fast Text embeddings on the corpus as the reposts contain non-standard names and words that denote high meaning (e.g. "George Floyd" or "Dr. Fauci"). From the dataset, I conducted a 40-60 test-train split for the LSTM model. The LSTM model consists of an embedding layer, an LSTM layer, and a linear layer with a dropout (.2) between all layers. By fine-tuning the hyperparameters, I achieve an 83% accuracy after 19 epochs. However, the model had a low precision and recall, as seen in the confusion matrix included in Appendix Figure 16. I include

image and text in deriving meaning from an Instagram post as well as details in a user's profile, such as their image or the number of posts on their page, that provide a subtext for classification. For these reasons, past studies (Hjorth & Adler-Nissen, 2019; Golovchenko et al., 2018) of Russian information operations have often relied on manual coding as opposed to automated analysis. However, such intensive manual coding is not only costly but introduces subjectivity, a limitation which is further discussed in 8.2 Methodological Discussion.

To code the data, I categorize all accounts that reposted Maffick Media accounts in 2020, resulting in a dataset of 1,870 accounts. In developing the codebook of account type, I adapt Golovchenko et al.'s (2018) codebook (Appendix Table 7) that was used to analyze Twitter accounts that disseminated disinformation around the MH17 downing. I pilot Golovchenko et al.'s (2018) codebook on a random sample of 50 accounts that have reposted Russian provocation content. Based on the pilot analysis, I condensed certain categories from Golovchenko et al.'s (2018) codebook as well as added certain account types that were more relevant to this dataset, such as activists. Having fine-tuned the codebook into nine possibilities, I work with a social media manager and marketing assistant to code 1,870 accounts. Prior to coding, the coder was made familiar with the material and the codebook. Given her previous experience in the industry as well as work on similar projects, she was uniquely positioned to aptly categorize these accounts. I double coded a random sample of 50 accounts to assess the inter-rater reliability of the coding procedure. The interrater reliability for the account type, captured with an unweighted Cohen's Kappa, was .61. According to McHugh (2012), this value is considered "moderate" in agreement. The final codebook of account types includes:

details of this LSTM model as further support for the decision to use manual classification beyond field norms.

Table S1: Categories for Repost Account Type

Group	Description
Activist or Civil Society Groups	NGOs, grassroots organizations, social movement accounts, non-profits, and activists
Citizens	Individual users who are not journalists, politicians, or public officials (includes users whose profiles lack self-description)
Citizen Journalist/Alternative or Independent News Outlet	Self-described journalists, bloggers, and journalists working for non-profit/unverified volunteer news sites or independent news outlets
Commercial and State Media and Journalists	Commercial or state media, including news websites, TV channels, and radio stations, or journalists, editors, etc. associated with these networks (excludes profiles who describe themselves as "citizen" or "independent" journalists or bloggers)
Private Company	For-profit company account
Public Figure	Artists, actors, influencers, sports players, etc.
State Institution or Public Official	Governmental institutions or public officials working in public institutions, such as representatives, advisors, or legislators
Misc	Profiles that fall outside of the above-listed categories
N/A	Profiles that are no longer available as they have been suspended, deleted, or removed

To understand what kind of accounts were reposting which content, I use further manual coding of topic groups. I categorize posts to provide a baseline prevalence of the different topics and to allow for further analysis of which topics are particularly likely to be reposted. However, before manually coding the posts and reposts, I first clean the repost data. After collecting the text of all reposts on Instagram, I use a Fast Text package, Language Identification, to identify the language of the repost. While all three accounts post in English, the corpus of reposts contains 22 different languages, including English, Spanish, Serbian, Indonesian, German, and Italian (for a breakdown of languages in the repost corpus, see Appendix Figure 18). Using the Google Translate API, I translate all non-English reposts (N = 736, 16.1% of the repost dataset) into English.

In part, the wide range of languages may be attributed to the international focus of these accounts, which highlight global social justice issues. As evidenced in Figure 5, which captures the locations tagged by the three Maffick Media accounts in posts, the accounts cover a broad array of international issues and current events, despite being written in English. Notably, despite the accounts being Russian-directed, Russia was never tagged as a location in the 3,532 posts.

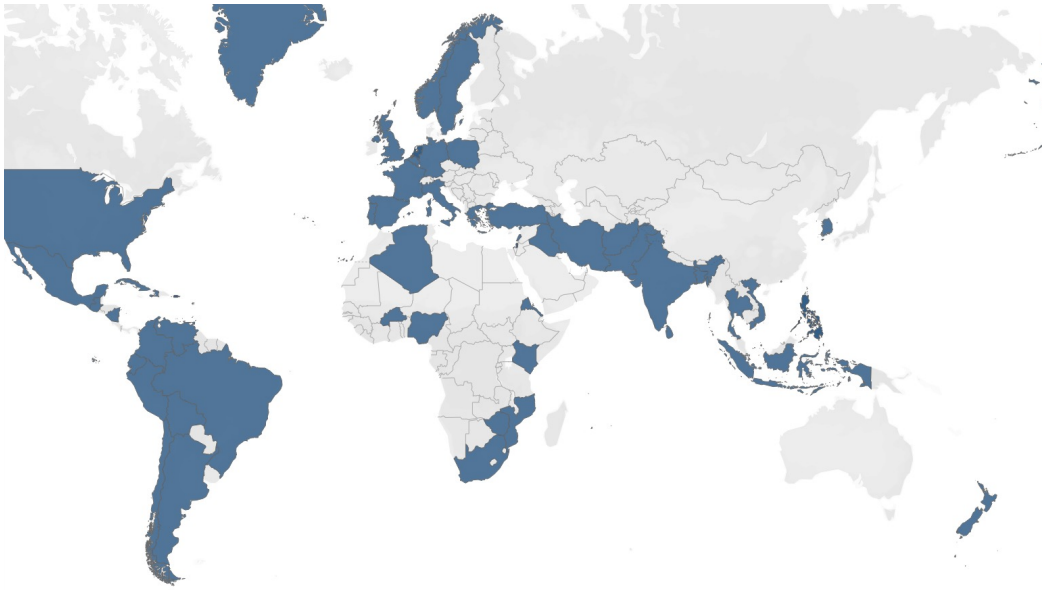


Figure 5: Locations Tagged in Russian Provocation Posts

I follow a similar protocol to the categorization of account types to categorize the topics. To develop a codebook of topics, I first collect an expansive list of possible topic groups based on the 2016 investigation into IRA activity on social media (Howard et al., 2019). Specifically, Howard et al. (2019) detail the interests targeted by IRA ads on Facebook and Instagram, such as immigration and Black identity and nationalism (Appendix Table 8). From this list of 20 possible topics, I pilot a test classification of a random sample of 100 posts and reposts. Based on the pilot analysis, I condense certain categories, such as imperialism and immigration, as well as add certain categories, such as protests. I then randomly sample 50% of the posts (N = 1,767) and 50% of the reposts (N = 2,283) to categorize by hand. While content inevitably can touch on a variety of themes, I assign each post to the dominant topical category. To provide a check on the reliability of the coding, I follow a similar protocol to the account type categorization. However, because I was responsible for the coding of 4,050 posts and reposts, I did not have a second coder available; as a result, I blind re-coded a random sample of 50 posts and reposts one month after the coding procedure had finished. The interrater reliability for the account type, captured with an unweighted Cohen's Kappa, was .60. Similarly, this value is considered "moderate" in agreement. The final codebook of post and repost types includes:

Table S2: Categories for Repost and Post Topic Theme

Topic Group	Description
Black Lives Matter (BLM)	Content touches on issues specific to Black Lives Matter, including race relations, racism, and police violence against the Black community
Capitalism/Socialism	Content explicitly discusses socialist or capitalist political thought or influential theorists
Climate	Content discusses issues relating to climate and climate change, including renewable energy, oil and gas industries, and carbon footprint
COVID-19	Content focuses around COVID-19, including the pandemic broadly, lockdown measures (e.g. mask-wearing, sanitizing, social distancing), and the vaccine
Imperialism & Immigration	Content focuses on issues of colonialism, neocolonialism, modern and historical imperialism as well as contemporary conversations around immigration and refugee resettlement
Police	Content discusses issues of police-state, police brutality, and over-policing (N.B. no mention of BLM)
Protests	Content touches on international protests around various issues and causes (N.B. no mention of BLM or women & LGBTQ+)
US Politics	Content highlights US political issues, including discussions around the US Presidential Election, President Trump, or Congress
Women & LGBTQ+	Content highlights women's or LGTBTQ+ rights, history, or discrimination
Misc.	Content has less than three words in the text and cannot be classified

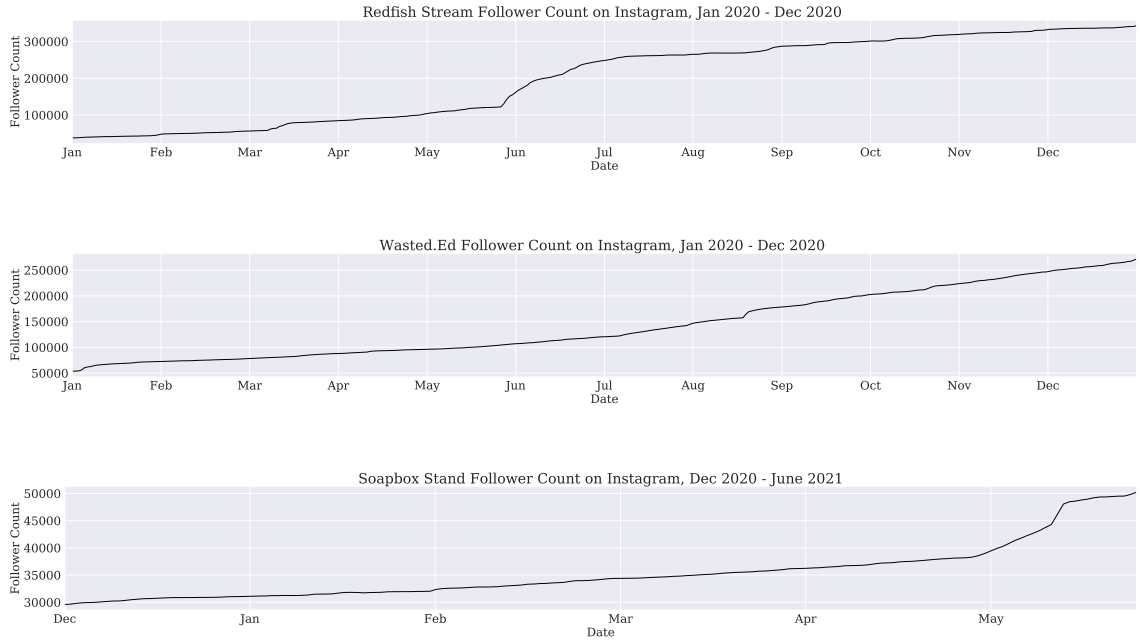
6.3 Measures

6.3.1 Follower Count

I extract the daily follower count of each of these three accounts from January 1st, 2020 to December 31st, 2020. To collect historical follower count data, I use Social Blade's API, a social media statistics and analytics website. For dates where Social

Blade's API did not have a daily follower count, I use data imputation techniques, specifically linear interpolation. As previously highlighted, due to Social Blade's data limitations, I could not collect follower data for Soapbox Stand in 2020; as a result, the dates of Soapbox Stand's follower count ranged from December 2020 – June 2021.

As can be seen in Figure 6, Redfish Stream and Soapbox Stand had dramatic peaks of follower gains, while Wasted.Ed had a more gradual gain of followers. For example, Redfish Stream had a dramatic gain of followers from June to July of 2020, during which time the account nearly doubled its follower count. I speculate that Redfish Stream and Soapbox Stand's viral follower growth may be due to the more general focus of each account around current events; for instance, during the June to July growth in followers, Redfish Stream may have been able to harness current events, specifically discussion of BLM in light of the George Floyd protests across the US, to gain momentum for its account. In contrast, Wasted.Ed may have had less viral, news-based coverage as the account focuses on general environmental issues; as a result, the account seems to have had a more gradual follower gain.



N.B.: While the X-axes for Redfish Stream and Wasted.Ed are consistent, the X-axis for Soapbox Stand is slightly different, ranging from Dec. 2020 - June 2021.

Figure 6: Maffick Media Cumulative Follower Gain on Instagram, 2020-2021

6.3.2 Number of Reposts

I compute the daily number of reposts by counting how many times each Russian account was reposted by other Instagram accounts. While this metric captures every repost or tag of these accounts in a public feed, it does not include temporary reposts or reshares of content to users' stories or private reposts. On average, Wasted.Ed had 8.47 reposts per day, Redfish Stream had 3.65 reposts per day, and Soapbox Stand had .77 reposts per day.

6.3.3 Size of Repost Reach

Similar to Zhang et al. (2021) and Wells et al. (2020), I compute a size of repost reach metric by summing together the follower counts of all accounts that repost Maffick Media content on a given day. The measure does not distinguish the account's message spreading to a wide and thin audience or reverberating through a close and dense ecosystem. However, out of users that reposted Maffick Media

content, both the median and the mode number of reposts is one, suggesting that Russian provocation content was spread through a wide and varied audience rather than a dense network of frequent reposters (for a breakdown of the count of reposts, see Appendix Figure 18). Additionally, there may be overlap in the audience between accounts, which is not accounted for. Despite these limitations, this measure represents the amount of potential exposure or impressions each Russian-run account gained daily through reposting.

6.3.4 Reposting by Influential Actors

To categorize accounts as "influential," I use Instagram's verified status, denoted by a blue checkmark badge visible in search and on an account's profile. Though online network studies have often used measures such as page range centrality or degree centrality to measure the influence of a user (Riquelme & González-Cantergiani, 2016), I use verified status to capture not just influence, which might be partly attributed to the number of followers, but also social status. Similar to Twitter, Instagram's verification program (2021) has a number of criteria, including that the account is a "well-known, highly searched for person, brand or entity." By virtue of these criteria, Instagram's verified check, much like Twitter, is often considered a status symbol, signifying social influence and personal clout (Dash, 2013). Given that the underlying variable I am trying to capture is status, particularly whether a user is seen as notable, Instagram's verified status is thus a more appropriate measure than centrality in a network for the purpose of this study.

I identify all verified Instagram users who reposted the Russian Instagram accounts in 2020 and categorize them as "influential actors." These influential actors ranged from activists, media organizations, and companies to actors, musicians, and public personalities. Notably, while 66 verified accounts reposted Maffick Media content, none of the three Maffick Media accounts themselves, despite having hundreds of thousands of followers, are verified. I compute the number of times each account

was reposted by influential actors, per day.

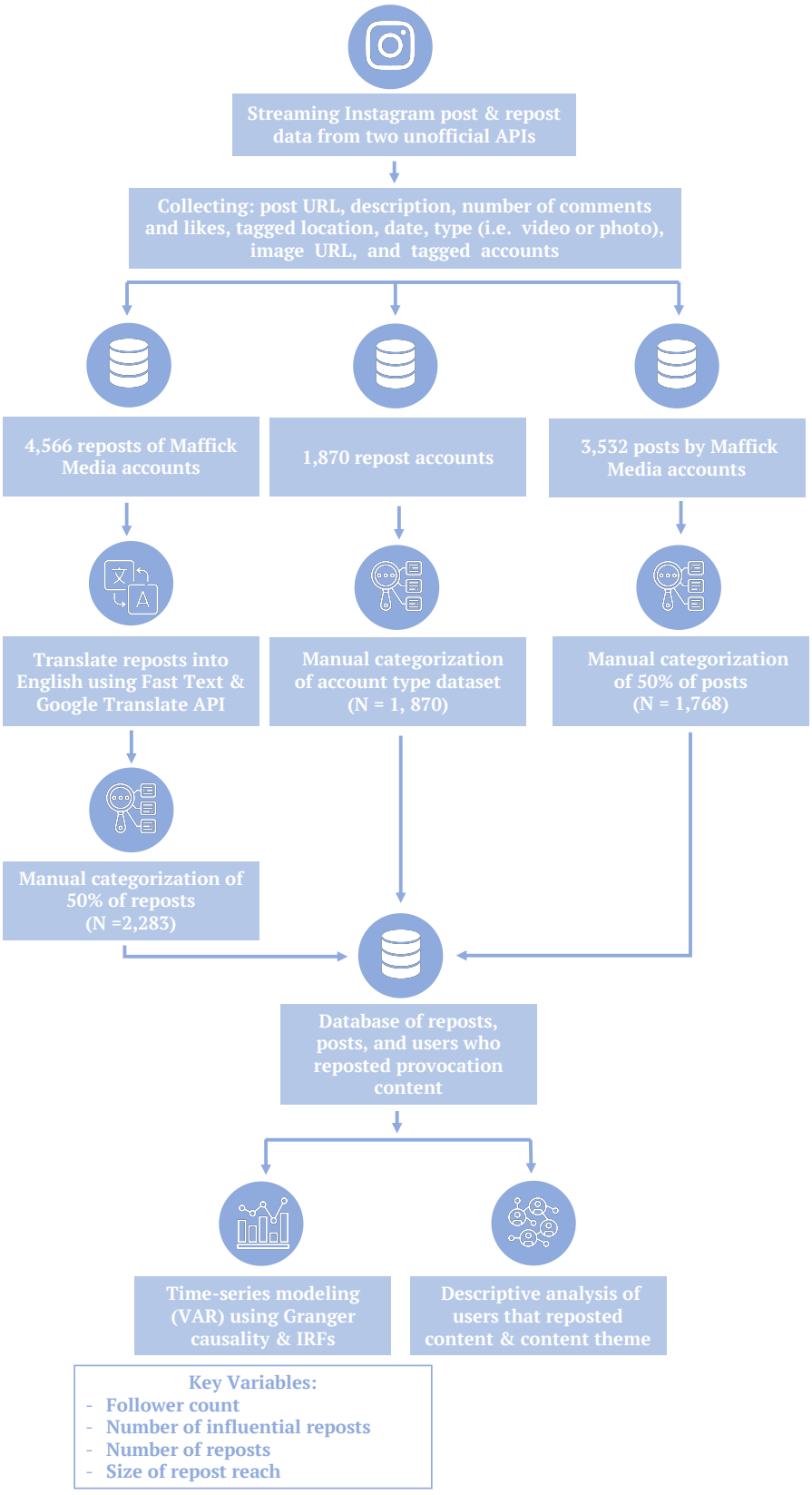


Figure 7: Schematic Overview of Data Collection and Analysis Pipeline

7 Results

7.1 Hypothesis Testing: Time Series Modelling

This section contains empirical results on the hypotheses derived in Chapter 5, as well as descriptive results for RQ2. Before presenting the results of the time series analysis, I first provide the mean and standard deviation of each measure by account (Table 3).

Table S3: Summary Statistics: Mean and Standard Deviation of Variables by Account

	Measures							
	Follower Count per Day (Δ)		Number of Reposts per Day		Size of Repost Reach per Day		Reposts by Influential Actors per Day	
Account	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Redfish Stream	837.75	1,342.23	3.64	3.60	171,065	795,773	.11	.39
Wasted.Ed	597.41	603.66	8.44	4.02	357,776	641,816	0.06	.26
Soapbox Stand	117.24	226.97	0.76	1.10	22,709	152,105	0.01	.11

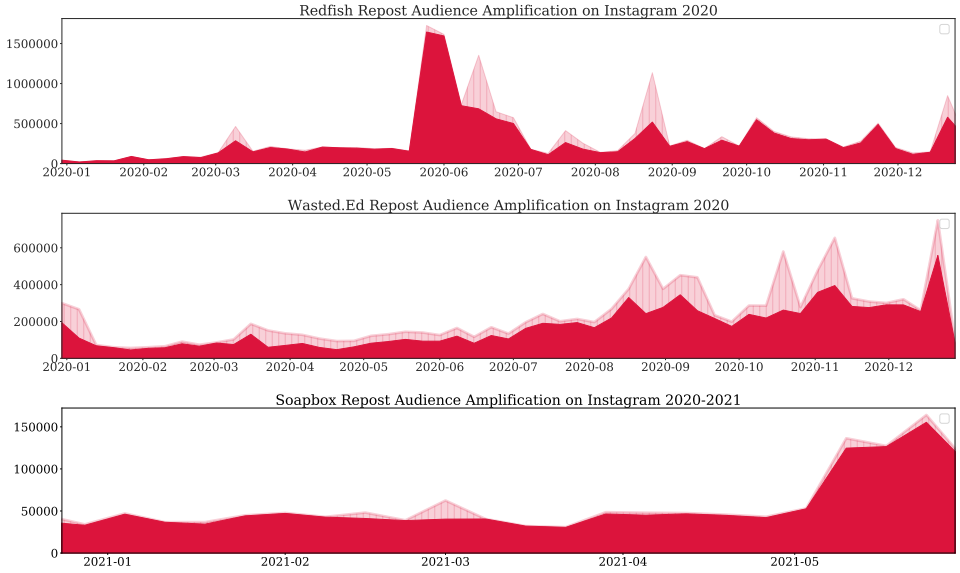
As seen in Table 3, Wasted.Ed had the highest mean number of reposts per day, while Redfish Stream had the highest number of reposts by influential actors. Despite this contrast, Wasted.Ed and Redfish Stream had a similar mean size of repost reach, suggesting that Redfish Stream benefited from low quantity but high quality (e.g. "influential") reposts while Wasted.Ed benefited from high quantity but low quality reposts.

Table S4: Breakdown of Reposts, Engagement, and Amplification

Account	Avg. Engagement of Post	Avg. Engagement of Repost	Amplification Factor = Engagement of Reposts/ (Engagement of Reposts + Posts)
Redfish Stream	42,592	7,315	17.00%
Soapbox Stand	7,937	483	6.08%
Wasted.Ed	2,905	1,050	36.12%

I also examine the engagement of reposted content, as measured by the number of likes and comments on a repost. On average, reposted content resulted in 2,949 engagements. Across all three channels, Wasted.Ed had the highest number of reposts per day at an average daily rate of 8.44, totaling over 3,000 reposts in 2020. However, Redfish Stream had the highest engagement on each of its posts as well as the highest engagement on reposted content. This finding further supports the high quality versus high quantity explanation for the difference between reposting behaviour on Redfish Stream and Wasted.Ed.

Comparing the engagements of posts by the Russian provocation accounts to reposts by other actors, I observe an average amplification factor, as measured by the extra engagement reached through reposts divided by the total engagement (i.e. sum of engagement from posts and from reposts), of 19.73%. The amplification factor is captured in Figure 8; the lighter transparent red represents the additional engagement from reposts from followers, which provided a broader reach for Russian provocation accounts' content. While certain accounts, like Soapbox Stand, had a more minimal audience amplification factor through reposts, accounts like Wasted.Ed saw their audience reached increase by 36.12%. This finding corroborates previous work on Twitter reposts by Snegovaya and Watanabe (2021), which concluded that the reach of Russian Twitter accounts was highly supported by followers who propagate their narratives.



N.B.: The engagement of Maffick Media accounts is shown in dark red, and the added engagement from reposts is shown in a lighter transparent red. While the X-axes for Redfish Stream and Wasted.Ed are consistent, the X-axis for Soapbox Stand is slightly different, ranging from Dec. 2020 - June 2021.

Figure 8: Audience Reached through Repost Amplification for Redfish Stream, Wasted.Ed, and Soapbox Stand on Instagram

To test the two potential models, I use Wald two-way Granger causality tests. While H1 hypothesizes that influential users reposting Russian provocation content will lead to follower growth (i.e. top-down, broadcast model), H2 and H3 hypothesize that large numbers of reposts or a significant repost reach would lead to follower growth (i.e. bottom-up, social contagion model).

Table S5: Granger Causality Tests for Maffick Media Accounts

		Redfish Stream		Soapbox Stand		Wasted.Ed		
	Relationship	F	p-value	F	p-valuue	F	p-value	Conclusion
H1	Reposts by Influential Users ->Followers	19.899	0.000***	27.578	0.000***	5.107	0.024*	Granger causality
	Followers ->Reposts by Influential Users	0.274	0.601	0.484	0.487	0.095	0.758	No causality
H2	Number of Reposts ->Followers	50.495	0.000***	33.598	0.000***	1.555	0.213	Mixed results
	Followers ->Number of Reposts	2.415	0.121	2.860	0.093.	0.073	0.788	No causality
H3	Size of Repost Reach ->Followers	2.326	0.128	9.996	0.002**	4.377	0.037*	Mixed results
	Followers ->Size of Repost Reach	0.264	0.608	0.626	0.484	0.142	0.707	No causality

*p <.05, **p <.01, ***p <.001

As standard with Wald two-way Granger causality tests, I include both possible directions of Granger causality; each hypothesis (A ->B) is displayed above in white, while its inverse (B ->A) is listed below in grey. The conclusion of each hypothesis – Granger causality, no causality, and mixed results – is included under conclusion.

All three accounts' Granger causality tests demonstrate that influential actors

reposting content led to statistically significant growth in followers, thus confirming H1. For each of the accounts, a one standard deviation increase in reposts by influential users would increase followers at time t+2 by about 100 for Redfish Stream (SD = .4), 60 for Wasted.Ed (SD = .3), and 50 for Soapbox Stand (SD = .1). In other words, one repost by an influential actor for Redfish Stream would result in 256 followers gained (roughly 25% of the account's daily average follower growth), 231 for Wasted.Ed (38% of daily follower growth), and 454.5 for Soapbox Stand (388% of daily follower growth). All three Granger causality tests are significant at the .001 level for Redfish Stream and Soapbox Stand and the .01 level for Wasted.Ed. IRFs, which further control for other variables in the model, partially corroborate the effect of influential users reposting content on follower growth (Figure 9, Figure 10, Figure 11). Across all three accounts, the IRFs display similar duration and magnitude; however, the 95% confidence interval band of the shock did cross the 0 y-axes for all three accounts. Thus, I find evidence of Granger causality between influential actors reposting content and follower growth for Russian provocation accounts, which I partially corroborate with further IRFs.

Examining the two bottom-up, social contagion model hypotheses (H2 & H3), Wald two-way Granger causality tests show varied results across all three accounts. For H2 – that the number of reposts would lead to follower gain – the results from two of the three accounts (Redfish Stream and Soapbox Stand) are significant at the .001 level. IRFs for Soapbox Stand and Redfish Stream corroborate the effect of reposts. For Soapbox Stand, a one standard deviation increase in reposts (SD = 1) would increase its followers by about 70 at time t+1 and about 160 at time t+2. Similarly, a one standard deviation increase in reposts of Redfish Stream (SD = 3) would increase their followers by around 250 at time t+1. The magnitude of this change in followers is comparable to the effect size found for reposts of @TenGOP, one of the most successful accounts on Twitter, by Zhang et al. (2021). Moreover, both IRFs for Soapbox Stand and Redfish Stream on the effect of the number of

reposts on follower gain are consistently above 0. However, while the magnitude of this relationship is greater than the relationship between reposts by influential users and followers gained, I find no statistically significant Granger causality relationship between the number of reposts and follower gain for Wasted.Ed.

For H3 – that the size of the repost reach would lead to follower gain – I find similarly mixed evidence: two of the three accounts (Soapbox Stand and Wasted.Ed) are significant at the .01 and .05 level, respectively. However, the IRFs cannot corroborate this finding for Soapbox Stand, suggesting a weak relationship. Thus, I find insufficient evidence to demonstrate H3 – the greater the size of the reposting reach of accounts, the more followers Russian left-leaning accounts will gain.

In sum, while the resulting follower gain is larger for H2 than for H1, the results of the Granger causality test are more consistent for H1 than H2. In other words, I find stronger evidence in terms of size of follower growth that the number of reposts drives follower gain, but I find stronger evidence in terms of consistency across all case studies that the number of influential reposts drives follower gain.

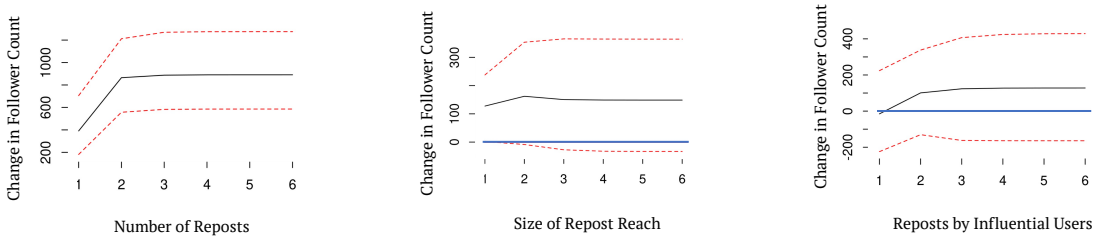


Figure 9: IRFs for Redfish Stream

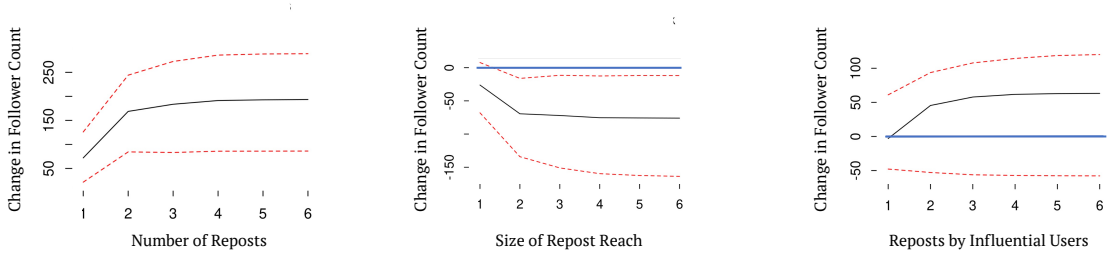


Figure 10: IRFs for Soapbox Stand

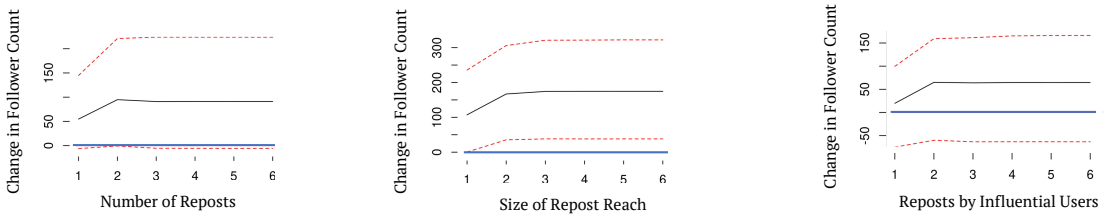


Figure 11: IRFs for Wasted-Ed

The 95% confidence interval is represented in the dashed red; effect shown in black and 0 y-axis shown in blue (where applicable).

7.2 Account Types & Reposted Content

Having examined the relationship between reposts and follower gain, I discuss the results of RQ2, namely analyzing the kind of account that reposted Russian provocation content as well as the theme of the reposted content. As seen in Figure 12, the most frequent users that reposted Russian provocation content included: 1) Activist and Civil Society Groups (1,469), 2) Citizen Journalist/Alternative or Independent News Outlets (1,088), and 3) Private Companies (651). In contrast, state institutions and public officials as well as commercial and state media rarely reposted Russian provocation content. While activist and civil society groups were the most frequent reposting accounts of Russian provocation content, private companies garnered the highest engagement on their reposts. Private companies were also responsible for six of the ten highest engaging reposts of Russian provocation content. On average, reposted Russian content by private companies resulted in 3,334 engagements compared to 1,745 engagements on reposts by activists and civil society groups (Figure 12). Private company accounts ranged from small boutique clothing stores to million-dollar conglomerates in sports, clothing, and other consumer goods. My results thus bring to light the key role of private companies and activist and civil society groups in generating engagement for Russian provocation accounts.

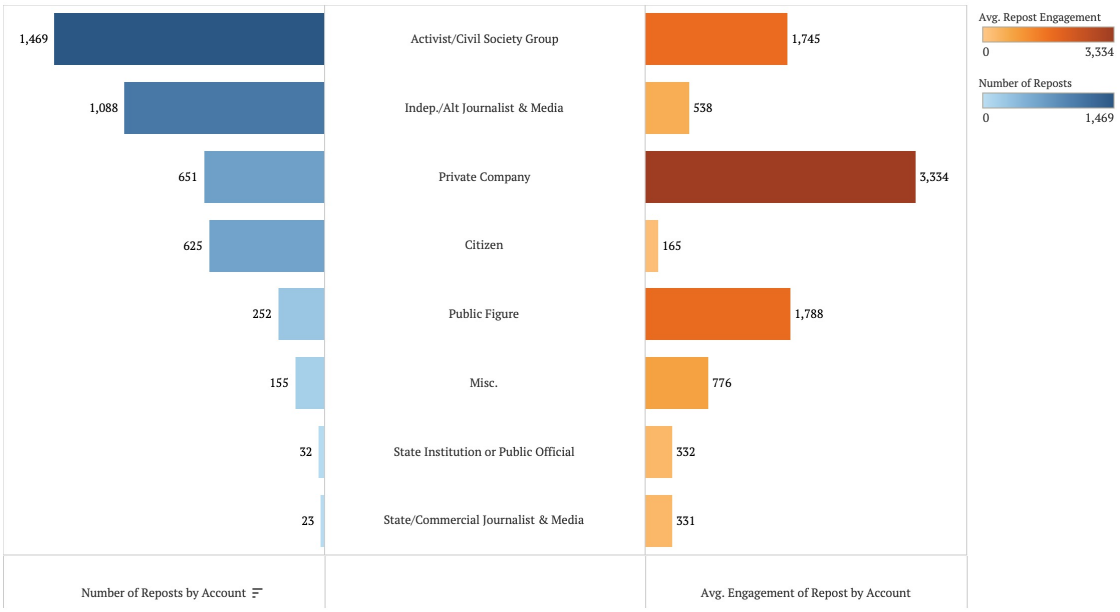


Figure 12: Number of Reposts of Maffick Media by Account Type and Average Engagement, 2020-2021

Examining topic themes, the most common themes in reposted content included: 1) Climate change (1353), 2) COVID (246), and 3) BLM (200) (Table 6). However, comparing the breakdown of content themes in reposts to a baseline of content themes in posts, I find that climate change was both the most frequently reposted topic as well as the most frequently posted topic by Russian provocation accounts. The high number of climate-themed posts and reposts may be attributed to the fact that one of the three accounts has an overt climate focus (Wasted.Ed); moreover, Wasted.Ed had the highest number of daily reposts, suggesting a very active dissemination network. As seen in Figure 13, content around BLM and COVID-19 was frequently shared by Russian provocation accounts and their networks of reposting accounts.

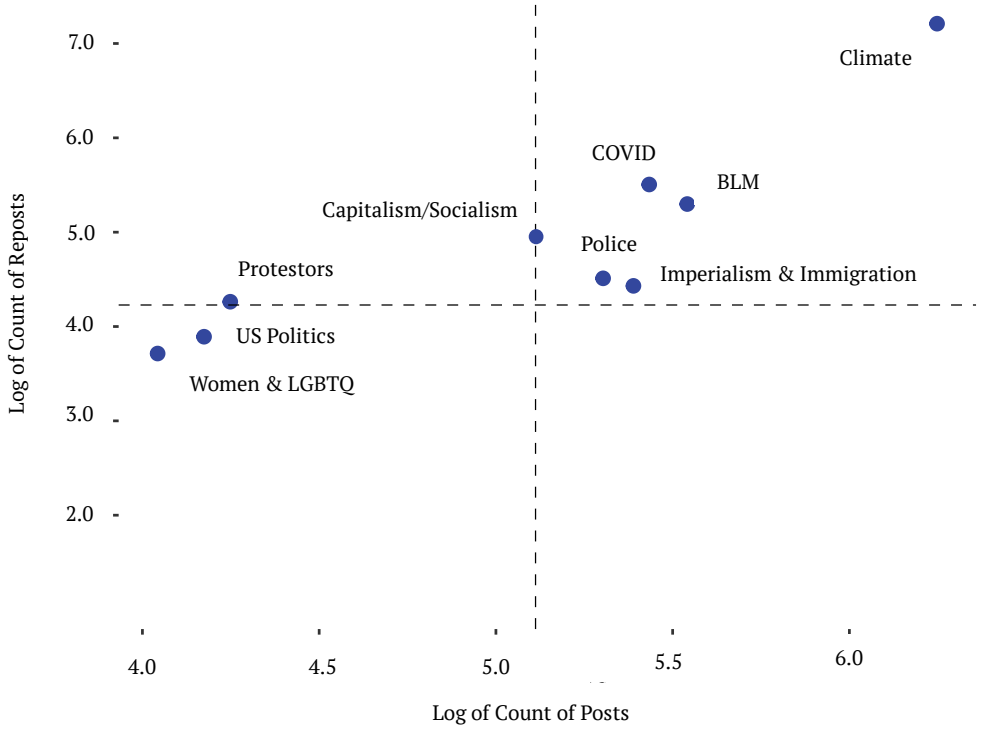


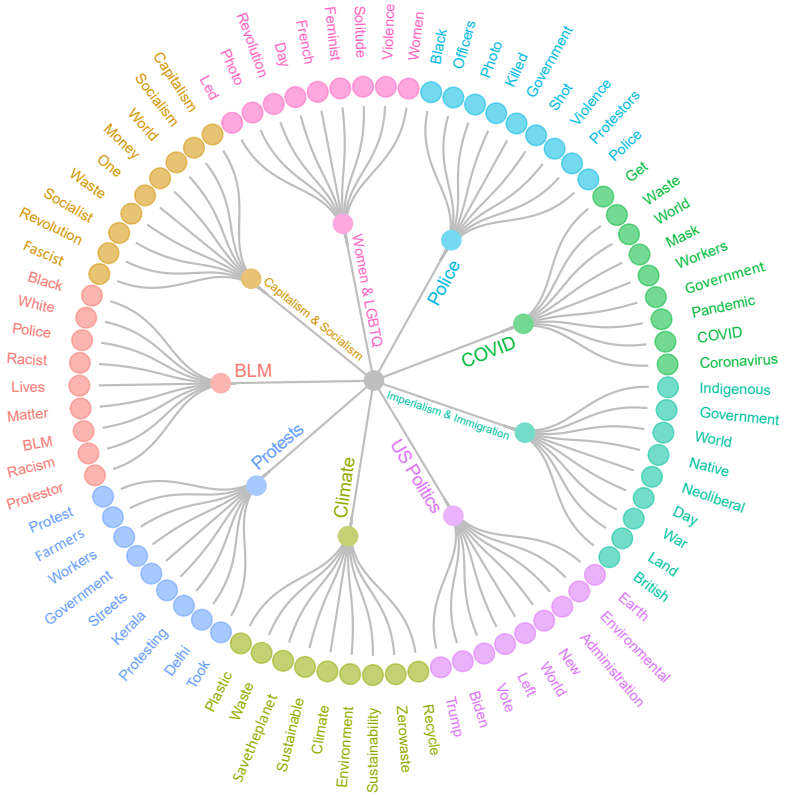
Figure 13: Log of Count of Posts vs. Reposts of Russian Provocation Content by Topic

As seen in the sample posts included in Table 6 as well as the top words per topic captured in Figure 14, discussion around BLM frequently highlighted protests around the US in the wake of George Floyd's murder, drew comparisons of modern-day race-relations to the Civil Rights era and even US slavery, and compared advocates for racial change, such as Martin Luther King Jr., to socialist figures, such as Stalin.

Table S6: Breakdown of Reposts by Topic, Counts, and Sample Reposted Content

Topic	Count	Sample Post
Climate	1353	"Why do we still continue to have plastic bags? We know they are incredibly damaging to the natural environment, especially in our Oceans!...Follow the link in our bio to support the cause!"
COVID	246	"Elder man wearing a #facemask seats reading a book in...a working class neighbourhood, west of #athens. This street used to be full of people...The #covid_19 #pandemic [has] changed our lives completely."
BLM	200	"..A step in the right direction is to not share racist memes and stand up strong to your friends when they send them to you. A lot of racism can be unlearned just by making the choice to educate oneself...Most of the text comes from the caption in a post by @redfishstream...good accounts to follow btw."
Capitalism /Socialism	148	"Via @redfishstream Hustle culture is just deeply internalized capitalism! Humans naturally want to work together, seeing each other as competition to beat is alienating us from connecting!"
Police	91	"Here we go again, police brutality in California, LA, in its rawest form. Why the f*** is it necessary to beat a disabled man in a wheelchair? They're monsters just like the politicians they are serving for... what a sick society we're living in, it's disgusting!"
Imperialism and Immigration	84	"U.S. Independence Day today, but for the Native Nations there isn't much to celebrate"
Protests	71	"As Indonesian students & workers are rising up, here's a short history of the last mass uprising...: the one that toppled the US-backed right-wing dictator Suharto in 1998..."
US Politics	49	"Still not convinced billionaire Mike Bloomberg is trying to buy the U.S. elections? Here is an idea of how much the presidential candidate is worth compared with the other candidates."
Women & LGBTQ+ Rights	41	"How should We Define the Difference between sex and gender?Is there a difference? #genderandsexuality #transrights #transgendergirl"

While one of the coded topic groups included explicit conversations around capitalism and socialism, Figure 14 highlights how pervasive such political commentary was; for example, one of the most commonly occurring words within imperialism and immigration was "neoliberalism." Though each topic had a unique set of most commonly occurring words, certain keywords frequently appeared throughout all topics. Across all topics, "revolution" and "protest" were frequently used, often to highlight unrest within a democratic or capitalist system or to call for a socialist revolution. Similarly, one of the most frequently used words across topics was "day," as posts often used awareness days or other holidays as springboards for topical discussions, to add a sense of timeliness. For example, "day" was commonly used within posts on imperialism and immigration to motivate discussion of native or refugee rights on national holidays, such as Columbus Day in the US.



N.B.: Inner nodes represent the topic categories, while the outer nodes represent the most frequent words in each topic group.

Figure 14: Top Words per Topic within Russian Provocation Posts and Reposts

Furthermore, I also examine the relationship between the type of repost and the type of user reposting. As seen in Figure 15, the various repost account types had similar distributions of reposted content themes, specifically clustering around BLM and climate, with the exception of state/commercial journalists and media and state institutions and public officials. Specifically, the top two repost categories for all other users were climate or BLM. In sum, despite differences in audience size or account type, reposting users tended to repost similarly-themed content.



Figure 15: Correlation Matrix of Repost User Type by Repost Content Theme

8 Discussion

8.1 Substantive and Theoretical Significance

This thesis aimed to contribute to the academic literature both substantively and methodologically by drawing together a novel dataset of Instagram reposts, social network theory, and time-series modelling to examine how three Maffick Media channels built their amplification network of followers. In light of this aim, I subsequently draw both analytical and descriptive inferences by testing two theoretically-motivated models explaining different determinants of follower gain and surveying the kind of accounts that reposted Russian provocation content as well as the content of those reposts.

To begin with, the story of how Maffick Media built three successful Instagram accounts in 2020 reveals several dynamics at work in the attention-driven political communication ecosystem. First, I confirm previous work demonstrating the significance of network effects internal to social media in driving follower gain. Though not a consistent result across all three case studies, the number of reposts was associated with the strongest effect on follower growth (H2). Building on the findings of Golochenko et al. (2018) as well as Zhang et al. (2021), the results thus demonstrate the role average users play in spreading Russian provocation content. Additionally, while previous literature on Twitter has highlighted the bottom-up network effect, this study on Instagram dynamics reveals the novel significance of influential users reposting Russian provocation content. While the effect of reposts by influential users' on follower growth was smaller than that of number of reposts, the results for all three accounts show the validity of a top-down, broadcast model. With verified users reposting their content and tagging them in posts on Instagram, these Russian provocation accounts transformed an audience into an amplifier. As outlined in our theoretical discussion, as Russian left-leaning provocation accounts turn influencers, politicians, and the like into amplifiers, followers of verified accounts that repost

their content may be more likely to follow the Russian accounts themselves due to the follow-worthiness signalled by the established account. Thus, through this top-down network, influential actors bridged the tie between Russian provocation accounts and future followers by signalling to their own followers that Maffick Media accounts were authentic and credible sources.

In addition to the analytical inferences, this thesis further contributes a set of descriptive findings. My results highlight the role of private companies, activist and civil society groups, and alternative and independent news outlets in disseminating Russian provocation content on Instagram. First and foremost, the role of private companies is a novel finding within the literature. Marketing literature (Fredrikson, 2021) has captured a rising trend of companies speaking out on social issues as a form of "values and identity-driven targeted marketing" (Hsu, 2020); by aligning corporate causes with customers' values, companies seek to build a sense of trust and foster personal connection over shared values. While companies have been critiqued for "greenwashing" and "rainbow washing," a kind of PR where companies present as caring about environmental or LGBTQ+ issues even as they are engaging in unsustainable or homophobic practices (Watson, 2016), the reposting of Russian provocation content by private companies, particularly around issues like climate change and BLM, brings to the fore further concerns around corporate "performative activism" and "virtue-signalling." Additionally, the finding around the role of activist accounts, which were often organized around left-leaning issues, such as BLM, police violence, and imperialism and immigration, supports previous evidence that IRA efforts targeted left-leaning voters as well as minority or marginalized groups on Facebook and Instagram in 2014-2017 (Howard et al., 2019). Furthermore, the high number of alternative and independent news outlets reposting Russian provocation efforts speaks to likely a wider web of Russian Fringenet accounts that easily harness a media multiplier effect to create "media storms" (Global Engagement Center, 2020).

Moreover, the topical focus around left-leaning issues, such BLM and LGBTQ+ and women's rights, as well as the polarizing discourse, which called for revolutions and protests, supports the framing of Russian information operations on Instagram as a strategy of provocation in the style of Cold War-era *aktivnye meropriyatiya*. While certain posts and reposts can be classified as misinformation, such as misleading insinuations that Democrats and Republicans in the US are conspiring to "rig" the "whole system" (Figure 2), much of the content may be factually accurate but incendiary, such as sharing graphic content of a disabled man being beaten by police (Table 6). My results thus support the framing of Russian information operations as less a strategy of misinformation, using models like the "firehose of falsehoods," and more a strategy of provocation. Failing to recognize the importance of such provocative but subjectively "true" content obscures the aim of contemporary Russian information operations. A focus solely on objectively false disinformation – and the potential implication that mere fact-checking could serve as a counter – overlooks the way Russian accounts are not merely passive purveyors of false information to frame incendiary conversations but rather active participants fanning the flames of ongoing political discourse.

Furthermore, not only were the topics identified similar to other case studies' findings but also the intertwinement of pop content with political commentary is archetypical of Russian information efforts. To the former, topics such as police violence and critiques of US politics were similar to other case studies' findings, including highlighting racial tensions in the US, accusations of foreign interference and war crimes committed by the US, and capitalism-induced suffering (Nimmo, 2020). To the latter, as seen in Figure 14, Maffick Media's accounts interwove content on news issues, such as refugee crises and climate change, with a lexicon of anti-capitalist, pro-revolutionary vocabulary. Such a careful intertwinement of "objective" clickbait news reporting with underlying political commentary has been noted in the literature as archetypical of other accounts like Ruptly (Snegovaya

& Watanabe, 2021; Hanlon & Morley, 2019). Therefore, my findings support the framing of Russian information operations on Instagram as a strategy of provocation and corroborate previous findings around the intertwinement of news content with political commentary.

8.2 Methodological Discussion

While this thesis sheds light on the dynamics of Russian provocation accounts on Instagram, it also highlights the challenges inherent in the data set itself. Namely, though leveraging Instagram data allows for unique insights which can complement current social media research on Twitter and Facebook, the dataset was also rife with limitations. To begin with, due to the limitations of the Instagram API, this thesis leverages two unofficial Instagram APIs, which may be less accurate than collecting the raw data. By using unofficial Instagram APIs, I introduce a potential processing error, which may bias the results of this thesis. Similarly, because the API only includes data on certain accounts and for certain dates, I could not examine all of Maffick Media's accounts, nor could I collect data from the same date range (i.e. January 1, 2020 – December 31, 2020) for one of the accounts in this thesis (Soapbox Stand).

Due to the nature of the dataset and the platform, the risks of a nonresponse or missing data error arises, which degrades the data and limits the replicability of the study. As noted previously, the data collection was limited to public tags on an account's page, thereby excluding private reposts or temporary reposts of content to a user's Instagram story. As a result, the number of reposts metric may be lower than the true value due to unaccounted data points. In addition to concerns around missing data, using historical Instagram data limits the future replicability of the study. Even between the weeks of the data collection process, 53 repost accounts, roughly 2.83% of the accounts, were deleted or suspended. As a consequence, certain accounts had to be removed from the sample as data could no longer be collected

on the account, which gives rise to a missing data error issue. Finally, the use of historical Instagram data also limits future replicability of this thesis, as accounts and posts may be removed or deleted over time.

Similarly, the selection of just three accounts raises concerns of sampling error; while these three accounts are among the most successful Russian directed provocation accounts on Instagram, hundreds of other Russian accounts remain unstudied. Moreover, as evidenced by the difference in the number of followers and, by extension, the number of reposts, Soapbox Stand has a much smaller following and repost reach; indeed, the dataset for this account was significantly smaller than Wasted.Ed or Redfish Stream. Consequently, not only may there be other dynamics that I am not able to capture using this sample of just three accounts but also the three accounts themselves may have had differing dynamics at play.

Furthermore, the dataset broke with many social science conventions, such as random selection and representativeness. Due to the covert tactics of the Fringenet described previously, I cannot assess whether this sample is representative of the population as a whole, as the underlying population is unknown. Moreover, given that the sample was limited to left-leaning accounts, this thesis' results should not be generalized to Russian provocation accounts across the political spectrum. While this thesis has high internal validity, it is therefore challenging to generalize the findings to broader dynamics of the Russian Fringenet on Instagram. However, given that these accounts were among the most successful in terms of the number of followers, they provide a rich dataset that aptly fits the purpose and scope of this study.

Finally, as highlighted previously, I use a system of manual coding as opposed to automated analysis to classify the repost account type and topic. While manual coding allowed for a more accurate and nuanced analysis of content, particularly accounting for both image-based content and text, my process was time-intensive and costly. Though other studies of comparable size and scope within disinforma-

tion studies frequently use manual coding over automated analysis (Golovchenko et al., 2018), manual coding limits the scalability of the methodology and introduces elements of researcher subjectivity.

8.3 Practical Implications

In addition to the substantive and methodological contributions to the academic community, the findings of this study also bear practical implications for policymakers and social media companies like Instagram. Primarily, this work brings to light the criticality of using a multiplayer analysis to study Russia's strategy of polarization and provocation. These findings not only highlight the importance – beyond the novel or symbolic – of a famous actor or well-known company resharing Russian provocation content to their followers for building these accounts' followings but also suggest that simply marking an account as state-media may not be enough. Rather, countless users were exposed to Russian provocation content from users within their network with no mention of the original source's bias. In sum, this work draws attention to the role of *aktivnye meropriyatiya*-style tactics, in which provocative but accurate information is spread by state actors through inauthentic media accounts. Our findings demonstrate to policymakers that mere fact-checking or flagging false information may not be sufficient to counter Russian provocation efforts that leverage misleading accounts and networks of unwitting followers to amplify incendiary discourse and exacerbate tensions.

9 Conclusion

In sum, this thesis has advanced both theoretical and methodological contributions. In terms of theoretical contributions, this thesis answers two research questions. Its first probe, RQ1, was analytical, posing the question of whether reposts by "fellow travellers" generate follower growth for left-leaning Russian provocation accounts on Instagram. Drawing on two dominant network dynamic models, namely a bottom-up social contagion model and a top-down broadcast model, I find evidence that reposts by influential users drove follower growth across all three channels (H1). This finding supports a top-down broadcast model, where influential users connected Maffick Media accounts with future followers by signalling that these accounts were authentic and credible sources. Further, my results also demonstrate a mixed but strong effect for one of the two bottom-up social contagion model hypotheses, namely that the number of reposts drove follower gain (H2). The bottom-up social contagion model finding corroborates evidence from Twitter studies that have demonstrated the role of average users in driving Russian provocation content (Golovchenko et al., 2018; Zhang et al., 2021). A second research question aimed to answer what kind of content, and which accounts reposted Russian left-leaning provocation efforts on Instagram in 2020. Throughout 2020, Maffick Media pushed content on issues such as climate change and BLM, which was readily reposted by a network of "fellow travellers." These findings brought to light a web of likely witting followers of alternative and independent journalists and news sites that amplify Russian provocation accounts as well as the likely unwitting civil society and activist organizations and private companies that reshared provocation content.

Beyond its substantive work, this thesis also advances methodological work in the field. Building on approaches by Wells et al. (2020) and Zhang et al. (2021), I use time-series analyses to test the effect of reposts on follower growth. Additionally, the novel dataset of Instagram post and repost data complements previous research that has investigated Twitter and Facebook; this data collection process also provides an

example methodology for how to collect data from Instagram, despite its official API limitations, without a paywall.

While I have addressed a few methodological limitations in 8.2 Methodological Discussion, I turn to a series of epistemological limitations and discuss potential opportunities for future research. First, a variety of methodological challenges arise when studying the Fringenet. Due to the covert tactics of the Russian Fringenet and accounts like Maffick Media, I cannot separate organic mechanisms, such as reposts by authentic accounts, from inorganic, coordinated mechanisms. Consequently, several metrics, such as the number of reposts or the repost reach, may be skewed by inauthentic activity, which does not reflect organic reposts on Instagram. While a variety of efforts have attempted to identify Russian bot accounts on Twitter as well as develop machine learning based bot classification methods, future research could continue to improve and to extend machine learning based bot classification to Instagram as well as develop a list of suspected Russian bot accounts on the platform.

Furthermore, as previously highlighted, all data included in this research is limited to public reposts on users' feeds, as opposed to private tags or story tags. While the discussion section has spoken to this limitation as a missing data error, I also note the possibility of moving from historical or static Instagram studies to live case studies. Future research might consider live case studies around Russian provocation accounts; for example, collecting story data from these accounts to analyze how they respond in live time to current events as well as how followers reshare their content to stories on Instagram.

Finally, although derived from social network theory, the VAR models are relatively simplistic, both due to the scale and scope of this thesis as well as to make findings interpretable. As such, future extensions could develop more complex measures of certain variables, such as controlling for overlap in network reach or comparing different measures, such as betweenness centrality, for influential users. Further-

more, future extensions might also investigate additional variables through manual or automated coding, such as the accuracy of Russian provocation posts, engagement (e.g. likes, comments) on content, and tone of posts and reposts.

In the time of COVID-19, social media has become even more of a playing field for discussions around social justice issues and advocacy. Within this sea of content, Russian provocation content continues to seamlessly blend in with a network of followers that repost and reshare their content. Though such content does not necessarily squarely fall within definitions of misinformation or disinformation at times, the aims of the account as well as the lack of transparency and deliberate intent to hide such aims are of critical importance. Given the covert tactics of the Fringenet, however, it can be challenging to study how these accounts build their following, appeal to their audiences, and even use their audiences as further amplifiers of their content. This thesis has sought to make a small step in the direction of answering these questions by developing and testing a theoretical framework of follower growth and quantifying the impact of reposts on the follower count of Russian provocation accounts. By shedding light on these dynamics, this thesis thus brings to light the importance of a multiplayer analysis in understanding Russian provocation efforts. While future work will continue to add detail to how these accounts position themselves within a broader left-leaning ecosystem on social media, this thesis reminds us all to think before we repost.

References

1055782. (2021). *Examining the "Ins and Outs" of the Russian Fringenet: A Social Network Analysis of Russian Alternative News Sites on Twitter.*

Abrams, S. (2016). Beyond propaganda: Soviet active measures in putin's russia. *Connections*, *15*(1), 5–31. Retrieved from https://www.jstor.org/stable/26326426 (Publisher: Partnership for Peace Consortium of Defense Academies and Security Studies Institutes)

Aral, S., Muchnik, L., & Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*, *106*(51), 21544–21549. Retrieved from http://www.pnas.org/cgi/doi/10.1073/pnas.0908800106 doi: 10.1073/pnas.0908800106

Arceneaux, K., & Johnson, M. (2015). How does media choice affect hostile media perceptions? evidence from participant preference experiments. *Journal of Experimental Political Science*, *2*(1), 12–25. Retrieved from https://www.cambridge.org/core/journals/journal-of-experimental-political-science/article/how-does-media-choice-affect-hostile-media-perceptions-evidence-from-participant-preference-experiments/C364D8A6FA2B3B9C472D75F81FCCD8F1 doi: 10.1017/xps.2014.10

Arif, A., Stewart, L. G., & Starbird, K. (2018). Acting the part: Examining information operations within #BlackLivesMatter discourse. *Proceedings of the ACM on Human-Computer Interaction*, *2*, 1–27. Retrieved from https://dl.acm.org/doi/10.1145/3274289 doi: 10.1145/3274289

Badawy, A., Addawood, A., Lerman, K., & Ferrara, E. (2019). Characterizing the 2016 russian IRA influence campaign. *Social Network Analysis and Mining*, *9*(1), 31. Retrieved from https://doi.org/10.1007/s13278-019-0578-6 doi: 10.1007/s13278-019-0578-6

Badawy, A., Ferrara, E., & Lerman, K. (2018). Analyzing the digital traces of political manipulation: The 2016 russian interference twitter campaign. In *2018 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)* (pp. 258–265). (ISSN: 2473-991X) doi: 10.1109/ASONAM.2018.8508646

Bail, C. A., Guay, B., Maloney, E., Combs, A., Hillygus, D. S., Merhout, F., ... Volfovsky, A. (2020). Assessing the russian internet research agency's impact on the political attitudes and behaviors of american twitter users in late 2017. *Proceedings of the National Academy of Sciences*, *117*(1), 243–250. Retrieved from http://www.pnas.org/lookup/doi/10.1073/pnas.1906420116 doi: 10.1073/pnas.1906420116

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in american politics.* Oxford University Press. (Google-Books-ID: MVRuDwAAQBAJ)

Bennett, W. L., & Segerberg, A. (2012). The logic of connective action. *Information, Communication & Society*, *15*(5), 739–768. Retrieved from https://doi.org/10.1080/1369118X.2012.670661 (Publisher: Routledge _eprint:

https://doi.org/10.1080/1369118X.2012.670661) doi: 10.1080/1369118X.2012.670661

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy*, *102*(5), 841–877. Retrieved from https://www.journals.uchicago.edu/doi/abs/10.1086/261957 (Publisher: The University of Chicago Press) doi: 10.1086/261957

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, *100*(5), 992–1026. Retrieved from https://www.journals.uchicago.edu/doi/abs/10.1086/261849 (Publisher: The University of Chicago Press) doi: 10.1086/261849

Bonifacic, I. (2020). Instagram starts labeling 'state-controlled media' accounts and posts. *Engadget*. Retrieved from https://www.engadget.com/instagram-state-controlled-media-application-220248779.html

Box-Steffensmeier, J. M., Freeman, J. R., Hitt, M. P., & Pevehouse, J. C. W. (2014). *Time series analysis for the social sciences*. Cambridge University Press. (Google-Books-ID: WLaLBgAAQBAJ)

Boyd, D. (2017). Google and facebook can't just make fake news disappear. *Wired*. Retrieved from https://www.wired.com/2017/03/google-and-facebook-cant-just-make-fake-news-disappear/

Cardenal, J. P., Kucharczyk, J., Mesežnikov, G., & Pleschová, G. (2017). *Sharp power: Rising authoritarian influence.* Retrieved from https://www.ned.org/wp-content/uploads/2017/12/Sharp-Power-Rising-Authoritarian-Influence-Full-Report.pdf

Chiovaro, M., Windsor, L. C., & Paxton, A. (2021). Vector autoregression, cross-correlation, and cross-recurrence quantification analysis: A case study in social cohesion and collective action. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*(43). Retrieved from https://escholarship.org/uc/item/49d0w3rn

Clark, B. (2016). *Political economy: A comparative approach, 3rd edition: A comparative approach*. ABC-CLIO. (Google-Books-ID: oCyyCwAAQBAJ)

Collier, K., & Dilanian, K. (2020). *Russian internet trolls hired u.s. journalists to push their news website, facebook says.* Retrieved from https://www.nbcnews.com/tech/tech-news/russian-internet-trolls-hired-u-s-journalists-push-their-news-n1239000

Dash, A. (2013). *What it's like being verified on twitter.* Retrieved from https://anildash.com/2013/03/01/what_its_like_being_verified_on_twitter/

Davis, C. (2018). 'grassroots' media startup redfish is supported by the kremlin. *The Daily Beast*. Retrieved from https://www.thedailybeast.com/grassroots-media-startup-redfish-is-supported-by-the-kremlin

Dawson, A., & Innes, M. (2019). How russia's internet research agency built its disinformation campaign. *The Political Quarterly*, *90*(2), 245–256. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-923X.12690 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-923X.12690) doi: 10.1111/1467-923X.12690

De Cock Buning, M. (2018). *A multi-dimensional approach to disinformation :*

report of the independent high level group on fake news and online disinformation. Publications Office of the European Union. Retrieved from https://cadmus.eui.eu//handle/1814/70297 (Accepted: 2021-03-02T15:34:50Z) doi: 10.2759/739290

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., ... Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*(3), 554–559. Retrieved from http://www.pnas.org/lookup/doi/10.1073/pnas.1517441113 doi: 10.1073/pnas.1517441113

DFR Lab. (2018). *Question that: Rt's military mission.* Retrieved from https://medium.com/dfrlab/question-that-rts-military-mission-4c4bd9f72c88

Dilanian, K., & Ramgopal, K. (2020). Facebook blocks russia-backed accounts, but other sites keep churning out content aimed at americans. *NBC News*. Retrieved from https://www.nbcnews.com/politics/2020-election/facebook-blocks-russia-backed-accounts-other-sites-keep-churning-out-n1242683

DiResta, R., Shaffer, K., Ruppel, B., Sullivan, D., & Matney, R. (2019). The tactics & tropes of the internet research agency. *Congress of the United States at DigitalCommons@University of Nebraska*, 103.

Faris, R., Roberts, H., Etling, B., Bourassa, N., Zuckerman, E., & Benkler, Y. (2017). *Partisanship, propaganda, and disinformation: Online media and the 2016 u.s. presidential election* (SSRN Scholarly Paper No. ID 3019414). Retrieved from https://papers.ssrn.com/abstract=3019414

Fluegel, S., & King, K. (2021). #workfromhome: how multi-level marketers enact and subvert federal language policy for profit. *Language Policy*. Retrieved from https://doi.org/10.1007/s10993-021-09589-x doi: 10.1007/s10993-021-09589-x

Fredrikson, K. (2021). Woke-washing or altruism? : Understanding the phenomenon of brand activism by studying the media effects and audience responses to the practice. Retrieved from https://www.doria.fi/handle/10024/181575 (Accepted: 2021-08-02T11:56:03Z)

Fu, J. S. (2016). Leveraging social network analysis for research on journalism in the information age. *Journal of Communication*, *66*(2), 299–313. Retrieved from https://doi.org/10.1111/jcom.12212 doi: 10.1111/jcom.12212

Gainous, J., & Wagner, K. M. (2013). *Tweeting to power: The social media revolution in american politics*. Oxford University Press. (Google-Books-ID: cc48BAAAQBAJ)

Gillespie, T. (2011). *Can an algorithm be wrong? twitter trends, the specter of censorship, and our faith in the algorithms around us – culture digitally.* Retrieved from https://culturedigitally.org/2011/10/can-an-algorithm-be-wrong/

Global Engagement Center. (2020). *Pillars of russia's disinformation and propaganda ecosystem* [GEC Special Report].

Goel, S., Anderson, A., Hofman, J., & Watts, D. J. (2016). The structural virality of online diffusion. *Management Science*, *62*(1), 180–196.

Retrieved from https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2015.2158 (Publisher: INFORMS) doi: 10.1287/mnsc.2015.2158

Golovchenko, Y., Hartmann, M., & Adler-Nissen, R. (2018). State, media and civil society in the information warfare over ukraine: citizen curators of digital disinformation. *International Affairs*, *94*(5), 975–994. Retrieved from https://doi.org/10.1093/ia/iiy148 doi: 10.1093/ia/iiy148

Graham, T., Broersma, M., Hazelhoff, K., & van 't Haar, G. (2013). Between broadcasting political messages and interacting with voters. *Information, Communication & Society*, *16*(5), 692–716. Retrieved from https://doi.org/10.1080/1369118X.2013.785581 (Publisher: Routledge _eprint: https://doi.org/10.1080/1369118X.2013.785581) doi: 10.1080/1369118X.2013.785581

Grimmelmann, J. (2015). The law and ethics of experiments on social media users. *Colorado Technology Law Journal*, *13*, 219. Retrieved from https://heinonline.org/HOL/Page?handle=hein.journals/jtelhtel13&id=243&div=&collection=

Groshek, J. (2011). Media, instability, and democracy: Examining the granger-causal relationships of 122 countries from 1946 to 2003. *Journal of Communication*, *61*(6), 1161–1182. Retrieved from https://doi.org/10.1111/j.1460-2466.2011.01594.x doi: 10.1111/j.1460-2466.2011.01594.x

Hanlon, B., & Morley, T. (2019). *Russia's network of millennial media.* Retrieved from https://securingdemocracy.gmfus.org/russias-network-of-millennial-media/ (Section: Interference Matters)

Hindman, M., & Barash, V. (2018). Disinformation, 'fake news', and influence campaigns on twitter. *Knight Foundation*, 62.

Hjorth, F., & Adler-Nissen, R. (2019). Ideological asymmetry in the reach of pro-russian digital disinformation to united states audiences. *Journal of Communication*, *69*(2), 168–192. Retrieved from https://doi.org/10.1093/joc/jqz006 doi: 10.1093/joc/jqz006

Howard, P. N., Ganesh, B., Liotsiou, D., Kelly, J., & François, C. (2019). The IRA, social media and political polarization in the united states, 2012-2018. , 48.

Hsu, T. (2020). Corporate voices get behind 'black lives matter' cause. *The New York Times*. Retrieved from https://www.nytimes.com/2020/05/31/business/media/companies-marketing-black-lives-matter-george-floyd.html

Hutto, C., Yardi, S., & Gilbert, E. (2013). A longitudinal study of follow predictors on twitter. *CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 821–830. Retrieved from https://doi.org/10.1145/2470654.2470771

Instagram. (2021). *Verified badges.* Retrieved from https://help.instagram.com/854227311295302

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, *60*(11), 2169–2188. Retrieved from https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/asi.21149 (_eprint: https://asistdl.onlinelibrary.wiley.com/doi/pdf/10.1002/asi.21149) doi: 10.1002/asi.21149

Kaczmarska, K., & Keating, V. C. (2018). Russia's influence is much more than propaganda and fake news. *Euro News*. Retrieved from https://www.euronews.com/2018/04/04/russia-s-influence-is-much-more-than-propaganda-and-fake-news-view

Katz, E., & Lazarsfeld, P. F. (2017). *Personal influence: The part played by people in the flow of mass communications*. Routledge. doi: 10.4324/9781315126234

Liebowitz, S. J., & Margolis, S. E. (1994). Network externality: An uncommon tragedy. *Journal of Economic Perspectives*, *8*(2), 133–150. Retrieved from https://www.aeaweb.org/articles?id=10.1257/jep.8.2.133 doi: 10.1257/jep.8.2.133

Litterman, R. B. (1986). Forecasting with bayesian vector autoregressions: Five years of experience. *Journal of Business & Economic Statistics*, *4*(1), 25–38. Retrieved from https://www.jstor.org/stable/1391384 (Publisher: [American Statistical Association, Taylor & Francis, Ltd.]) doi: 10.2307/1391384

Lukito, J., Suk, J., Zhang, Y., Doroshenko, L., Kim, S. J., Su, M.-H., ... Wells, C. (2020). The wolves in sheep's clothing: How russia's internet research agency tweets appeared in u.s. news as vox populi. *The International Journal of Press/Politics*, *25*(2), 196–216. Retrieved from https://doi.org/10.1177/1940161219895215 (Publisher: SAGE Publications Inc) doi: 10.1177/1940161219895215

Maares, P., & Hanusch, F. (2020). Exploring the boundaries of journalism: Instagram micro-bloggers in the twilight zone of lifestyle journalism. *Journalism*, *21*(2), 262–278. Retrieved from https://doi.org/10.1177/1464884918801400 (Publisher: SAGE Publications) doi: 10.1177/1464884918801400

Marwick, A., & Lewis, R. (2019). Media manipulation and disinformation online. *Data & Society*, 106.

McCrow-Young, A. (2021). Approaching instagram data: reflections on accessing, archiving and anonymising visual social media. *Communication Research and Practice*, *7*(1), 21–34. Retrieved from https://www.tandfonline.com/doi/full/10.1080/22041451.2020.1847820 doi: 10.1080/22041451.2020.1847820

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, *22*(3). Retrieved from https://pubmed.ncbi.nlm.nih.gov/23092060/ (Publisher: Biochem Med (Zagreb))

Metaxas, P. T., & Mustafaraj, E. (2012). Social media and the elections. *Science*, *338*(6106), 472–473. Retrieved from https://science.sciencemag.org/content/338/6106/472 (Publisher: American Association for the Advancement of Science Section: Policy Forum) doi: 10.1126/science.1230456

Mundt, M., Ross, K., & Burnett, C. M. (2018). Scaling social movements through social media: The case of black lives matter. *Social Media + Society*, *4*(4), 2056305118807911. Retrieved from https://doi.org/10.1177/2056305118807911 (Publisher: SAGE Publications Ltd) doi: 10.1177/2056305118807911

Nimmo, B. (2020). The breakout scale: Measuring the impact of influence opera-

tions. *Brookings Institute*, 15.

Nye, J. S. (1990). Soft power. *Foreign Policy*(80), 153–171. Retrieved from https://www.jstor.org/stable/1148580 (Publisher: Washingtonpost.Newsweek Interactive, LLC) doi: 10.2307/1148580

O'Brien, L. (2017). Twitter ignored this russia-controlled account during the election. team trump did not. *Huffington Post*. Retrieved from https://www.huffpost.com/entry/twitter-ignored-this-russia-controlled-account-during-the-election_n_59f9bdcbe4b046017fb010b

O'Sullivan, D., Griffin, D., Devine, C., & Shubert, A. (2019). Russia is backing a viral video company geared at american youth. *CNN*. Retrieved from https://www.cnn.com/2019/02/15/tech/russia-facebook-viral-videos/index.html

Parmelee, J. H., & Bichard, S. L. (2011). *Politics and the twitter revolution: How tweets influence the relationship between political leaders and the public*. Lexington Books. (Google-Books-ID: ZcVxGdegAJ8C)

Paul, C., & Matthews, M. (2016). *The russian "firehose of falsehood" propaganda model: Why it might work and options to counter it*. RAND Corporation. Retrieved from http://www.rand.org/pubs/perspectives/PE198.html doi: 10.7249/PE198

Pew Research Center. (2021). *Use of online platforms, apps varies – sometimes widely – by demographic group.* Retrieved from https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2021/04/PI_2021.04.07_social-media_0-03.png

Ratkiewicz, J., Conover, M. D., Meiss, M., Goncalves, B., Flammini, A., & Menczer, F. (2011). Detecting and tracking political abuse in social media. , 8.

Redfish Stream. (2021). *Redfish youtube channel*. Retrieved from https://www.youtube.com/channel/UCHgnQzZY7T9TxhI40BmKJwQ

Riquelme, F., & González-Cantergiani, P. (2016). Measuring user influence on twitter: A survey. *Information Processing & Management*, *52*(5), 949–975. Retrieved from http://arxiv.org/abs/1508.07951 doi: 10.1016/j.ipm.2016.04.003

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, *311*(5762), 854–856. Retrieved from https://science.sciencemag.org/content/311/5762/854 (Publisher: American Association for the Advancement of Science Section: Report) doi: 10.1126/science.1121066

Select Committee on Intelligence. (2019). *Report of the select committee on intelligence united states senate on russian active measures campaigns and interference in the 2016 u.s. election* [Senate Intelligence Committee Report]. Retrieved from https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume1.pdf

Selvage, D. (2019). Operation "denver": The east german ministry of state security and the KGB's AIDS disinformation campaign, 1985–1986 (part 1). *Journal of Cold War Studies*, *21*(4), 71–123. Retrieved from https://doi.org/10.1162/jcws_a_00907 doi: 10.1162/jcws_a_00907

Silva, M., Giovanini, L., Fernandes, J., Oliveira, D., & Silva, C. S. (2020).

Facebook ad engagement in the russian active measures campaign of 2016. *arXiv:2012.11690 [cs]*. Retrieved from http://arxiv.org/abs/2012.11690

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, *48*(1), 1–48. Retrieved from https://www.jstor.org/stable/1912017 (Publisher: [Wiley, Econometric Society]) doi: 10.2307/1912017

Snegovaya, M., & Watanabe, K. (2021). *The kremlin's social media influence inside the united states: A moving target.* Retrieved from https://www.4freerussia.org/the-kremlin-s-social-media-influence-inside-the-united-states-a-moving-target/

Stubbs, H. (2020). Russian operation masqueraded as right-wing news site to target u.s. voters. *Reuters*. Retrieved from https://www.reuters.com/article/usa-election-russia-disinformation-idUSKBN26M5OP

Susarla, A., Oh, J.-H., & Tan, Y. (2011). Social networks and the diffusion of user-generated content: Evidence from YouTube. *Information Systems Research*, *23*(1). Retrieved from https://pubsonline.informs.org/doi/abs/10.1287/isre.1100.0339

Swanson, N. R., & Granger, C. W. J. (1997). Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association*, *92*(437), 357–367. Retrieved from https://doi.org/10.1080/01621459.1997.10473634 (Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01621459.1997.10473634) doi: 10.1080/01621459.1997.10473634

Thejas, G. S., Kumar, K., Iyengar, S. S., Badrinath, P., & Sunitha, N. R. (2019). AI-NLP analytics: A thorough comparative investigation on india-USA universities branding on the trending social media platform "instagram". In *2019 4th international conference on computational systems and information technology for sustainable solution (CSITSS)* (Vol. 4, pp. 1–8). doi: 10.1109/CSITSS47250.2019.9031050

T-Online. (2019). *Russland attackiert t-online.de nach kritischen berichten.* Retrieved from https://www.t-online.de/-/85177670

Tucker, J. A., Guess, A., Barbera, P., Vaccari, C., Siegel, A., Sanovich, S., ... Nyhan, B. (2018). *Social media, political polarization, and political disinformation: A review of the scientific literature* (SSRN Scholarly Paper No. ID 3144139). Retrieved from https://papers.ssrn.com/abstract=3144139 doi: 10.2139/ssrn.3144139

Tufekci, Z. (2013). "not this one": Social movements, the attention economy, and microcelebrity networked activism. *American Behavioral Scientist*, *57*(7), 848–870. Retrieved from https://doi.org/10.1177/0002764213479369 (Publisher: SAGE Publications Inc) doi: 10.1177/0002764213479369

Vargas, J. M., Rivera-Rogel, D., & Carrión-Salinas, G. (2020). Ecuadorian educational youtubers a perspective from the educational vision. In *2020 15th iberian conference on information systems and technologies (CISTI)* (pp. 1–6). (ISSN: 2166-0727) doi: 10.23919/CISTI49556.2020.9141077

Walker, C., Kalathil, S., & Ludwig, J. (2020). The cutting edge of sharp power. *Journal of Democracy*, *31*(1), 124–137. Retrieved from https://muse.jhu.edu/article/745959 doi: 10.1353/jod.2020.0010

Walton, C. (2016). *"active measures": a history of russian interference in US elections.* Retrieved from https://www.prospectmagazine.co.uk/science-and-technology/active-measures-a-history-of-russian-interference-in-us-elections

Wardle, C., & Derakhshan, H. (2017). *INFORMATION DISORDER: Toward an interdisciplinary framework for research and policy making.*

Watson, B. (2016). The troubling evolution of corporate greenwashing. *the Guardian*. Retrieved from http://www.theguardian.com/sustainable-business/2016/aug/20/greenwashing-environmentalism-lies-companies (Section: Guardian Sustainable Business)

Watts, D. J., & Dodds, P. S. (2007). Influentials, networks, and public opinion formation. *Journal of Consumer Research*, *34*(4), 441–458. Retrieved from https://academic.oup.com/jcr/article-lookup/doi/10.1086/518527 doi: 10.1086/518527

Webster, J. G. (2014). *The marketplace of attention: How audiences take shape in a digital age*. MIT Press. (Google-Books-ID: iBNmBAAAQBAJ)

Weisburd, A., Watts, C., & Berger, J. (2016). Trolling for trump: How russia is trying to destroy our democracy. *War on the Rocks*. Retrieved from https://warontherocks.com/2016/11/trolling-for-trump-how-russia-is-trying-to-destroy-our-democracy/

Wells, C., Zhang, Y., Lukito, J., & Pevehouse, J. C. W. (2020). Modeling the formation of attentive publics in social media: The case of donald trump. *Mass Communication and Society*, *23*(2), 181–205. Retrieved from https://doi.org/10.1080/15205436.2019.1690664 (Publisher: Routledge _eprint: https://doi.org/10.1080/15205436.2019.1690664) doi: 10.1080/15205436.2019.1690664

Zhang, Y., Lukito, J., Su, M.-H., Suk, J., Xia, Y., Kim, S. J., ... Wells, C. (2021). Assembling the networks and audiences of disinformation: How successful russian IRA twitter accounts built their followings, 2015–2017. *Journal of Communication*, *71*(2), 305–331. Retrieved from https://doi.org/10.1093/joc/jqaa042 doi: 10.1093/joc/jqaa042

10 Appendix

10.1 Figures and Tables



Figure 16: LSTM Classification Matrix for Instagram Topic Reposts

Table S7: Model for Repost Account Type by Govlochenko et al. (2018)

Group	Description
State Institutions	Governmental or intergovernmental institutions
Public Officials	Ministers, governmental advisers, and public officials working in intergovernmental institutions
Politicians	Members of, or candidates for, national/local parliaments, and politicians in international or supranational bodies
Commercial and State Media	Commercial or state-owned newspapers, radio stations, TV channels or news websites
Journalists	Journalists, editors, correspondents, columnists, etc. (excludes profiles who describe themselves as 'citizen journalists' as well as bloggers and journalists working for non-profit volunteer news sites)
Civil Society Group	NGOs, grassroots organizations, social movements, non-profit research centres, non-profit volunteer news sites, and citizen journalist groups
Citizens	Individual users who are not journalists, politicians or public officials (includes users whose profiles lack self-description)
Other	Profiles that fall outside any of the categories listed above (e.g. universities, think-tanks and political parties)
Removed from Twitter	Profiles no longer available on the platform.

Table S8: Model for Repost and Post Type by Howard et al. (2019)

Group	Number of Ads Purchased (e.g. "Size" of Topic)
African American Politics and Culture	841
Black Identity and Nationalism	246
Campus Politics	16
Conservative Politics and Culture	166
Immigration	60
Incarceration	10
Internet Culture	165
Latin American Culture	143
LGBT Rights & Social Liberalism	70
Music Streaming	61
Muslim American Politics and Culture	57
Native American History and Culture	18
None	753
Other	4
Patriotism	24
Pro-gun Politics	66
Self Defense	20
Social Justice	81
Texan Identity	11
Veterans & Policing	43

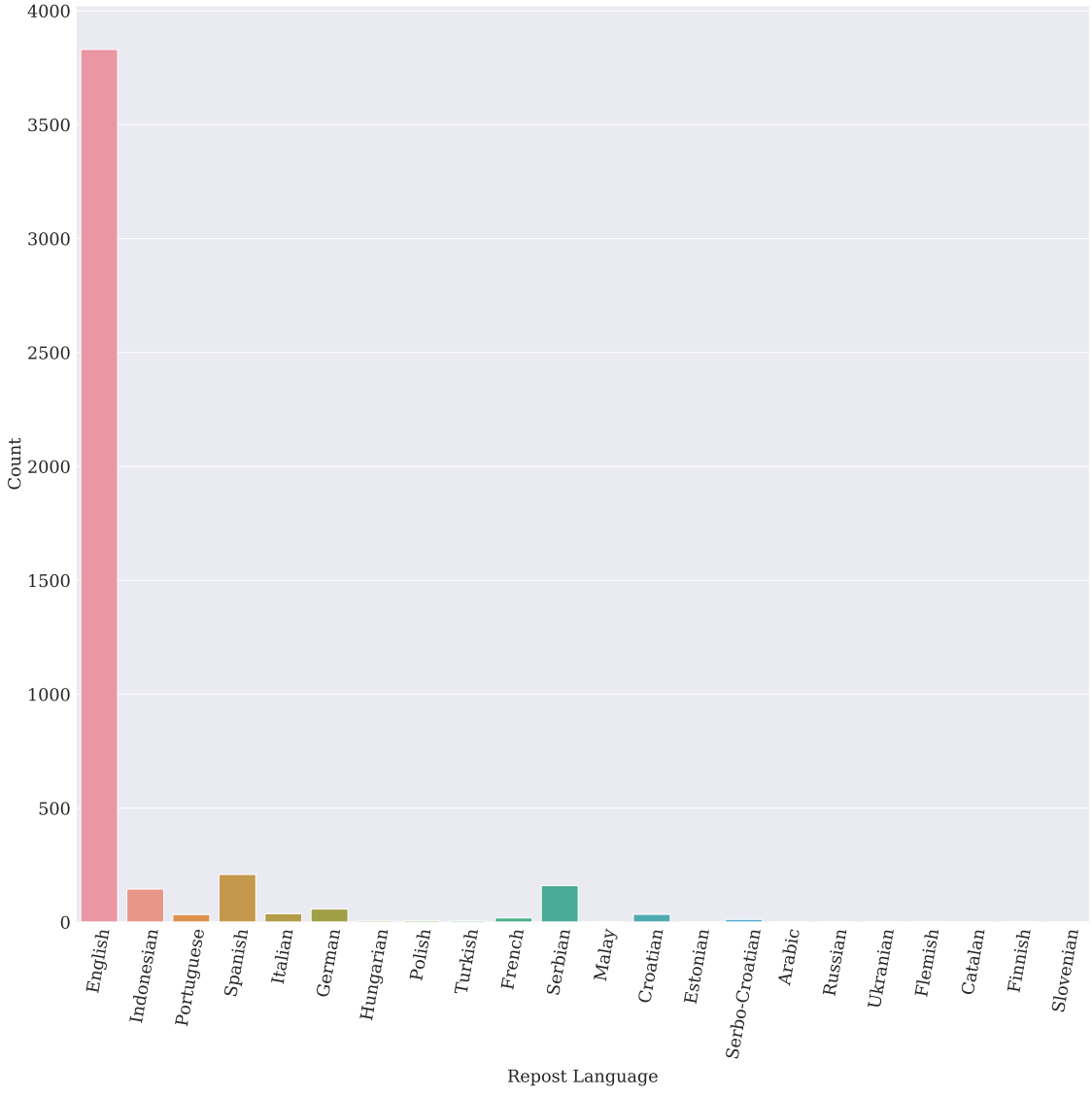
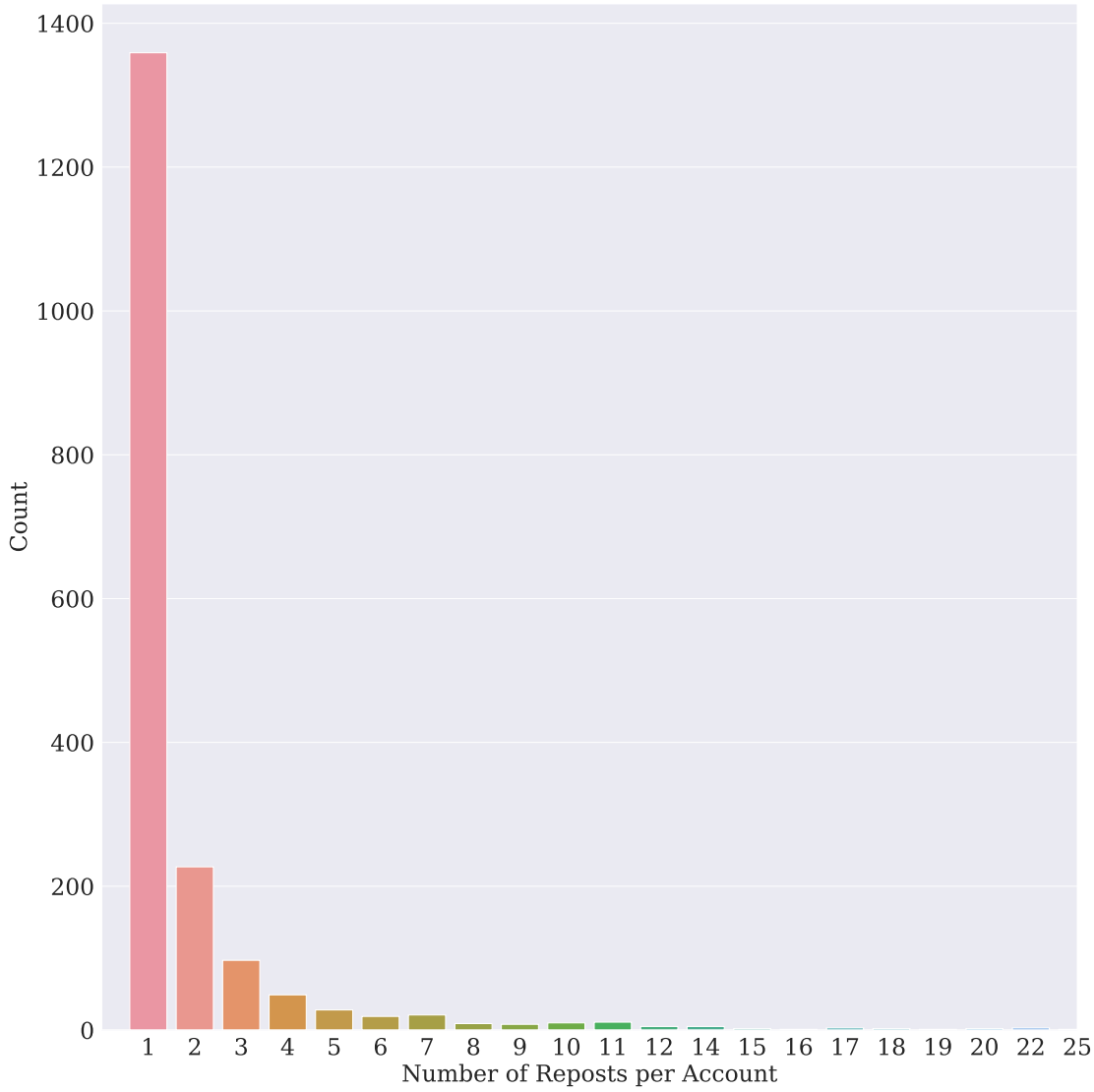


Figure 17: Breakdown of Languages in Reposts of Maffick Media Content



N.B.: Outliers have been removed from visualization.

Figure 18: Breakdown of Number of Reposts by Users that Reposted Maffick Media Content

10.2 Code Details

Given that one of the main contributions of this thesis is methodological, namely the collection of a novel data set from Instagram, I detail here the breakdown of code. The Github contains a total of 24 scripts in R and Python. I further use Tableau notebooks for some of the visualizations (e.g. Figure 5 and 12), which can be made available upon request. While the Python scripts were used to collect and visualize the data, the statistical analysis was conducted in three main R scripts. All code is appropriately cited. By publishing these scripts on Github, I hope to create a public and equitable pipeline for future research on Instagram and to lower the barriers to access.