Revisiting the disinformation resilience framework: the context of deliberate disinformation attacks

Dinara Karimova

Department of Political Science
McGill University, Montreal

December 2023
First published on September 2024

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Arts in Political Science

Abstract

The full-scale invasion of Ukraine by Russia in February 2022 was accompanied by a massive global disinformation campaign that the country launched to justify its act of aggression. Despite the measures taken, the Russian disinformation rhetoric persists in the public discourse. This thesis aims to revisit the disinformation resilience framework by analyzing the macro- and micro-level factors that enhance or debilitate resilience to disinformation in the context of a purposeful information attack by an external actor. I use the case of the ongoing Russian war on Ukraine and the YouGov-Cambridge Globalism Project survey and argue that aside from the established disinformation resilience factors, such as social media use and political ideology, it is important to account for attitudes toward the country that disseminates disinformation — Russia, in the context of this work. As such, I show that unfavorable attitudes towards Russia, measured through perceptions of the country as a threat, and the support of sanctions on Russia's energy sector are correlated with higher resilience to Russian disinformation. Other Russia-specific factors, such as volume of trade with Russia and geographical distance to it, do not reveal a significant correlation with resilience.

Résumé

L'invasion à grande échelle de l'Ukraine par la Russie en février 2022 s'est accompagnée d'une vaste campagne mondiale de désinformation que le pays a lancée pour justifier son acte d'agression. Malgré les mesures prises, la rhétorique de désinformation russe persiste dans le discours public. Cette thèse vise à revisiter le cadre de résilience à la désinformation en analysant les facteurs macro et micro qui renforcent ou affaiblissent la résilience à la désinformation dans le contexte d'une attaque d'information délibérée par un acteur externe. J'utilise le cas de la guerre actuelle de la Russie contre l'Ukraine et l'enquête de YouGov-Cambridge Globalism Project, et je soutiens que, outre les facteurs établis de résilience à la désinformation tels que l'utilisation des médias sociaux et l'idéologie politique, il est important de tenir compte des attitudes envers le pays qui diffuse la désinformation — la Russie dans le contexte de ce travail. Ainsi, je démontre que des attitudes défavorables envers la Russie, mesurées par la perception du pays comme une menace, et le soutien aux sanctions visant le secteur énergétique russe, sont corrélées à une plus grande résilience face à la désinformation russe. D'autres facteurs spécifiques à la Russie, tels que le volume des échanges commerciaux avec la Russie et la distance géographique par rapport à celle-ci, ne révèlent pas de corrélation significative avec la résilience.

Table of Contents

Table of Figures

Table of Tables

Acknowledgments

I am endlessly grateful to all the people who supported my effort on this thesis and believed in the value of this work.

Thank you to my supervisor, Prof. Dietlind Stolle, whose patience and support helped me find the research question that sparked my genuine scientific interest. I appreciate greatly your help with data, your article recommendations, your directions on the appropriate research methods and your belief that I can use them.

Likewise, I am grateful to Political behavior laboratory, led by Prof. Stolle, for being my patient audience and providing valuable feedback on my research project.

I would also like to thank Prof. Aaron Erlich and Prof. Leonardo Baccini who dedicated their time to listen to my ideas and provide some critical feedback and useful suggestions at the initial stages of my work.

Special thanks to the Data Science Solution Hub at CSDI that provided useful consultations on the methods and answered all my questions with great patience.

I would have not been able to accomplish this task without all the wonderful people that I am fortunate to have in my life. Thank you Will, Scout and Erin for always being there for me and making yourself available to brainstorm my research ideas and reassure their validity. Thank you, Oksana, for constant mentoring and mental support. Thank you to my friends back home, Adelya, Veronica and Nastya for sending support from afar and reminding me of my strong sides. Words could not tell how grateful I am to Alexis for her infinite support in my academic and research struggles, tireless advice, countless proofreads, and our friendship.

Finally, thank you to my family, who despite all the challenges and hardships, through all the distances and time zones always and forever gives me a feeling of unconditional support. I am carrying through because I feel your love across the thousands of kilometers that separate us.

1 Introduction

Over the past several years, *disinformation*, along with notions such as *fake news*, *misinformation*, and *propaganda*, has become a true buzzword, entering both the popular discourse and the field of scientific research (Kuo & Marwick 2021). Despite such wide interest and extensive academic coverage, research on certain areas of disinformation remains rather scarce and the impact of disinformation on society and various political processes is far from being completely understood. In particular, the phenomenon of disinformation resilience is one understudied topic in the field, despite the growing concern about the proliferation of disinformation.

Given the complexity of disinformation dissemination, studying factors affecting resilience to disinformation requires an equally complex approach, namely one that considers factors at both the individual and country-context levels on a given individual's degree of disinformation resilience. Nevertheless, most studies so far have focused predominantly on the individual-level of analysis, comparing resilience to disinformation among various socio-economic groups across a limited number of countries. As such, scholars have studied how factors like gender, age, education, political values, news-consumption patterns, etc. impact one's propensity to believe and/or share disinformation (see Guess 2019; Madrid-Morales 2021). Meanwhile, it is highly important to underline that macro-level factors, such as general patterns of social media use, societal levels of trust in the mass media, and the degree of political polarization in a society, among others, also have the potential to inform cross-national studies of disinformation resilience.

Moreover, deliberate disinformation attacks by external actors should also be considered as an important factor in the resilience framework, given that scholars systematically record how certain states launch international disinformation campaigns seeking to achieve their strategic objectives abroad (Arayankalam & Krishnan 2021). Being one of the examples of such information attacks, the ongoing Kremlin-backed disinformation campaign that surrounds Russia's invasion of Ukraine will provide the basis for testing the factor of deliberate disinformation attack in this thesis.

While capturing the nature of disinformation resilience at a global scale is an ambitious undertaking, a generalized global framework could provide valuable insights for a number of reasons. First and foremost, in the digital era of interconnected media systems, various forms of information disorder follow events of a global scale, and scholars frequently detect surges in disinformation following major events that leave a lasting global impact (Nelson & Patino 2021), such as the outbreak of COVID-19 in 2020 (Apuke & Omar 2021) or the ongoing war in Ukraine (Yablokov 2020). This global nature of disinformation dissemination encourages comparative analysis of the different mechanisms and outcomes of disinformation campaigns across the globe. Secondly, the individual-level factors discussed in the previous works, such as patterns of social media use, news consumption, political leanings, trust in media and others, do not exist outside of the larger socio-political context, such that the effects of individual-level factors are moderated or altered by larger political processes.

In light of this academic and sociopolitical context, this thesis seeks to contribute to the development of a comprehensive disinformation resilience framework analyzing both individual- and aggregate-level factors of resilience and whether and how the relationships between the individual-level factors and disinformation resilience are moderated by the broader socio-political context. In other words, this work seeks to offer a framework that takes into account the effect of both individual- and aggregate-level factors on disinformation resilience. In doing so, I seek to bridge the gap between research that approaches questions of resilience at either the individual or the country-level, and thereby offer a more complete picture of how resilience is built and strengthened. This will also allow me to extend some of the existing findings to a wider sample of political and media systems, and consider the contemporarily relevant factor of deliberate disinformation attacks by an external actor, such as Russia. With these goals in mind, this work will examine the ongoing global disinformation campaign that was launched by the Russian state to justify its aggression against Ukraine.

Importantly, given its focus on the deliberate information attacks by external actors, this thesis recognizes a crucial distinction between government-led disinformation campaigns and grassroot disinformation narratives (Wardle & Derakshan 2017). Consequently, this research assumes that resilience to government-led disinformation is based on different factors as compared to factors of resilience to grassroot disinformation, and therefore, only concerns the former case and the factors specific to it.

The government- and grassroot-led disinformation narratives are frequently interlinked, with similar rhetoric featured across both types of disinformation efforts and with the actors behind these campaigns providing support to each other (Bueno 2023). However, there exist persuasive arguments that governments of particular regime types, which are not held accountable for providing truthful information to their citizens, possess superior capacities to organize elaborate disinformation campaigns by using sophisticated methods (Paul & Matthews 2016), investing significant financial and other resources (Mohan & Wall 2019) and pursuing specific goals (La Cour 2020). In contrast, the sporadically flaring grassroots disinformation narratives frequently do not pursue specific goals and only occasionally reach broad audiences in both free and unfree political regimes due to the proliferation of communication technologies

(Shu 2020). As a result, the conclusions of this work should not be extrapolated to all instances of disinformation proliferation but rather they should inform future work on factors of resilience in a specific context of government-led disinformation campaigns.

To my knowledge, there exists only one widely cited study that attempts to incorporate some of the factors mentioned above, and which, therefore, will be the foundation of this thesis. In their work, "Resilience to Online Disinformation: A Framework for Cross-National Comparative Research," Humprecht et al. (2020) seek to build a general framework of disinformation resilience that looks at how populism, polarization, social media use, public media availability and other factors that will be discussed in more detail below, relate to people's exposure to disinformation across eighteen Western democracies. The study reveals that aside from factors such as trust in news media and social media use for news consumption, the size of online media market correlates significantly with the circulation of disinformation (Humprecht et al. 2020).

The goal of this thesis is to propose an updated and generalized version of the existing resilience framework. To do so, I argue that both individual-level characteristics and institutional specificities are important factors of resilience to disinformation. Moreover, I put forward some original hypotheses concerning the factors of resilience to the Kremlin-led disinformation campaign, such as the perception of Russia as a threat and support of sanctions on its energy sector, to investigate whether and under which circumstances deliberate informational attacks by other states might correlate with lower or higher resilience. I base my research on the data from the YouGov-Cambridge Globalism Project and implement multilevel modeling to provide empirical evidence on the correlation between the (1) levels of disinformation resilience among the respondents in 25 countries regarding the reasons behind the ongoing Russian aggression against Ukraine, (2) individual-level characteristics such as political leanings, perception of Russia as a threat, and energy-sector policy preferences, and (3) macro-level factors such as societal levels of social media use, levels of media freedom, repressiveness of political regime, share of total trade share with Russian and distance to it. As such, I seek to show that, aside from the established factors of disinformation resilience, it is important to consider whether disinformation is spread deliberately by a foreign state and what the relationships are between that state and a country under disinformation attack. Thus, my thesis aims to address some of the existing gaps in the current literature. In particular, I aim to complement the Humprecht et al. resilience framework by extending the research in three ways: (1) by using the current framework to test the case of the ongoing Russian war in Ukraine, (2) by considering additional predictor variables, such as perception of Russia as a threat, support of sanctions on Russian energy sector, trade with and distance to Russia; and (3) by applying the model to a larger subset of countries, including democratic and non-democratic regimes.

The remainder of the thesis proceeds as follows. I will first review the position of existing literature on factors that impact the process of information acquisition, addressing some of the most common explanations as to why humans often fail to acquire correct information. Connecting it to the topic of disinformation, I will provide a theoretical discussion on disinformation, misinformation, and fake news, explaining the importance of conceptual differentiation. I will then elaborate on the concept of resilience, and discuss the literature on the disinformation resilience framework. While describing the framework, I will point out some of its potential omissions and propose solutions that might help enhance its analytical power. I will postulate several hypotheses to verify and update the framework accordingly. After describing my proposition and hypotheses, I will move on to the description of the data that will be used, justifying the chosen sources along the way. I will then describe my selected methodology, explaining the mechanism behind and advantages of multilevel modeling, as well as the assumptions that underlie the method. This will be followed by a step-by-step description of model building and an interpretation of the results. I will conclude with the implications of the results, major limitations, and remarks on potential future work.

2 Literature Review

This thesis is situated at the intersection of information theory and disinformation literature. As such, I will review the main ideas related to each domain, starting from information theory and conceptual discussion of disinformation, moving on to a detailed discussion on disinformation resilience framework studies. I will also highlight some of the limitations of the Western-oriented nature of the research and link it to the discussion of the topic of deliberate disinformation attacks by illiberal states, by illustrating it with the overview of the ongoing Russian disinformation campaign that surrounds the war in Ukraine.

2.1 Information and Failure of Its Acquisition

Information is one of the most valuable assets in the contemporary world. Although the bulk of information is acquired from lived experiences, people rely a lot on second-hand expertise when cognizing about the broader world due to limited human learning capacities and time resources. Even though a classic interpretation of the *marketplace of ideas* theory states that in a "free market" reliable information will inevitably overshadow false information, it has been proven that the failure of the marketplace of ideas occurs more often than one would think (Stanley 2018).

There are a few reasons why people fail to acquire correct information. More specifically, scholars have identified cognitive and social factors that are associated with the propensity to hold false beliefs (Ecker et al. 2022). Importantly, however, some of the findings are rather debatable and the results are often inconclusive. For example, researchers are still debating whether people's ability to update incorrect information depends on their analytical skills and their degree of involvement with the information, or is determined predominantly by their partisan sentiments (Wood & Porter 2019). For example, Nyhan and Reifler (2010) conclude that updating false information is conditional on one's political values, and the incongruence of the corrective information with one's political beliefs leads to a significantly increased conviction in the previously held false information - a phenomenon known as the "backfire effect."

However, Pennycook and Rand (2019) and other scholars in subsequent studies (Erlich et al. 2023), have used the Cognitive Reflection Test (CRT) to convincingly demonstrate that one's ability to update false information correlates with one's analytical thinking capacities, irrespective of the compatibility of the information with one's political ideology. This finding has an important implication for the research on disinformation resilience. In particular, it calls for a more nuanced consideration of the relationship between political ideology and disinformation resilience than is commonly accepted in some of the existing literature. I will further address this consideration when discussing my hypotheses.

Trust is another important individual-level factor. Generally, research has shown that low levels of trust in science, media and government institutions are associated with a higher propensity to believe false information (see Lewandowsky & Oberauer 2016; Kim & Cao 2016; van der Linden et al. 2021).

Aside from the individual-level explanations for the human propensity to believe in false information, scholars also identify structural and environmental factors. For example, the estimated social consensus on the validity of information (Lewandowsky et al. 2012) and the polarization of political elites over a given topic (Flynn et al. 2017) impact how people perceive

information. Other studies look at how the corruption of institutions responsible for the provision of credible information and expert knowledge could lead to greater levels of disinformation circulation. In this regard, scholars highlight the role of mainstream news media and how — in their attempt to keep up with topical agendas and sometimes following their ideological inclinations — news media become a source of false information (Yariv et al. 2020).
On a large scale, Livingston and Bennett (2020) point out that the retrenchment of the public sector, characteristic of the neoliberal era, leads to a decline in so-called "authoritative institutions" defined by their transparency of norms and practices (such as academia and science, journalism, traditional news media, labour unions, etc.), which leads to increasing levels of disinformation circulation.

As is made clear, scholars identify both individual and societal factors that determine successful or unsuccessful acquisition of correct information: from analytical skills, political values, and levels of trust to levels of societal polarization and the quality of institutions of media and science. These studies take on different approaches and are situated in different contexts: whereas Nyhan and Reifler (2010) and Pennycook and Rand (2019) conduct their experimental studies in the US, Erlich et al. (2023) run their research in the context of post-Soviet Ukraine. In turn, Lewandowsky et al. (2012) and Flynn et al. (2017) provide a more theoretical account of the structural mechanisms of information acquisition. Despite the disparities in the methods and contexts, these studies, nonetheless, provide valuable explanations of why people acquire false information, contributing to our understanding of the phenomenon of disinformation.

False information, however, is a generic term that encapsulates many notions that are frequently used interchangeably, such as disinformation, misinformation, fake news, and others. Since this thesis focuses on *disinformation* resilience, it is important to provide a brief conceptual discussion on some of the major terms before going into a detailed review of the framework.

2.2 Conceptualizing Disinformation, Misinformation and Malinformation

In a widely cited report, Wardle and Derkhshan (2017) identify three types of information disorder: misinformation, disinformation and malinformation. Looking at both falseness and harmfulness of information, they conclude that the intentionality of erroneous information forms a basis for the conceptual differentiation between misinformation, *which is false but shared with no intention to cause harm*, and disinformation, *which is false and shared purposefully to cause damage*. Malinformation, according to this typology, is information that can be factually true but causes damage because it is not meant to be shared with third parties. Along the same lines, Hameleers et al. state that misinformation is potentially harmful "...inaccurate information…[that] does not offer a clear indication about the motivations…underlying its creation or motivation," while disinformation is "...a more extreme and politically motivated type of pseudo-information that specifies the harmful intentions of the sender of falsehood" (Hameleers et al. 2022, p. 239). Looking at disinformation as a systemic issue, Bennett and Livingston (2018) consider this "disruptive communication" to be a fundamental part of national politics in certain socio-political contexts and specific political and media settings. Importantly, disinformation does not always aim to instill beliefs in a specific piece of false information among the target audience, but rather to amplify information pollution and sow confusion, distrust, and the conviction that truthful information is never feasible (Snegovaya 2015).

While many scientists note the close nature of the two concepts and identify misinformation as an umbrella term for any type of inaccurate information within which disinformation represents a special case of intentional misinforming (see Fallis 2015; Hameleers et al. 2022), they give different assessments of their effects. In fact, some scholars claim that misinformation is almost inevitable as it lies at the heart of the gradual learning process, in which parts of the newly accepted information could be refuted if the updated knowledge proves them wrong (Lewandowsky et al. 2012). Disinformation, on the other hand, is frequently portrayed as one of the greatest security threats of the digital age that endangers the very foundations of democracy by sowing distrust in political institutions (Tenove 2020). The term itself invokes strong security concerns, as it comes from the Cold War era, when *dezinformatsiya* ("дезинформация") was part of the KGB's strategic communication operations (Holland 2006). Built on the reflexive control theory, the Soviet disinformation campaigns were aimed at altering the adversary's perception of reality through information attacks in such a way that would force them to make disadvantageous decisions (Snegovaya 2015).

2.3 Theorizing Disinformation Resilience Framework

Paradoxically, the increasing recognition of the risks associated with the intentional spread of false information has been mirrored by a rise in disinformation dissemination (Nagasako 2020). Among the most striking recent examples are the surge of disinformation that followed the outbreak of COVID-19 in 2020 (Apuke & Omar 2021), the deliberate spread of false information by high-ranking officials during the 2016 Brexit referendum (Henkel 2021), and the notorious case of the elaborate disinformation campaign waged during the 2016 presidential elections in the United States, which, according to some analysts, facilitated the victory of Donald Trump. This surge engendered a period during which the very notions of disinformation, fake news, and misinformation became firmly integrated into the everyday lexicon (Tenove 2020; Allcott & Gentzkow 2017). Wars and conflicts, by causing major interruptions in the international order, are likewise responsible for provoking massive flows of disinformation, as was the case during the 2003 Iraq War (Lewandowsky et al. 2013); and is most recently the case with the ongoing Russian war in Ukraine (Yablokov 2022).

There exists some persuasive evidence that particular characteristics of the modern global society (e.g., the abundance of social media platforms, the liberalization of cyberspace, technological progress, and increasing political polarization, etc.), heighten its vulnerability to the spread of disinformation (see Lazer et al. 2018; Bennett & Livingston 2018). Nonetheless, while the global community came to realize the importance of studying the factors of resilience to disinformation, few studies offer a general framework that could serve as a useful tool for analyzing, for example, whether and how certain political regimes facilitate higher levels of the circulation of disinformation, while others successfully prevent the deterioration of the information space and propagate resilience.

Amidst this academic scarcity, Humprecht et al. (2020) offer one of the most comprehensive frameworks to date on the topic of disinformation resilience. Implying that certain countries tend to have higher levels of disinformation circulation, the authors posit the existence of so-called “resilient systems” (Humprecht et al., 2020, p. 497). Importantly, immunity to disinformation does not imply a complete absence of disinformation, but rather the presence of effective tools designed to detect and dismiss it.

Providing a theoretical basis for their model, the authors identify seven macro-level factors that are hypothesized to predict higher or lower levels of disinformation resilience across eighteen Western democracies. They are, in turn, combined into three larger categories: a) *political* - referring to 1) levels of societal polarization and 2) the presence of populist rhetoric, both of which are expected to correlate negatively with higher resilience to disinformation; b) *media* - referring to 3) societal levels of media trust, 4) the prevalence of public service broadcasting (PSB) and 5) the level of common news resources, which are hypothesized to predict higher resilience; and c) *economic* - referring to 6) the size of the advertisement market and 7) the prevalence of social media as a news source, which are assumed to be negatively correlated with higher resilience (Humprecht et al., 2020, pp. 498-501). As a first step of their analysis, the authors create seven indices for the aforementioned factors of disinformation resilience. To do so, they use aggregate-level data from sources such as Timbro Authoritarian Populism Index (to measure populism) and the Varieties of Democracies Project (to measure polarization). Where data is available only on individual-level (media trust and social media use), they create an aggregate-level indices by averaging the individual-level measures to enable cross-country comparison (Humprecht et al., 2020, pp. 503). By comparing the values of the indices across eighteen countries, the authors reveal significant differences in levels of resilience to disinformation and group the countries accordingly. As such, the high-level resilience or "media supportive" cluster consists of Northwestern European countries[1] and Canada; the moderately resilient or "polarized" cluster consists of the PIGS countries (Portugal, Italy, Greece, and Spain); and the USA forms a separate category, as a country that is least resilient to disinformation (Humprecht et al., 2020, pp. 505-506).

To further test their theoretical assumptions, the authors study how these factors are related to the presence of disinformation. To do this, they run an OLS model with the aforementioned seven aggregate-level indices as predicting variables. The outcome variable is the aggregate-level exposure to online disinformation, which the authors also construct by averaging the individual-level responses concerning self-assessed exposure to disinformation (Humprecht et al., 2020, pp. 509). In doing so, they focus on the broader, macro-level examination of country-level differences in the levels of disinformation. Interestingly, out of the seven indices, the OLS model suggests statistically significant relationships between media trust,

[1] Finland, Denmark, the Netherlands, Germany, the UK, Norway, Belgium, Sweden, Switzerland, Ireland, Austria, France.

social media use, market size and nationally-averaged levels of self-reported exposure to disinformation (Humprecht et al., 2020, p. 507). Contrary to the theoretical expectations, lower levels of polarization and higher shares of common media sources used by the respondents are associated with higher levels of self-reported exposure to disinformation (Humprecht et al., 2020, p. 507). In an attempt to deal with these inconsistencies, Humprecht et al. evaluate the model with an alternative outcome variable, namely, a disinformation index based on external expert opinion about the levels of domestic disinformation disseminated by the government and political parties (Humprecht et al., 2020, pp. 509). With these alterations, the model shows a significant association between higher levels of disinformation and the presence of populist rhetoric and social media use, but not the other, previously significant variables.

Recognizing the limitations of reliance on self-reported levels of exposure to disinformation and other, supposedly more objective assessments of its presence in society, a subsequent study by Humprecht et al. (2021) probes the resilience framework with what the authors consider to be a more precise measurement of resilience. Rationalizing that concrete actions vis-a-vis disinformation are more indicative of resilience than subjective self-reported accounts of exposure to it, the authors define resilience as “the unwillingness to share, like, and comment on disinformation” (Humprecht et al., 2021, pp. 3). Additionally, the authors focus on the micro-level analysis of country differences and therefore measure the previously defined factors on the individual level (societal polarization is changed to extreme political leanings). As such, the authors conclude that populist party support, alternative media use, the use of social media, and extreme ideology are correlated with the respondents’ willingness to interact with disinformation (Humprecht et al., 2021, pp. 9-12).

At this point, it should be noted that one of the major challenges in quantitative studies of disinformation resilience concerns operationalizing the phenomenon itself, as there seems to be important conceptual differences between exposure to, belief in, and dissemination of disinformation – all of which are used to assess general levels of resilience to disinformation in a given society. Though all of them are indicative of the presence of resilience to disinformation, they might, in theory, point to different aspects of the same phenomenon. For example, some people might recognize that they are exposed to disinformation without believing in it substantively, while others might believe in the content of disinformation and therefore not recognize its false nature. Equally important, people might interact with disinformation for opposite motivations – for example, some would share and comment on disinformation online following their firm belief in it, while others would share false information to warn others or leave their critical opinion in the form of a comment. Such differences are not always caught by the original resilience framework, partially explaining discrepant results.

Addressing this issue, the study by Boulianne et al. (2022) seeks to look at three different aspects of misinformation[2]: *awareness* of the existence of certain misinformed topics; self-assessed *exposure* to misinformation; and its *sharing* on social media. Introducing the "awareness" measurement is beneficial as it attempts to go beyond the subjective measures that depend on one's perception of information as either false or true (such as self-reported exposure to and sharing of false information). Such a nuanced approach increases the robustness of the results since the resilience is now dependent on the three separate measures. This measurement, however, does not tell whether respondents *believe* in the disinformation. This aspect of the

[2] Here, I use the term "misinformation" to stay consistent with the authors' terminology. It could be argued that the research on misinformation is relevant to the research on disinformation, as some scholars identify misinformation as an umbrella term for any type of false information, within which disinformation represents a special case of intentional misinforming (Hameleers 2022, p. 239).

phenomenon (one that is missing from the existing studies discussed above) must be considered if the goal of a study is to trace the depth of the impacts of disinformation on individual respondents.

With this nuanced measurement of misinformation in hand, Boulianne et al. run separate multivariate models for each of the outcome variables (awareness, exposure and sharing) with the predicting variables of trust in national news media, consumption of the public broadcasting media, following various news organizations on social media, as well as right-left wing ideology and some control variables. They conclude that, contrary to the expectations of the resilience model, trust in public broadcasting is associated with higher awareness and sharing of misinformation. They also establish a statistically significant relationship between following news organizations on social media and all three outcome variables, meaning that controlling for the factors of age, gender and education, social media news use predicts higher levels of awareness about the existence of misinformed topics, higher levels of self-assessment of exposure to misinformation and greater sharing of misinformation (Boulianne et al., 2022, p. 179). They also find that right-wing ideology is linked to greater awareness and sharing of misinformation in some countries. Contrary to the theoretical predictions that the consumption of public broadcasting media increases resilience, their findings do not show a consistent relationship between the consumption of public media sources and the three measurements of disinformation resilience (Boulianne et al., 2022, p. 178). As such, from the summary of the existing research to date, it becomes apparent that the current resilience framework generates inconsistent findings. Such discrepancies can be partially explained by the different approaches in operationalizing disinformation resilience, as some studies rely on individual assessments of the presence of disinformation and others on individual actions toward it. Additionally, various levels of measurements (individual and aggregate) require different levels of analysis, leading to incomparable results. Importantly, all the studies are limited to a small sample of affluent Western liberal democracies: eighteen countries in the original work (Humprecht et al. 2020), six countries (Belgium, France, Germany, Switzerland, the UK, and the US) in the following work (Humprecht et al. 2021); and France, Canada, the UK, and the US in the work by Boulianne et al. (2022). The following section will further elaborate on some of the limitations of the existing research.

2.4 Limitations of Western-oriented Research

One of the major drawbacks of the previous studies is their strong Western-centric focus, in which the US takes on an exceptional position. This makes generalization of the findings impossible, as the results are conditioned by the existence of a free information space, which is crucial, considering the significant differences that exist between the western liberal democratic media markets and the closed media systems of illiberal regimes. In particular, we can expect that whereas trust in major media sources might be associated with resilience to disinformation in democratic systems, it might signify lower resilience in non-democratic systems. This assumption is highly plausible because in illiberal media systems, all major media sources are controlled by a government interested in dominating the information environment and propagating its own agenda. Hence, we can assume that people who are more critical and less trustful of official, government-sanctioned information are in fact more resilient to disinformation. Unfortunately, existing studies cannot test this hypothesis because their samples do not include countries with centralized control over the media, censorship, and suppression or marginalization of independent media, all of which are associated with contemporary authoritarian regimes (Guriev & Treisman 2019).

Additionally, the discussed research looks at the phenomenon of disinformation from a "domestic" perspective, perhaps implicitly assuming that in the era of globalization, information and media systems do not permeate each other at the interstate level. Yet, scholars have not only established how media systems around the world influence each other's agendas (Golan 2006), but also how certain states engage in deliberate information attacks globally (Arayankalam & Krishnan 2021). In this regard, it might be informing to study whether some of the countries are more vulnerable to foreign disinformation campaigns than others, and if so, what factors are associated with lower resilience to foreign disinformation attacks. For this thesis, I take on this approach as it might help to enrich the current studies on the disinformation resilience framework by adding an important factor of deliberate external influence on the domestic information environment.

2.5 Russian Disinformation Campaign and War in Ukraine

As discussed across political science literature, authoritarian regimes strive to maintain control over the information space in which they operate, and while they accomplish this task most successfully domestically, they also try to influence the international agenda and the media coverage of other states (Walker 2016). While disinformation is not inherent to illiberal political regimes and many instances of the so-called "computational propaganda" are shown to accompany political processes in various liberal democracies (Woolley & Howard 2018), scientists unanimously recognize a number of "not free" political regimes that have earned notorious reputations as the major 'invaders' of foreign media; most infamously Russia and China, but also Iran, Saudi Arabia and Venezuela among others (Walker 2016; Hameleers 2023; Woolley & Howard 2018 b). To make their voices heard abroad and to sow doubt about certain liberal-democratic ideals, these countries 'rely on trolls, cyberattacks and disinformation to achieve their objectives' (Walker 2016, p. 58). Aside from promoting their own political narratives through locally established media headquarters, such as RT and CCTW, both Russian and Chinese governments have been proven to interfere with various political processes across the globe, including national elections (Mohan & Wall 2019).

While it is not entirely clear whether these authoritarian regimes are themselves more vulnerable to each other's disinformation campaigns than their shared audience in democratic political systems, it is argued that their information operations are mutually influenced. This way, despite different objectives and tactics, the networks of authoritarian disinformation are "…amplifying, cross-pollinating, and learning from one another" (Kalathil 2020, p. 33). Indeed, it is plausible to assume that their disinformation narratives are interlinked, mutually influencing and find reciprocal support for their arguments, since fundamentally they all seek to compromise political principles and values of democratic systems and sow distrust and doubt in the liberal order in both consolidated and emerging/developing democratic systems (Walker 2016; Lyammouri & Eddazi 2020). These considerations were foundational in the construction of this work's hypotheses.

Looking at the Russian case specifically, the global scale of its disinformation efforts has become acutely evident after the 2016 U.S. presidential elections, for which various Russian-backed actors used social and other media channels to distort public opinion (Erlich & Garner 2023). As mentioned previously, Russian disinformation campaigns have inherited the traditions of the Soviet doctrine of "active measures" composed of various informational activities, whose goal was to "…weaken the West…, to sow discord among allies, to weaken the United States in the eyes of the people in Europe, Asia, Africa, Latin America, and thus to prepare ground in case the war really occurs" (Snegovaya 2015, pp. 14-15; see also Holland 2010; Kux 1985). When talking about the current Russian information operations, scholars often refer to the so-called

"4D approach" of dismissing, distorting, distracting, and dismaying the true facts (Snegovaya 2015, p. 13). As such, the Russian disinformation campaign is often aimed at the obfuscation of public opinion and distraction from the central arguments (e.g., "whataboutisms") (Headley 2015). For this, the government commissions a great number of channels to ensure a rapid and continuous flow of narratives devoid of any consistency and objectivity (Paul & Matthews 2016).

Importantly, previous research has shown that different information environments are either more or less vulnerable to pro-Kremlin disinformation, depending on the corresponding state's relationship with Russia. For example, research has shown that Poland, due to its specific position in the Russian-Ukrainian conflict (and Russian-European relationships in general), has become a frequent target of Russian disinformation attacks (Gorwa 2018). Similarly, the historical relationships between Germany and Russia, as well as extensive economic and business links between the two countries, have made the former one of the major targets of Russian information operations (Neudert 2018).

Arguably, however, none of the historic and contemporary targets of the Kremlin disinformation attacks have been hit as hard as Ukraine (Erlich et al. 2023). Situated next to one of the major disinformation-filled regimes on the planet and linked to it culturally, linguistically and economically, Ukraine has been experiencing massive disinformation attacks that at times have accompanied on-the-ground military actions, as Russia has attempted to redeem the remnants of past influence on the increasingly Western-oriented post-Soviet state (Snegovaya 2015; Zhdanova & Orlova 2018). Interestingly, studying the credibility of pro-Kremlin news coverage among Ukrainians prior to the full-scale invasion, Erlich and Garner (2023) have established certain individual-level correlates of the propensity to believe in Russian disinformation, which could be grouped into two overarching categories: contextual traits (or links to Russia) and political sophistication and awareness (Erlich & Garner 2023, p. 6).

The fact that disinformation campaigns often accompany on-the-ground military operations is nothing new. Warring parties engage in these operations partially to obscure their concrete military goals and intentions, but also to justify the war in the eyes of the global public. In the case of Russian aggression, the combination of conventional and information warfare, or so-called "hybrid warfare," is also crucial because it has allowed Russia to compensate for the limitations of its military capacities (Snegovaya 2015). The Russian state utilized this strategy during its war on Georgia in 2008 and during the annexation of Crimea in 2014 (Erlich & Garner 2023; Snegovaya 2015). It is hardly surprising, therefore, that in its ongoing war on Ukraine, the Russian government has relied on the same strategy of hybrid warfare (OECD, 2022). As such, the Russian state along with its state-controlled media has initiated a massive disinformation campaign to provide justification for the invasion of the sovereign state. In particular, the Russian president claimed on numerous occasions that the so-called "special military operation" was inevitable due to the existential security threat that NATO allies posed to Russia with their proxy-war through Ukraine (Government of Canada, n.d.); the seizure of political power in Ukraine by neo-Nazi factions (Azarov et al. 2023); and the "genocide" of Russian-speaking groups in Ukraine (Hinton 2022), among other false justifications.

With this focus on the Russian disinformation strategy laid out, this study seeks to revisit the resilience model by looking at the disinformation campaign that surrounds the ongoing Russian war in Ukraine. While studying more explicit links to Russia, such as the support of pro-Russian political parties or ethnic and linguistic proximity, is not feasible due to data availability, it is plausible to assume that certain cultural-historical and economic connections and even geographic proximity might all be relevant for the prevalence of Russian disinformation in a given country (see Kilkenny 2021). Given the theoretical discussion, I expect that unfavorable attitudes toward Russia, mirrored in the perception of the country as a threat and the support of sanctions on the Russian energy sector predict higher resilience to the pro-Kremlin disinformation. At the same time, per the regime literature above, I anticipate resilience levels to be lower in illiberal political regimes. I will now elaborate on these propositions.

3 Hypotheses

Based on the conceptual framework introduced above, I outline my hypotheses beginning with factors that explain disinformation resilience in general, described by Humprecht's framework, and moving on to factors of resilience specific to the Kremlin-led disinformation campaign covering the ongoing war in Ukraine. I will differentiate between hypotheses based on individual-level factors and those based on country-level factors.

3.1 Factors of Resilience from the Resilience Framework

In their studies, both Humprecht et al. (2021) and Boulianne et al. (2022) argue that ideology constitutes an important factor of disinformation resilience. Importantly, the authors look at ideology from different perspectives (e.g., Humprecht et al. measure extremist views, while Boulianne et al. look at how right-wing extremism specifically is correlated with resilience compared to left-wing extremism) and do not arrive at comparable results. Specifically, the results for comparing the effects of right-wing and left-wing ideologies are not consistent across the studied countries (Boulianne et al. 2022). Givent that the literature suggests that disinformation feeds on political and ideological disagreements in general and that left-wing extremism is as prone to disinformation as right-wing (Hameleers 2023; Bennett & Livingston 2018), I propose to look at how centrist political values (compared to all other political leanings) are correlated with resilience. In doing so, I seek to examine ideology from yet another perspective, and analyze whether and how moderate political values are correlated with resilience. It is informative, therefore, to see whether *centrist political leanings are associated with higher resilience to Russian disinformation (H1).* To test this individual-level hypothesis, I create a dummy variable with 1 indicating respondents with centrist political values and 0 indicating the rest to isolate the centrist respondents from others (see Table 3 in the Appendix).[3] Additionally, following on from the previous findings, I run two alternative models, comparing the effect of left-wing versus right-wing leanings on resilience in the first model, and combining the two extremes in the second (see discussion below).

Another factor featured in the resilience framework — the use of social media — has also shown to have fairly persistent statistically significant effects. Considering previous findings on the correlation between general social media use and increasing probability to encounter disinformation, I expect that *a lower percentage of social media use in a given society is associated with higher resilience to Russian disinformation (H2).* To test this country-level hypothesis, while accounting for the fact that social media systems vary from one country to another, I use data on social media prevalence by country compiled by the Digital News Report. Specifically, I calculate the proportion of the total population using the most popular social media platform in each country (see Table 4 in the Appendix).

Additionally, the resilience framework focuses on the effect of news media, looking at how the presence of public service media correlates with disinformation. Although the effect of this factor is not consistent throughout the studies, it is built on a compelling theoretical foundation that the availability of numerous information sources "increases the overall quality" of information, thereby decreasing the influence of disinformation (Humprecht et al. 2020, p.

[3] A full table of variables, their measurements/scales and sources are available in the Appendix section (Table 3).

119). Since the variety of information sources is an intrinsic quality of free media systems, I argue that *higher levels of media freedom predict higher resilience to Russian disinformation (H3).* This is plausible because media freedom facilitates the emergence of various public broadcasting and media sources that are forced to compete for an audience by, in part, increasing the quality of provided information and reducing the value of false information (Aalberg & Cushion 2016). To assess a country's level of media freedom, I create a five-point scale index from 1 (extremely unfree media systems) to 5 (fully free media systems) based on the World Press Freedom Index.

Following on the previous hypothesis, as well as on earlier discussion of modern authoritarian regimes, i.e., informational autocracies (Guriev & Treisman 2019), characterized by centralized control over media, I presume that *authoritarian and other non-free political regimes are associated with lower resilience to Russian disinformation (H4).* To assess the degree of freedom in the political regime, I construct a three-point scale index from 1 (unfree) to 3 (free), using the conventional Freedom House regime classification. Since the media freedom and political regime variables are likely highly correlated, including both variables in my model might lead to estimation problems linked to multicollinearity. With the additional data limitation concerning the political regime variable – namely that the authoritarian countries featured in the data set are located almost exclusively in the Middle East (with the exception of Thailand) – I also provide an alternate model that excludes political freedom and looks at the isolated effects of media freedom (see "Alternate (Final) Model Estimation (Excluding Political Regime Variable)" in the Appendix). At the same time, considering media and regime factors separately is justified because while control over media is one of the key characteristics of illiberal regimes, such regimes can debilitate their citizens' abilities to discern false information in other ways. For example, it is argued that individuals in non-democratic systems are often withdrawn from broader political and social processes and hence have limited motivation, resources, and tools for meaningful engagement with new information, whether it is true or false (Alyukov 2022).

3.2 Russia-Specific Factors of Resilience

The second set of hypotheses is contextual and concerns the message factors of the specific case of disinformation under study: its source (Russia), the subject (the war in Ukraine) and its target audience (respondents across the 25 sample countries). This work draws on the aforementioned study by Erlich and Garner (2023) which establishes a positive relationship between close links with Russia (such as the support of pro-Russian political parties or self-identification with the Russian ethnicity) and a higher propensity to believe Russian disinformation, to incorporate context-specific factors into the resilience model. Due to the difficulty of establishing a universal measurement of close links with Russia, I propose that favorable attitudes towards Russia and perceptions of it as a non-threat are indicative of the desire to maintain links with it, while perceiving Russia as a threat and supporting the establishment of anti-Russia defence infrastructure signals the opposite. Therefore, I argue that *an individual-level policy preference favorable of defence against Russian aggression (as opposed to maintaining trade and diplomatic relationships) predicts higher resilience to Russian disinformation (H5).* To test this proposition, I create a three-point scale variable that indicates whether a respondent prefers to invest in diplomacy with Russia (-1), does not have a strong preference for either diplomacy or defence (0), or prefers to invest in defence and security (1) based on a corresponding question from the YouGov-Cambridge Globalism Project survey.

The energy dependency of Russia's economic partners has also been actively exploited as a tool of political manipulation, with the Kremlin threatening to cut oil and gas supplies if the political decisions of importing states somehow contradict its interests (Collins 2017; Krickovic

2015). Unsurprisingly, the threat of a potential energy crisis has since been extensively used in the pro-Kremlin disinformation campaign, feeding and magnifying the voices of those worried about cold winters without Russian gas (Bergengruen 2023). Because possible Russian energy withdrawal (according to the Kremlin's narrative) might have potential tangible effects on people's everyday lives, I hypothesize that *individual-level willingness to bear higher energy costs due to sanctions on the Russian energy sector is associated with higher resilience to Russian disinformation (H6).* For this, I create a five-point scale of respondents' willingness to pay higher energy costs based on the corresponding question from the YouGov-Cambridge Globalism Project survey, from -2 (strong opposition to increasing costs), -1 (opposition to increasing costs), 0 ("don't know"), 1 (support of increasing costs) to 2 (strong support of increasing costs). Moreover, since Russia's disinformation campaign could effectively target any one of a country's economic sectors (e.g., agricultural sector) substantially reliant on Russian exports (e.g., wheat), it can be argued that *a lower fraction of total trade with Russia predicts higher resilience to Russian disinformation (H7)*, given that this limits the opportunities of the Russian government to manipulate trade-related information. In other words, individuals living in countries that are economically dependent on Russia are more likely to have a positive view of the Russian government and are therefore less likely to be resilient to Russian disinformation campaigns. To measure this country-level factor of trade volumes with Russia, I divide the dollar value of exports to and imports from Russia by the total value of all international trade for each country, based on the Michigan State University's Market Potential Index (MPI) (See Figure 4 in the Appendix).

Finally, it would be reasonable to look at how the factor of distance correlates with disinformation resilience. Geographic proximity to Russia can incentivize more critical assessments of any statements made by the country's officials or government-controlled media, since the Russian state has compromised its credibility on multiple occasions and is largely perceived as an existential threat by its immediate neighbors (Pezard et al. 2017). Therefore, I argue that *geographic proximity to Russia increases resilience to Russian disinformation (H8).*

Table 1: Hypotheses

	General factors of resilience	**Unit**
H1	*Centrist political leanings are associated with higher resilience to Russian disinformation.*	*1- centrists; 0-all the rest*
H2	*Lower percentage of social media use in a given society is associated with higher resilience to Russian disinformation.*	*proportion of population using the most popular social media platform*
H3	*Higher levels of media freedom predict higher resilience to Russian disinformation.*	*1-extremely unfree; 2-unfree; 3-partially free 4-free; 5-fully free*
H4	*Authoritarian and other non-free political regimes are associated with lower resilience to Russian disinformation.*	*1-unfree; 2-partly free; 3-free*
	Russia-specific factors of resilience	
H5	*Individual-level policy preference favorable of defence against Russian aggression predicts higher resilience to Russian disinformation.*	*-1- invest more in trade and diplomacy;* *0-don't know/neither of this;* *1-invest more in defence and security*
H6	*Individual-level willingness to bear higher energy costs due to sanctions on the Russian energy sector is associated with higher resilience to Russian disinformation.*	*-2-strongly oppose; -1-tend to oppose; 0 - Don't know; 1-tend to support; 2-strongly support*
H7	*Lower fraction of total trade with Russia predicts higher resilience to Russian disinformation.*	*proportion of imports/exports from/to Russia out of total exports/imports*
H8	*Geographic proximity to Russia increases resilience to Russian disinformation.*	*capitals' distance in km (thousands) to Moscow*

4 Methodology

4.1 Data and Measurements

In order to test the above hypotheses, I draw on multiple data sources. The majority of data comes from a nationally representative survey conducted by the data analytics firm, YouGov, for the YouGov-Cambridge Globalism Project.[4] The individual-level data was collected from August 24 to September 22, 2022, in the midst of acute military activity, and includes, among other things, questions about the reasons justifying the start of the war (including the reasons pushed by the Russian government in its disinformation campaign), the best possible outcome for Ukraine, and potential relationships between the surveyed countries and Russia. A total number of 25,939 responses were collected in the UK, Canada, the US, Australia, France, Germany, Italy, Spain, Sweden, Denmark, Poland, Hungary, Greece, Mexico, Brazil, Egypt, Turkey, Saudi Arabia, Nigeria, Kenya, South Africa, India, Japan, Indonesia and Thailand. Such a wide range of countries both permits to test some of the existing findings in the resilience literature and extend them beyond the traditional scope of Western democracies. It also enables the exploration of some of the original context-specific hypotheses discussed above, given the diversity of regime types, media freedom and relationships with Russia among the selected countries.

The construction of the disinformation resilience index used in this work is based on responses to the following prompts:

1. "Before the war started, ethnic Russians living in Ukraine were being subjected to mass murder – or 'genocide' – by Ukrainians";

---

[4] See https://docs.cdn.yougov.com/0ma5boayqk/Globalism%202022%20-%20The%20info%20war%20for%20Ukraine%20-%20All%20markets.pdf.

2. “Before the war started, the Ukrainian Government had fallen under the influence of militant extremists who supported the ideology of Nazism and Adolf Hitler”;
3. “Before the war started, Western countries were seeking to establish a military infrastructure in Ukraine in order to bully and threaten Russia”.

Given that these specific questions mimic the Kremlin disinformation rhetoric, I assigned 0 points on the resilience scale to the respondents who identified these options as “definitely true”; 1 for “probably true”; 2 for “either can be true”; 3 for “probably false”; and 4 for “definitely false.” I then combined the scores for these three items and got a 12-point scale of disinformation resilience from 0 (lack of any resilience) to 12 (fully resilient). The distribution of resilience scores and the average resilience score by country are shown in Figures 5 and 6 in the Appendix respectively.

In choosing the survey prompts I was guided by several principles. First of all, I picked prompts that contained disinformation discourses that are actively pushed by the Russian state and state media, rather than general speculations about the invasion of Ukraine. Secondly, I picked the prompts that offered the same structure of response options, to maintain the inner consistency of the index. Thus, I had to exclude the fourth question which contained a disinformation claim as a response option, rather than being contained in the prompt itself. To validate the internal consistency of the resilience index, I performed a factor analysis with a varimax rotation which showed that 54% of the variance in the disinformation resilience scale can be explained by one latent factor with an Eigenvalue of 1.63. After establishing that all the items measure the same latent variable, I ran a reliability test (i.e., Cronbach alpha) to assess how well the identified items measure this variable. With $\alpha = 0.78$, the test showed a stronger correlation between the responses to the first three questions and a higher internal consistency of the disinformation resilience scale, compared to when the fourth question is included.

To create the trade variable, I drew on the Michigan State University's Market Potential Index (MPI), which provides one of the most comprehensive overviews of economic activities for nearly all of the countries in the world (globalEDGE, 2021). In particular, it includes yearly data on countries' total exports and imports, as well as information on their import and export volumes to other countries, allowing me to assess their trade share with Russia out of total trade value.

For the media-freedom index, I used the World Press Freedom Index (WPFI) published by Reporters Without Borders (RWB). This index represents a five-level classification of countries based on their media freedom scores, with the lowest level representing countries with the least favourable conditions for media freedom, and the highest level representing countries with the most favourable conditions for media freedom (Reporters Without Borders, 2022).

For the overall democracy score, I used the standard Freedom in the World index by Freedom House, which assesses political rights and civil liberties in countries on a 7-point scale and ranks them as free, partially free and unfree in accordance with their rights and liberties scores (Freedom House, 2022).[5]

I used Centre d'Etudes Prospectives et d'Informations Internationales' (CEPII) data on the geodesic distances between countries' capitals to estimate the geodesic distance between Moscow and a given country's capital (Mayer & Zignago 2011). For scale-consistency considerations, I incorporated this data in thousands of kilometers and rounded to three decimal places. While using the distances between countries' capitals is justified on the ground that most

[5] I also use an alternative measurement of democracy by V-Dem. For the comparison of the results for the two measurements see Table 8 in the Appendix.

economic and political interactions usually occur between capitals, I acknowledge that this measurement is imperfect. In particular, a country might share an extensive border with Russia, while the capital might be located a considerable distance from Moscow. Future research might consider using the presence of a common border or border distances as an alternative measurement.

Social media use data proved to be most challenging as different countries have different social media ecosystems. I decided to consider the percentage of the population using the most popular social media platform in each country, with data provided by the Digital News Report on internet use (Digital News Report, 2022). More specifically, I divided the penetration rate for the most popular social media platform by the total internet penetration rate. For the missing data on Saudi Arabia and Egypt I used analogous data from Statista (Statista, 2022).

Importantly, all the aggregate-level data was gathered from 2022 to capture the trends that were unfolding as the main survey was conducted. The trade data is an exception in that it summarizes information from the year preceding Russia's full-scale invasion of Ukraine, to capture general patterns prior to major changes in global market dynamics, including the import of energy resources.

4.2. Multilevel Modeling: General Form and Reasons for Application

As was argued previously, resilience to disinformation is an individual-level characteristic that is conditioned by the broader societal context. Therefore, this work will employ multilevel modeling – a type of quantitative analysis that is commonly used when the subject of study assumes a relationship between individuals and the society they exist in (Hox et al., 2017).[6]

---

[6] Multilevel modeling possesses a few advantages over standard regression models when the data includes both individual- and aggregate-level observations. For example, standard statistical models rely on the assumption of

In its general form, a multilevel model constitutes "…a regression (a linear or generalized linear model) in which the parameters – the regression coefficients – are given a probability model" (Gelman & Hill 2006, p. 1). A multilevel model with a single independent variable on an individual level and an independent variable on a group level has the following structure (Woltman et al. 2012; Peugh 2010):

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij}$ (1)
Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}G_j + u_{0j}$ (1.1)
$\beta_{1j} = \gamma_{10} + \gamma_{11}G_j + u_{1j}$ (1.2)

In equations 1-1.3, *i* refers to an observation at the lowest level (individuals), and *j* refers to the group (countries, in my case), such that Y_{ij} is the value of an outcome variable for i^{th} observation in j^{th} group. ε_{ij} is a level-1 error term that follows a normal distribution *N* ~ (0; σ^2). It can be noted that the first-level equation (1.1) resembles a simple linear model, but the subindex *j* indicates that the first-level model is estimated for each unit of the group-level variable. In turn, second-level equations (1.2, 1.3) indicate that the level-1 intercept (β_{0j}) and the slope (β_{1j}) represent a function of level-2 parameters: γ_{00}, γ_{10}, (true means across an entire population) and γ_{01}, γ_{11} (average effects of group-level variable, *G*, across all level-2 units) that do not vary across level-2 groups, and level-2 error terms u_{0j}, u_{1j} that follow multivariate normal distributions with variances Var(u_{0j}) = τ_0^2 and Var(u_{1j}) = τ_1^2 accordingly. A combination of the two equations yields the following equation:

independence of observations. If this assumption is violated (which is the case with my data, where individuals are nested in countries), the estimates of standard errors have to be adjusted to reflect this nested data structure. Addressing this challenge, multilevel modeling allows for the combination of aggregate- and individual-level data without inflating statistical significance of highly correlated variables (Hox et al. 2017). Additionally, it allows one to avoid the issue of the ecological fallacy that was criticized in previous studies on the resilience framework (Boulianne 2022). Because multilevel modeling exploits both aggregate- and individual-level variables simultaneously, it does not bluntly apply the results of the aggregate-level analysis to individual-level outcomes (Subramanian et al. 2009).

$$Y_{ij} = \gamma_{00} + \gamma_{01}G_j + \gamma_{10}X_{ij} + \gamma_{11}G_jX_{ij} + \varepsilon_{ij} + u_{0j} + u_{1j}X_{ij} \quad (1.3)$$

4.3 Model Assumptions

Most of the assumptions that underlie multilevel modeling represent an extension of the assumptions of simple linear regression models, such as linearity, homoscedasticity, and independence of errors. In their work, Hox et al. (2017) offer to focus on the normality assumption, according to which residuals (at all levels) are normally distributed.

For the preliminary visual inspection of the assumption of the normality of residuals, I created QQ plots for the two final models (one with the Political Leanings variable (M3 Full) and one without (M3)) (see Figure 7 in the Appendix). These plots show the standardized residuals relative to the normal quantiles and can indicate if the distributions of residuals are not normal. It can be noted that there are some minor outliers at the first and last quintiles but in general the line looks straight. Nevertheless, I further inspected the distribution of individual- and country-level residuals (see Figures 8-9 in the Appendix) and confirmed that the outliers occurred because some country-level residuals are not normally distributed. This violation of the normality assumption is potentially problematic because the maximum likelihood estimation assumes normal distribution of error terms (Maas & Hox 2004). While this does not affect the estimates that remain asymptotically unbiased and consistent, standard errors are downward biased and confidence intervals cannot be trusted. Replacing the maximum likelihood estimation method with robust errors helps address the issue (ibid.).

It is also important to acknowledge that, due to the nature of the outcome variable as a bounded integer between 0 and 12, I could not confirm the linearity assumption. In fact, when residuals were plotted against the fitted values, a strong diagonal pattern emerged (Figure 10 in the Appendix). This trend is only natural if we consider that for the fitted value of 1.5, the only possible values of the residual are: -1.5, -0.5…9.5, 10.5. As a result, fitting the predictors with the Poisson or ordinal models might represent a more optimal strategy and any subsequent studies could benefit from experimenting with different model specifications. Within the scope of this project, however, I accept the limitations that the violation of this assumption implies and interpret the results accordingly.

4.4 Model Construction

Multilevel analysis assumes sequential testing of increasingly complex models (Gelman & Hill 2007, 251) and begins with the basic model of unconditional means (Peugh 2010):

Level 1: $Y_{ij} = \beta_{0j} + \varepsilon_{ij}$ (2)
Level 2: $\beta_{0j} = \gamma_{00} + u_{0j}$ (2.1)
Combined: $Y_{ij} = \gamma_{00} + \varepsilon_{ij} + u_{0j}$ (2.2)

In equation (2), Y_{ij} is the disinformation resilience for i^{th}person in country j and is modeled as a function of the mean disinformation resilience for country j (β_{0j}) plus as a residual term that represents the individual differences around the mean of the country j (ε_{ij}). Equation 2.1 defines the disinformation resilience mean for country j (β_{0j}) as the function of a grand-mean (or "true mean" in some literature) of the disinformation resilience score (γ_{00}) and as a country-specific deviation from the grand mean (u_{0j}). The error terms are assumed to follow normal distributions: $\varepsilon_{ij} \sim N(0;\ \sigma^2)\ and\ u_{0j} \sim N(\gamma_{00};\ \tau_0^2)$. Equation 2.2 combines the two equations.

While not considering individual- or country-level factors, this model constitutes a crucial first step because it allows for the estimation of the interclass correlation coefficient (ICC), which, in turn, serves as a quantitative criterion for the feasibility of using multi-level regression modeling in the first place, by showing the proportion of the total variance in the dependent variable that is explained by the between-group variance: $ICC = \frac{\tau_{00}}{\tau_{00}+\sigma^2}$.[7]

Gelman and Hill (2006) argue that ICC values starting at 0.05 justify the use of multi-level modelling (Gelman & Hill 2006, p. 449). With an ICC of 0.16 in my unconditional means model, I conclude that there is enough group-level variation to account for a multilevel structure and proceed with the multilevel analytical approach. The next intermediate steps of the model construction are detailed in the Appendix section ("Intermediate Steps of Multilevel Model Building").

In the final step, I explain the intergroup variation in intercepts by including the country-level predictors: social media use, political regime, media freedom, trade with and geographic distance to Russia:

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}Polit\_Lean_{ij} + \beta_{2j}Energy\_Policy_{ij} + \beta_{3j}Rus\_Policy_{ij} + \beta_{4j}Age_{ij} + \beta_{5j}Gender_{ij} + \beta_{6j}Edu_{ij} + \varepsilon_{ij}$ (5)

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}Social\_Med_j + \gamma_{02}Polit\_Reg_j + \gamma_{03}Media\_Freedom_j + \gamma_{04}Trade_j + \gamma_{05}Geo\_Dist_j + u_{0j}$ (5.1)

$\beta_{1j} = \gamma_{10} + u_{1j}$ (5.2)

...

$\beta_{6j} = \gamma_{60}$ (5.7)

Combined: $Y_{ij} = \gamma_{00} + \gamma_{01}Social\_Med_j + \gamma_{02}Polit\_Reg_j + \gamma_{03}Media\_Freedom_j + \gamma_{04}Trade_j + \gamma_{05}Geo\_Dist_j + \gamma_{10}Polit\_Lean_{ij} + \gamma_{20}Energy\_Policy_{ij} + \gamma_{30}Rus\_Policy_{ij} + \gamma_{40}Age_{ij} + \gamma_{50}Gender_{ij} + \gamma_{60}Edu_{ij} + u_{1j}Polit\_Lean_{ij} + u_{2j}Energy\_Policy_{ij} + u_{3j}Rus\_Policy_{ij} + u_{0j} + \varepsilon_{ij}$ (5.8)

[7] As ICC approaches 0, individual-level observations are practically independent and grouping them on a higher level provides no additional information. In contrast, as ICC approaches 1, this essentially means that "all members of a group are identical" and looking at them on an individual level provides no additional information (Gelman & Hill 2006, p. 258)

5 Results

5.1 Models

The linear mixed-effects model was fitted using the maximum likelihood estimation method. Predicting variables were scaled and grand-mean centered, in accordance with suggestions from the literature (Enders & Tofighi 2007). The parameters of the fixed part of the multilevel model (e.g., γ_{00}, γ_{10}) have their base value not at 0 but at grand means (or averages) which calls for the transformation of variables so that their zero values become means across all observations. This allows for coherent interpretation of the estimates and helps with model convergence.

Striving to avoid biased estimates as a result of the omission of random effects on the one hand, and overfitting the model with the inclusion of all of the predictors on the other, I followed a common multilevel modeling strategy in which all the parameters are added sequentially, and the resulting models are compared by AIC and BIC (indicators of goodness of fit) (Glatz & Eder 2020; Hox et al. 2017). Table 2 provides the results for the complete models of each of the steps of model building discussed in the previous section (excluding the empty model, equations 2-2.2). Since multilevel models consist of both fixed and varying parts, Table 2 represents these two parts. The fixed part shows the estimates for global averages across all countries (“Intercept”, “Energy policy”, etc.). The varying part shows the variation in estimates that comes from two sources of variation: the between-country variances (“Variance between countries (Intercept)”, “Variance between countries (Energy policy)”, etc.) and the within-country variance (“Variance between individuals within countries (Intercept))”). In other words, they reflect the amount of variation in variables that were allowed to vary on the country level – τ_0^2, τ_1^2, τ_2^2, τ_3^2 in the equations above; and the variation of individuals within the countries – σ^2.

Table 2: Coefficients of Multilevel Linear Regression Analysis

	Model 1	Model 1(Full)	Model 2	Model 2(Full)	Model 3	Model 3(Full)
Intercept	4.744***	4.853***	4.808***	4.875***	4.801***	4.758***
	[4.215, 5.273]	[4.322, 5.383]	[4.341, 5.274]	[4.383, 5.368]	[4.464, 5.139]	[4.352, 5.164]
Energy Policy	0.154***	0.190***	0.132*	0.177**	0.134*	0.177**
	[0.123, 0.185]	[0.157, 0.224]	[-0.029, 0.293]	[0.012, 0.342]	[-0.027, 0.295]	[0.012, 0.342]
Russian Policy Preference (No preference)	1.232***	1.204***	1.209***	1.204***	1.216***	1.205***
	[1.126, 1.338]	[1.088, 1.319]	[0.961, 1.457]	[0.935, 1.473]	[0.985, 1.447]	[0.942, 1.468]
Russian Policy Preference (Defence)	1.403***	1.530***	1.214***	1.394***	1.223***	1.394***
	[1.293, 1.512]	[1.413, 1.647]	[0.714, 1.714]	[0.891, 1.897]	[0.719, 1.728]	[0.898, 1.891]
Age	0.274***	0.289***	0.237***	0.258***	0.236***	0.257***
	[0.243, 0.305]	[0.256, 0.322]	[0.207, 0.267]	[0.226, 0.290]	[0.206, 0.267]	[0.225, 0.290]
Female	-0.125***	-0.133***	-0.052+	-0.071*	-0.055+	-0.072*
	[-0.208, -0.041]	[-0.223, -0.043]	[-0.134, 0.029]	[-0.158, 0.017]	[-0.136, 0.027]	[-0.159, 0.016]
Edu (post-secondary)	0.532***	0.545***	0.422***	0.459***	0.423***	0.458***
	[0.405, 0.659]	[0.409, 0.681]	[0.298, 0.545]	[0.326, 0.591]	[0.299, 0.546]	[0.326, 0.590]
Edu (tertiary)	0.809***	0.831***	0.626***	0.683***	0.630***	0.683***
	[0.683, 0.935]	[0.693, 0.969]	[0.503, 0.749]	[0.548, 0.818]	[0.506, 0.753]	[0.548, 0.818]
Political Leanings (Centrist)		-0.018		-0.064		-0.063
		[-0.122, 0.086]		[-0.236, 0.109]		[-0.237, 0.112]
Social media					0.557	0.288

	Model 1	Model 1(Full)	Model 2	Model 2(Full)	Model 3	Model 3(Full)
					[-3.185, 4.299]	[-3.890, 4.466]
Media freedom					0.428+	0.571*
					[-0.203, 1.059]	[-0.179, 1.321]
Political regime					0.075	-0.050
					[-0.871, 1.022]	[-1.167, 1.067]
Trade with Russia					-2.585	-2.197
					[-30.019, 24.849]	[-35.416, 31.022]
Distance to Russia					-0.053	-0.039
					[-0.177, 0.072]	[-0.178, 0.101]
Variance between countries (Intercept)	0.990	0.863	0.753	0.733	0.365	0.394
Variance between countries (Energy policy)			0.094	0.086	0.094	0.086
Variance between countries (Rus policy (no preference))			0.189	0.194	0.159	0.183
Variance between countries (Rus policy (defence))			0.895	0.790	0.912	0.769
Variance between countries (Polit leanings (centrist))				0.063		0.065
Variance between individuals within countries (Intercept))	6.640	6.765	6.210	6.366	6.215	6.366
Num.Obs.	25939	22919	25939	22919	25939	22919
R2 Marg.	0.082	0.096	0.065	0.086	0.135	0.151
R2 Cond.	0.201	0.198	0.239	0.181	0.183	0.201
AIC	122909.4	109026.9	121324.3	107834.9	121373.4	107829.6

	Model 1	Model 1(Full)	Model 2	Model 2(Full)	Model 3	Model 3(Full)
BIC	122991.0	109115.4	121479.4	107963.6	121528.5	107998.4

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table includes three pairs of models. Each pair consists of one model that includes Political Leanings as a predicting variable and one that does not. This is because data for the Political Leanings variable is missing for three countries from the survey – Turkey, Thailand, and Saudi Arabia. Since excluding these countries from analysis would have resulted in the loss of 3020 datapoints, I decided to run two separate models instead. My decision bears important considerations for the interpretation of the results. In particular, the results of the most complete model ('M3 Full') are only applicable to 22 out of the 25 countries on my list.

Aside from demonstrating regression coefficients, their confidence intervals, and *p*-values, Table 1 allows for the comparison of the explained variance across the models by looking at the marginal R^2 (explained variance by fixed factors) and conditional R^2 (explained variance by both fixed and random effects) (Glatz & Eder 2020). For the unexplained variance, it can be noted that the variance between individuals within countries decreases from 6.765 (Model 1(Full)) to 6.366 (Model 3(Full)), while the variance between countries decreases from 0.863 (Model 1(Full)) to 0.394 (Model 3(Full)).

In Table 2, Intercept reflects the grand mean average of disinformation resilience for all respondents across all countries, when all the predicting variables are held constant at their mean values. In other words, the predicted average resilience score is 4.758 when all the predictors are at their means (M3(Full)). Turning to the slopes, the results from the models show some notable statistically significant correlations between the outcome variable and individual and country-level predictors. The average effect of individual-level policy preferences for higher energy costs across all countries is positively associated with the resilience to disinformation in all the models, and its magnitude is bigger in models that include the Political Leanings variable (Model 1(Full) – Model 3(Full)). In other words, every one-unit increase in energy policy preference is associated with a 0.177-point increase in resilience score, on average (M3(Full)). Accounting for these considerations, I find statistical support for *H6*.

Compared to energy policy preferences, perceiving Russia as a threat shows stronger statistical correlation with resilience ($p < .001$ in all models) on all levels of this factor variable. As such, not having strong policy preferences toward Russia (in contrast to preferences of diplomatic relationships) correlates positively with higher resilience to disinformation, increasing the global resilience score by 1.205 points on average. The magnitude of this correlation increases for preferences of defence against Russian aggression and amounts to a 1.394-point increase on average compared to preferences for diplomatic relationships (Model 3(Full)). The effect persists across all the models (including the alternate estimate, Table 5 in the Appendix), providing strong support for *H5*.

On the country level, media freedom shows a statistically significant positive correlation with disinformation resilience, although the relationship is only significant at $p < .1$ (M3) and $p < .05$ (M3(Full)). However, when models are estimated without the Political regime variable (Table 5 in the Appendix), the correlation becomes statistically significant at $p < .001$ (M3) and $p < .01$ (M3(Full)). With these considerations in mind, I find statistical support for *H3* with some important stipulations that I will discuss below.

Curiously enough, the Political Leanings variable shows no statistically significant correlation with resilience. Moreover, in contrast to the expectations of the hypothesis, the direction of the correlation is negative, suggesting that respondents with centrist political values are less resilient toward Russian disinformation. As such, I fail to reject the null hypothesis for *H1*. Potential explanations for this result will be discussed further. At the same time, when looking at extreme political leanings in general (i.e., combining far-right and -left), the model shows statistically significant negative correlation with resilience, supporting the idea that irrespective of the direction of their leanings, people with extreme views are more prone to believe disinformation (Table 6 in the Appendix). Moreover, the negative effect is slightly stronger for left political views as compared to right. The comparison of models with the extreme leanings variable and far-left and -right variables is shown in Table 6 in the Appendix.

It is interesting to note that the country-level variables of social media, political regime (including the alternative measurement of V-Dem score, see Table 8 in the Appendix), trade with Russia, and distance to Russia did not show significant statistical correlation with disinformation resilience including the alternate estimates of the models, thus resulting in the failure to reject the null hypotheses for *H2, H4, H7,* and *H8*.

To provide a more visual interpretation of the results, I present marginal effects plots for the three factors that showed statistically significant correlation with the outcome variable (Figures 1-3). Here, the global average effects (across all the countries) of each of the independent variables are plotted with the *prediction* intervals that account for between-country variation. In other words, the shaded area here includes the average effect for all 25 countries while accounting for the variation between these countries. It could be noted that predicting intervals are slightly wider for Model 3(Full) than for Model 3, because the former includes the additional varying variable of Political Leanings which increases uncertainty. I also provide marginal effects with more conventional *confidence* intervals that surround the global average but do not account for between-country variation and are hence narrower due to lower uncertainty (see Figures 12-14 in the Appendix).

Figure 1: Marginal Effects for Energy Policy Preferences with Predicting Intervals

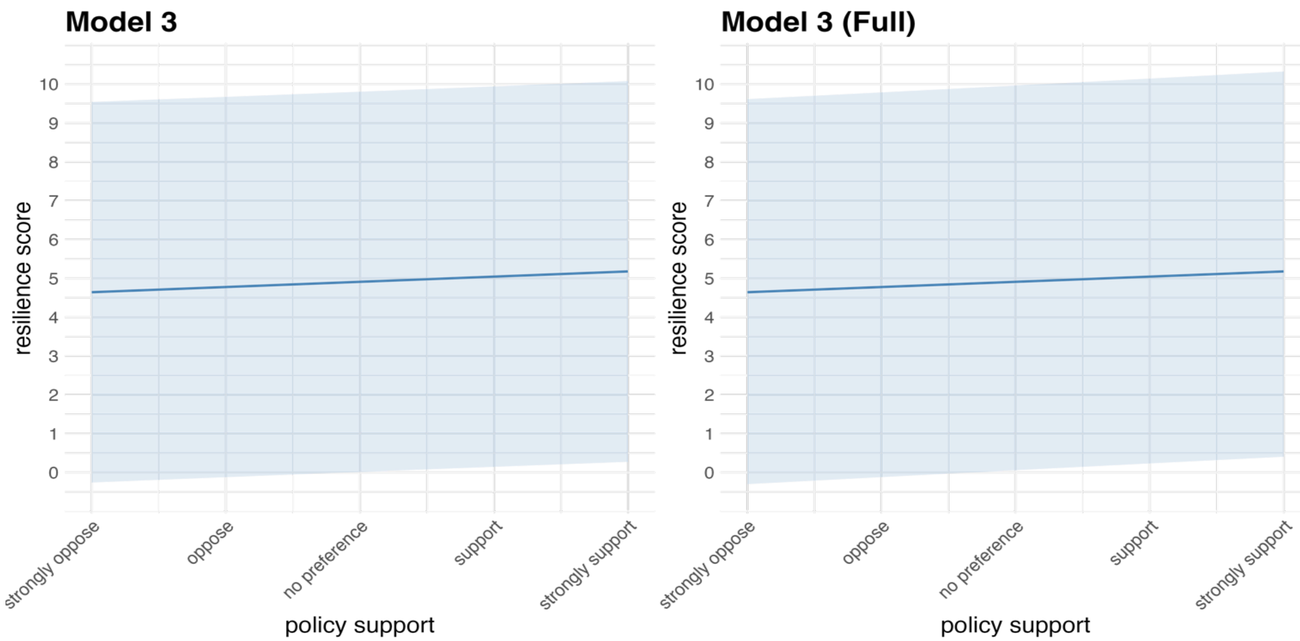


Figure 2: Marginal Effects for Attitudes Toward Russia with Predicting Intervals

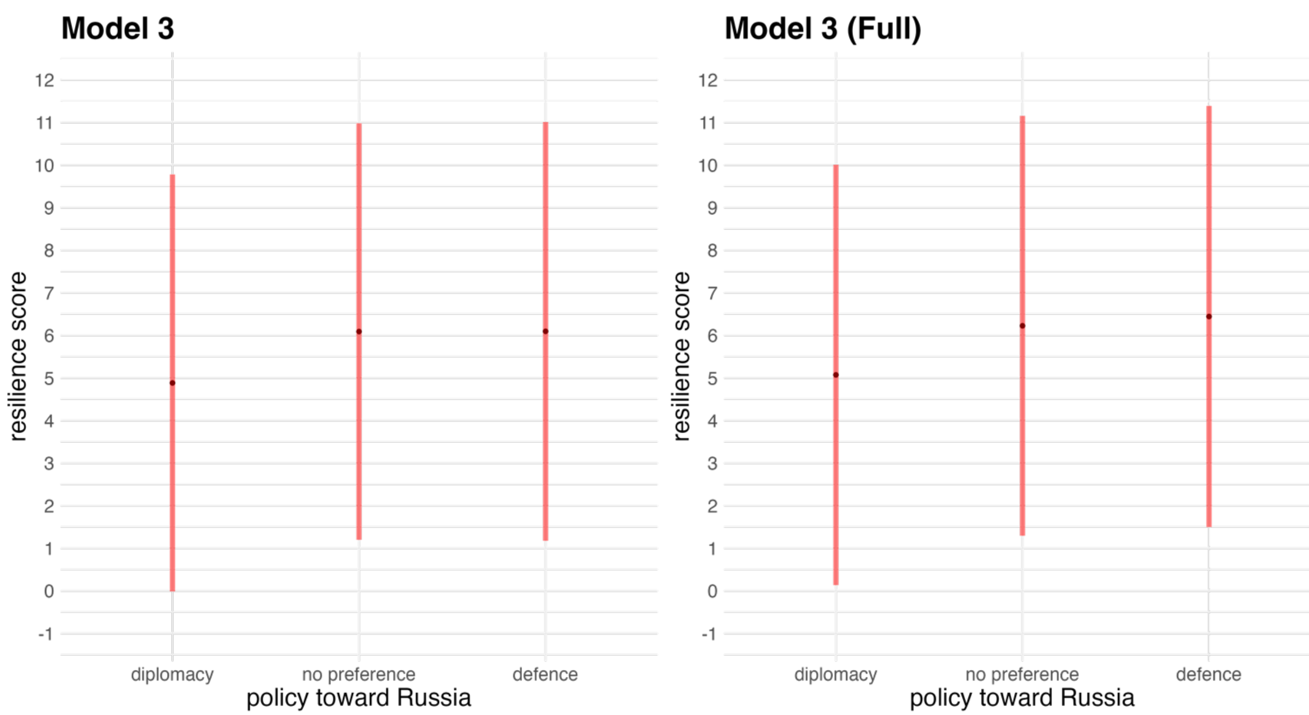
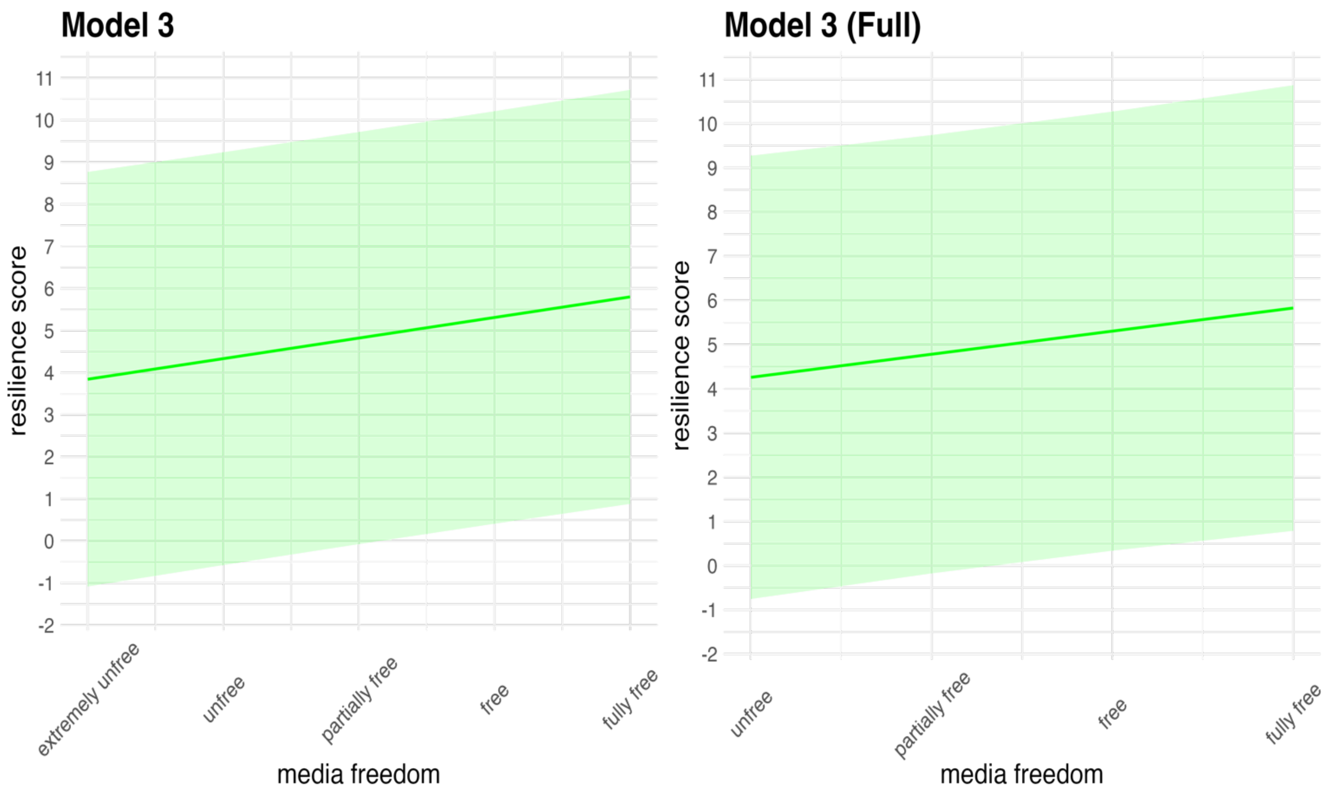


Figure 3: Marginal Effects for Media Freedom with Predicting Intervals



5.2 Robust Standard Errors

As was discussed previously, the violation of the normality assumption affects the precision of the parameters' estimation. While the coefficients remain unbiased, the asymptotic standard errors estimated with maximum likelihood are downward biased, making confidence intervals artificially smaller. This issue diminishes as the number of observations on both levels increases, and Maas and Hox (2004) suggest that more than 50 observations on the group level are required to consider this violation negligible. However, since my study includes only 25 countries, I must employ the Huber/White (or sandwich) estimator to produce robust standard errors. The results of robust error estimations can be found in Table 7 in the Appendix. Although this estimation method does change the confidence intervals, making them more precise and consistent, it does not provide any substantial changes in the values of the parameters.

Importantly, the sandwich estimator is only one of the proposed solutions for the violation of the assumption and is most often used as a diagnostic method aimed at detecting of some modelling imprecisions. More robust estimation methods, such as the bootstrap method or a Bayesian approach, would ideally address this issue in a more rigorous way, but are, unfortunately, beyond the scope of the present research.

6 Discussion and Limitations

The results above provide a few interesting findings worthy of discussion. I will, therefore, discuss the theoretical and practical implications of the hypotheses that found empirical support as well as those that did not. I will then point out some of the key limitations of the research.

6.1 Hypotheses

Given the solid theoretical background of this work, it is informative to compare the results of this research with the results of previous works on disinformation resilience discussed above. To start, both Humprecht et al. (2020, 2021) and Boulianne et al. (2022) emphasized the role of media in resilience to disinformation. Applying this factor to a broader sample of political and media regimes, I found a statistically significant association between higher disinformation resilience and media freedom, suggesting that respondents living in more free media systems are on average more resilient to disinformation compared to those who live in more restricted media systems. However, this finding implies an important reservation—only two countries in my sample, Saudi Arabia and Egypt, are in the lowest level of the media freedom index. The fact that these two countries are regional neighbors and share some political and cultural similarities invokes the issue of a spurious relationship. Future research could address this issue by looking at a more diverse sample of countries across levels of media freedom.

The next two factors are original to this work, although they too are based on broader theoretical foundations and therefore have some broader theoretical significance. As my research sought to analyze the specific case of Russian disinformation in the context of the ongoing war in Ukraine, I hypothesized whether and how people's attitudes toward Russia might be correlated with their resilience towards pro-Russian disinformation. I found that respondents' toleration of higher energy prices as a result of sanctions on the Russian energy sector is associated with higher resilience to disinformation. This finding is noteworthy, as it provides potential avenues for further research on how disinformation resilience is affected by the personal costs of accepting or not accepting this disinformation. In the same vein, I discovered that perceptions of Russia as a threat, expressed through respondents' support of building up defence against Russian aggression, predict higher resilience to disinformation compared to more friendly or undecided attitudes toward the country. These two findings could potentially enrich the growing body of work describing one's propensity to believe in the disinformation narratives of a disinformation disseminating state, depending on one's links and attitudes towards the state itself.

Given that refuting particular hypotheses can have value in and of itself, I now turn to a discussion of the hypotheses that did not find support in this work. First, I was not able to find a significant correlation between higher social media use and lower resilience. This could partly be explained by the nature of the data used to study this hypothesis. As information on social media use is not readily available, I had to estimate my own datapoints by combining various sources, which in the end could have affected the robustness of estimations. Given that social media use and propensity to believe in disinformation are widely studied and proved to be correlated, it would be beneficial to revisit the results of this work with more reliable data.

Although I was not able to arrive at a statistically significant correlation between centrist political values and disinformation resilience, I found negative statistical correlation between both far-right and -left political leanings (as well as extreme leanings in general) and resilience to disinformation, supporting some of the previous findings. While my decision to take an alternative approach and look at the centrist views is potentially informative, isolating centrism from other political leanings introduces uncertainty related to each respondent's understanding of centrism itself, as well as the propensity to relate *oneself* to more moderate political beliefs than one might actually uphold. In other words, varying understandings of centrism could easily allow for the inclusion of respondents with vastly different values, and notably, values that could also be shared by individuals with other political leanings. In terms of future research, I suggest that centrism either be identified with more precise and perhaps more objective measurements (such as voting behaviour), or be replaced with a detailed scale of political values instead of the two broad political groupings that a dummy variable presupposes.

To my surprise, neither trade with Russia, nor distance to it played a significant role in predicting disinformation resilience in the context of this study. While the weakness of the correlation between the Trade variable and disinformation resilience is more likely rooted in weaker theoretical foundations, the non-finding for the Distance variable might be caused by data imprecision related to the fact that capitals' distances do not always reflect the true distance between the countries. As was mentioned previously, future studies may benefit from using alternative approaches to operationalize distance, such as borders' distance or presence of a common border or cultural distance (Erlich & Garner 2023). Importantly, these findings might still bear important practical considerations in suggesting that neither distance nor economic relationships with Russia can predetermine people's disinformation resilience.

Finally, the political regime factor was flagged as a potentially problematic factor due to its skewed distribution (most of the countries are free or partially free) and high correlation with the factor of media freedom. It is of little surprise, then, that this factor did not show statistical significance. Nonetheless, the susceptibility of illiberal political regimes to disinformation, their mutual support and amplification of disinformation narratives, as well as the growing body of studies on informational autocracies discussed above, continue to suggest the relevance of this hypothesis and thereby encourages further studies with a more diverse sample of countries.

6.2 Limitations

The previous section alluded to some of the limitations of this work upon which I will now elaborate. While the quality and availability of data is a fundamental part of any valuable research project, this study was conducted under particular conditions of limited data availability (e.g., data on social media use). Addressing this issue, I opted to run two separate types of models, one of which excluded some important observations, implying certain limitations on the interpretation of the results such that they tend to exclude countries with illiberal political regimes from their application.

Additionally, some of the underlying assumptions of linear multilevel modeling were not met. In particular, some of the group-level residuals were not normally distributed, jeopardizing the precision of standard error estimates and the consequently the reliability of confidence intervals. While I attempted to address this issue with robust standard errors, more rigorous techniques, such as a Bayesian approach or bootstrap robust errors, were not implemented due to the limited methodological scope of the research. I also mentioned the violation of the linearity assumption due to the nature of the outcome variable. While the methods available to address this violation (Poisson and ordinal models) are beyond the scope of this research, I accept the results for cautious exploratory analysis, where predictive power of the model is not assumed.

This brings me to the major limitation of this work, which concerns its inability to claim any causal relationships and its restriction to preliminary findings. Disinformation resilience is a phenomenon that proves tricky to quantify and the strength of identified resilience factors is often constrained by available data. With this in mind and given the available data and the model at hand, this research seeks to provide cautious initial evidence for the existence of some correlations between the predicting and outcome variables. It cannot indicate the presence of causal relationships, and it cannot provide information about the direct effect of the factors on disinformation resilience.

Therefore, this analysis should be considered as a first step of identifying potential directions for future research on disinformation resilience frameworks that would likely benefit from incorporating both individual and country-level factors, as well as the factor of deliberate information attacks by foreign states.

7 Conclusions

In constructing the framework of disinformation resilience, Humprecht et al. (2020, 2021) and other scholars identify a few factors of resilience and provide support for their claims in empirical data. In this thesis, I sought to complement existing research by proposing some amendments to the prevailing resilience framework. In doing so, I aspired to enhance its analytical power by: (1) extending the current framework to test a case of the ongoing Russian war in Ukraine, (2) considering additional predictor variables that are related to the country that disseminates disinformation narratives, and (3) applying the model to a larger subset of countries. My research was based on the fundamental argument that both individual-level characteristics and country-level specificities are important factors of resilience to disinformation; and my goal was to investigate whether and under which circumstances deliberate informational attacks by other states might correlate with lower or higher resilience to disinformation.

While not all the hypotheses were supported, and despite certain model and data limitations, I presented interesting findings that offer several avenues for future work on disinformation resilience. In particular, I generated evidence that individual-level characteristics, such as perceptions of Russia as a threat and support of increased energy prices, are correlated with resilience to disinformation. As such, I provided support to my argument that aside from the established factors of disinformation resilience, it is important to consider attitudes about the source of disinformation (in the context of this work—the Russian state). I also confirmed the relationship between media and resilience, showing that free media systems predict higher resilience.

While I acknowledge the limitations of my study, the results can still be used to identify ideas for future research. For example, it would be valuable to check whether the relationship between the individual-level factors and resilience is moderated by the country-level factors, i.e., to look at the interaction terms between individual- and aggregate-level factors. Moreover, it would be interesting to apply this framework to a different disinformation disseminating state, such as China, to see if the results persist. Finally, these findings might have some practical value for policy implementation aimed at addressing the issue of deliberate disinformation attacks by foreign states.

As the deliberate disinformation attacks become a more pressing issue and governments around the world mobilize resources to find comprehensive solutions (Wintour 2024), it is important to have a critical discussion of whether a universal framework is feasible. The past works on the resilience framework suggest that a variety of factors determine the propensity to belief disinformation (many of which are country specific), and this study pushes this argument further by considering the source of disinformation. This discussion would benefit from further research and broader sample of countries. Finally, future research should move from identifying correlational links between the proposed factors to explaining potential causal mechanisms.

Appendix

Table 3: Variables

Variable	Measurement	Original Code	Hypo-sis	Source
Political Leanings	1- centrists; 0-all the rest	*"Some people talk about 'left', 'right', and 'center' to describe parties and politicians. With this in mind, where would you place yourself on this scale?"*:[1] Don't know, [2] Very left-wing, [3] Fairly right-wing, [4] Fairly left-wing, [5] Centre, [6] Slightly right-of-centre, [7] Slightly left-of-centre, [8] Very right-wing,[9] Prefer not to say	*H1*	EUI-YouGov survey
Energy Policy Preference	-2-strongly oppose; -1-tend to oppose; 0 - Don't know; 1-tend to support; 2-strongly support	*"EU/Market accepting higher energy costs due to the sanctions imposed on Russia"*: [1] Don't know,[2] Strongly support,[3] Tend to support,[4] Tend to oppose,[5] Strongly oppose	*H6*	EUI-YouGov survey
Russia Policy Preference	-1- invest more in trade and diplomacy; 0-don't know/neither of this; 1-invest more in defence and security	*"Thinking about market's/EU's relationships with Russia, which approach would you prefer to take?"*: [1] Don't know,[2] Neither of these,[3] should invest more in defence and security to defend against Russian aggression,[4] should invest more in trade and diplomacy with Russia to improve relations	*H5*	EUI-YouGov survey
Social Media	proportion of the users of the most popular social media out of the whole population of a country Alternative: proportion of Facebook users		*H2*	Digital News Report, Statista Alternative: World Population Review
Media Freedom	1-extremely unfree; 2-unfree; 3-partially free; 4-free; 5-fully free;	*"The level of freedom enjoyed by journalists and media in 180 countries and territories":85 - 100 points [good];70 - 85 points [satisfactory];55 - 70 points [problematic]; 40 - 55 points [difficult]; 0 - 40 points [very serious]*	*H3*	World Press Freedom Index
Political Regime	1-unfree; 2-partly free; 3-free;	1(Not Free); 2(Partially Free); 3(Free)	*H4*	Freedom House

Trade with Russia	proportion of imports/exports from/to Russia out of total exports/imports		*H7*	Market Potential Index (MSU)
Geographic Distance	capitals' distance in thousands km		*H8*	CPII
Control Variables				
Age	-2 – 18-24; -1 – 25-34; 0 – 35-44; 1 – 45-54; 2 – 55+	18-24; 25-34; 35-44; 45-54; 55+		EUI-YouGov survey
Gender	1-Female; 0-Male	Male; Female		EUI-YouGov survey
Education	1-less then primary, lower secondary; 2-upper secondary and post-secondary; 3-tertiary	less then primary, lower secondary: “yes”, “no”; upper secondary and post-secondary: “yes”, “no”; tertiary: “yes”, “no”		EUI-YouGov survey
Outcome Variable				
Resilience Index	Scale for each question: 0-definitely true; 1-probably true; 2-either true or false; 3-probably false ; 4-definitely false; Aggregated for three questions: 0-2-definitely true(no resilience); 3-5-probably true (weak resilience); 6-8-either true or false (some resilience); 9-11-probably false (strong resilience); 12-definitely false (full resilience)	***1) "Before the war started, ethnic Russians living in Ukraine were being subjected to mass murder – or 'genocide' – by Ukrainians"****: [1] Don't know either way – this may be true or may be false,[2] Probably false,[3] Probably true,[4] Definitely false,[5] Definitely true,* ***2)"Before the war started, the Ukrainian Government had fallen under the influence of militant extremists who support the ideology of Nazism and Adolf Hitler"****:[1] Don't know either way – this may be true or may be false,[2] Probably false,[3] Probably true,[4] Definitely false,[5] Definitely true, 3)* ***"Before the war started, Western countries were seeking to establish a military infrastructure in Ukraine in order to bully and threaten Russia"*** : [1] Don't know either way – this may be true or may be false,[2] Probably false,[3] Probably true,[4] Definitely false,[5] Definitely true. Sum of responses to the three questions	All	EUI-YouGov survey

Table 4: Social Media Use by Country

Country	Top Social Media	Proportion of Users out of Total Population
UK	WhatsApp	0.627
Australia	Facebook	0.5628
Brazil	WhatsApp	0.585
Canada	Facebook	0.6392
US	Facebook	0.522
Denmark	Facebook	0.7056
Egypt	Facebook	0.6715
France	Facebook	0.5612
Germany	WhatsApp	0.6528
Greece	Facebook	0.5548
Hungary	Facebook	0.7209
India	WhatsApp	0.4104
Indonesia	WhatsApp	0.6468
Italy	WhatsApp	0.7533
Japan	Line	0.3895
Kenya	WhatsApp	0.748
Mexico	Facebook	0.5293
Nigeria	WhatsApp	0.6059
Poland	Facebook	0.546
Saudi Arabia	WhatsApp	0.8232
South Africa	WhatsApp	0.4524
Spain	WhatsApp	0.7812
Sweden	Facebook	0.6912
Thailand	Facebook	0.672
Turkey	WhatsApp	0.5976

Figure 4: Trade share with Russia

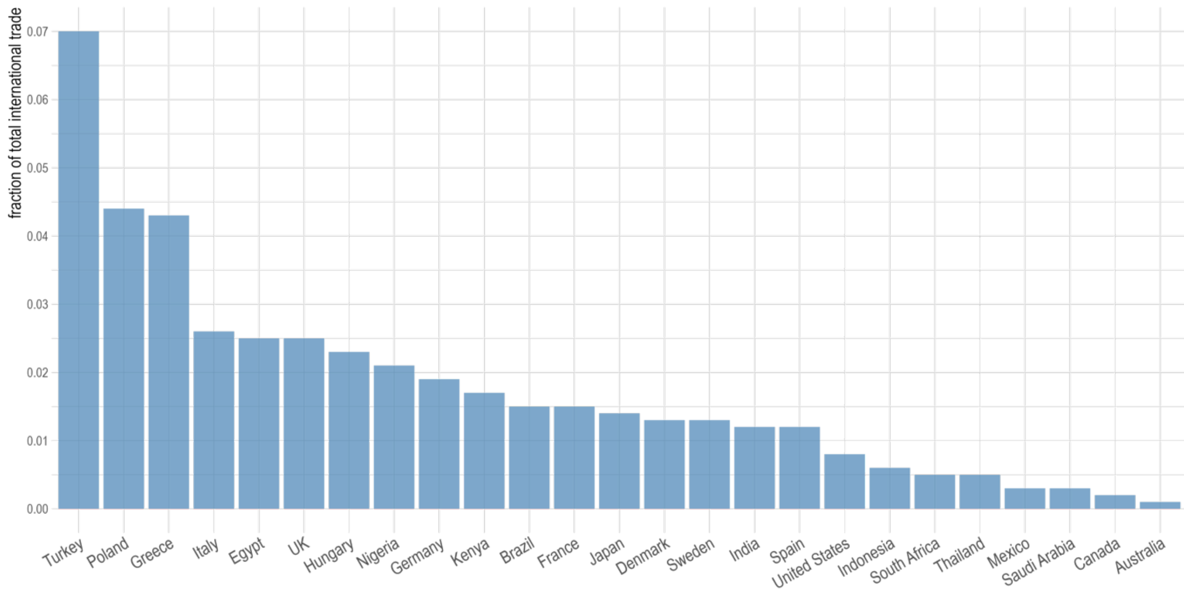


Figure 5: Disinformation Resilience Distribution by Country

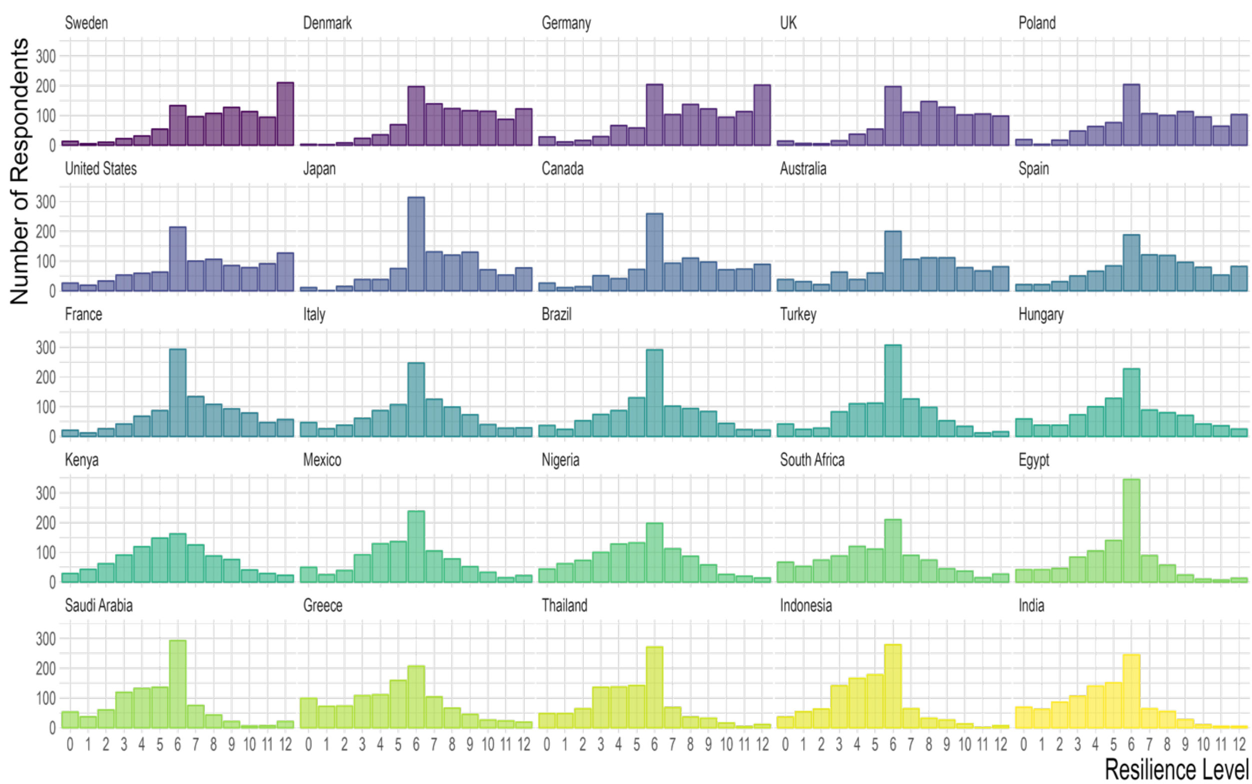


Figure 6: Average Resilience Score by Country

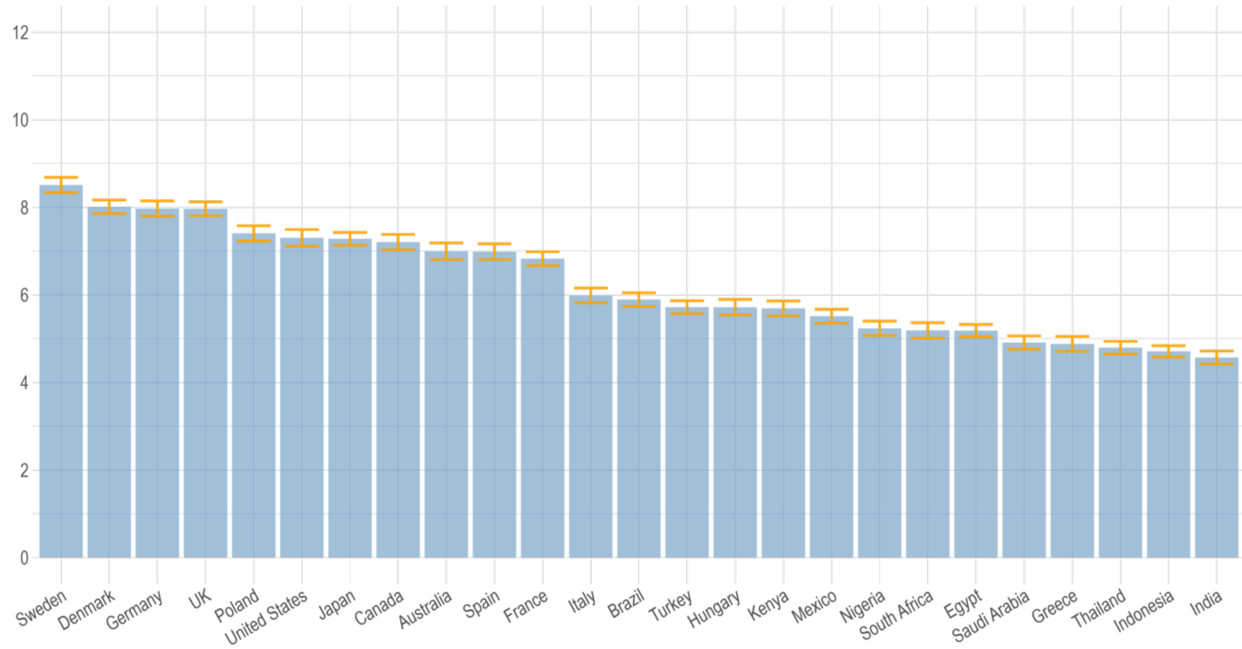


Figure 7: QQ Plots for M3 and M3(Full)

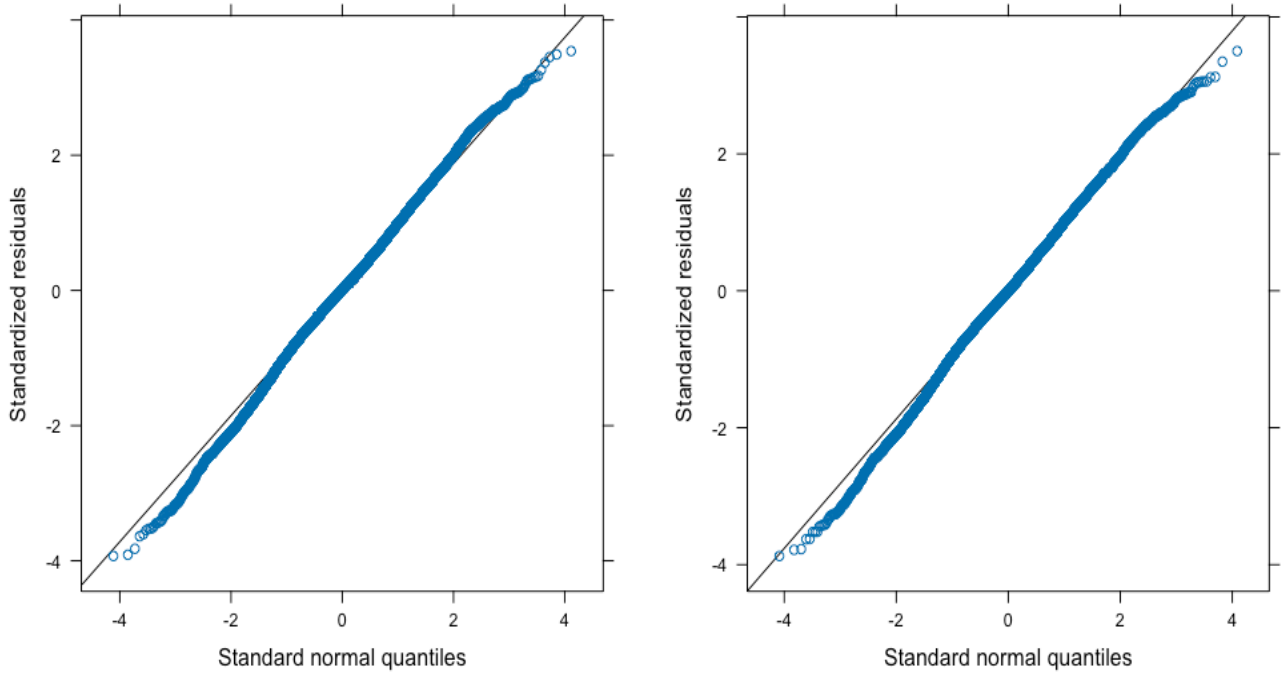


Figure 8: Distribution of Residuals, M3

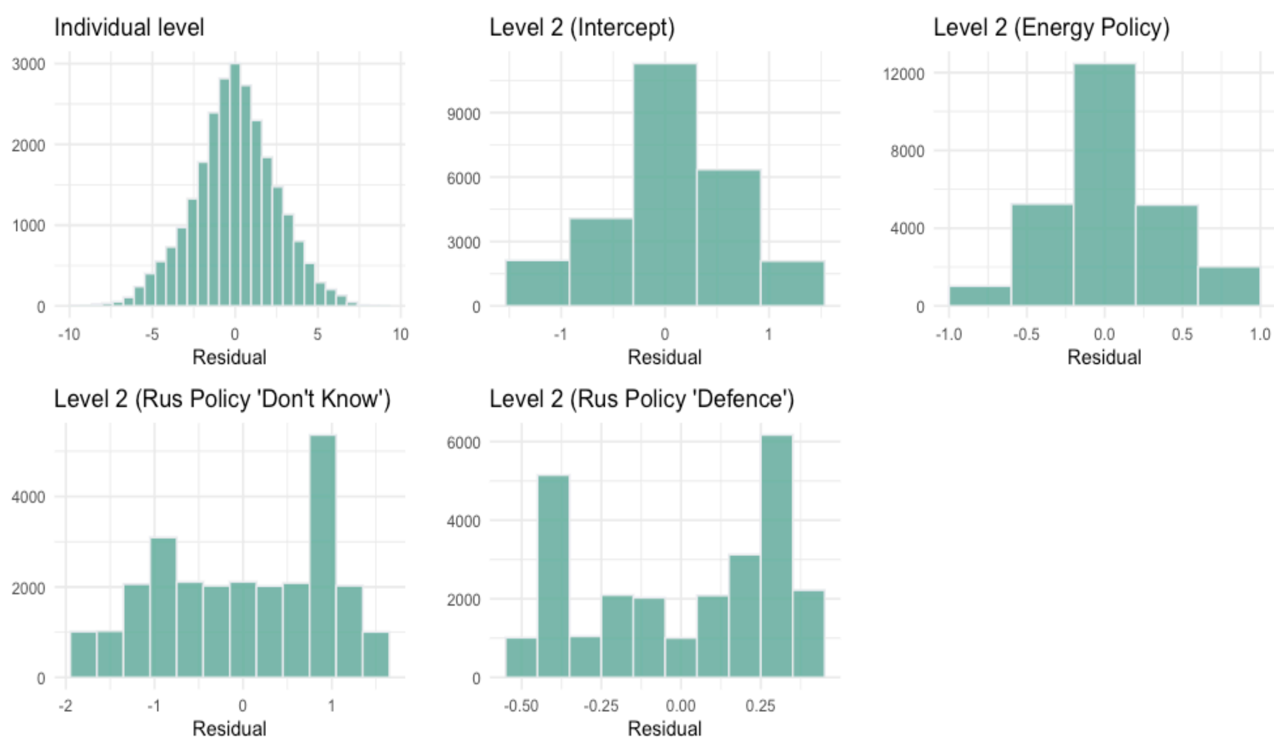


Figure 9: Distribution of Residuals, M3(Full)

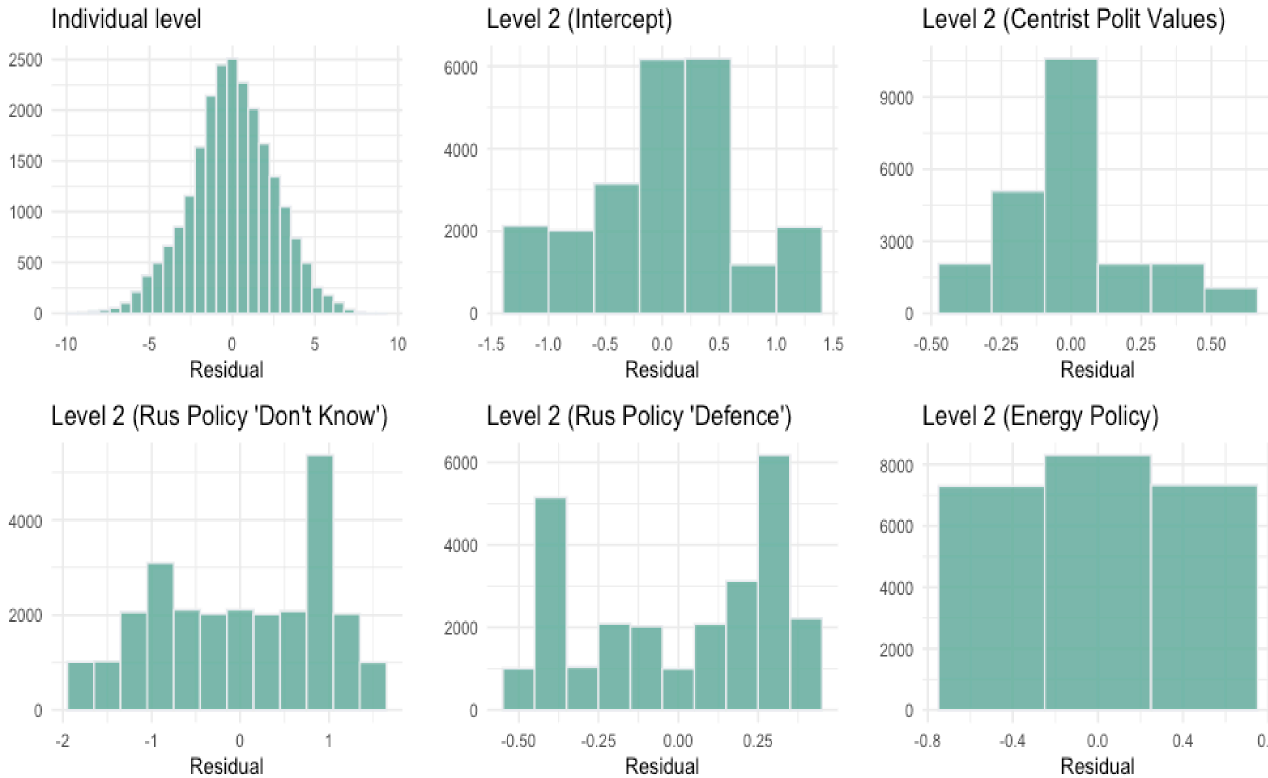
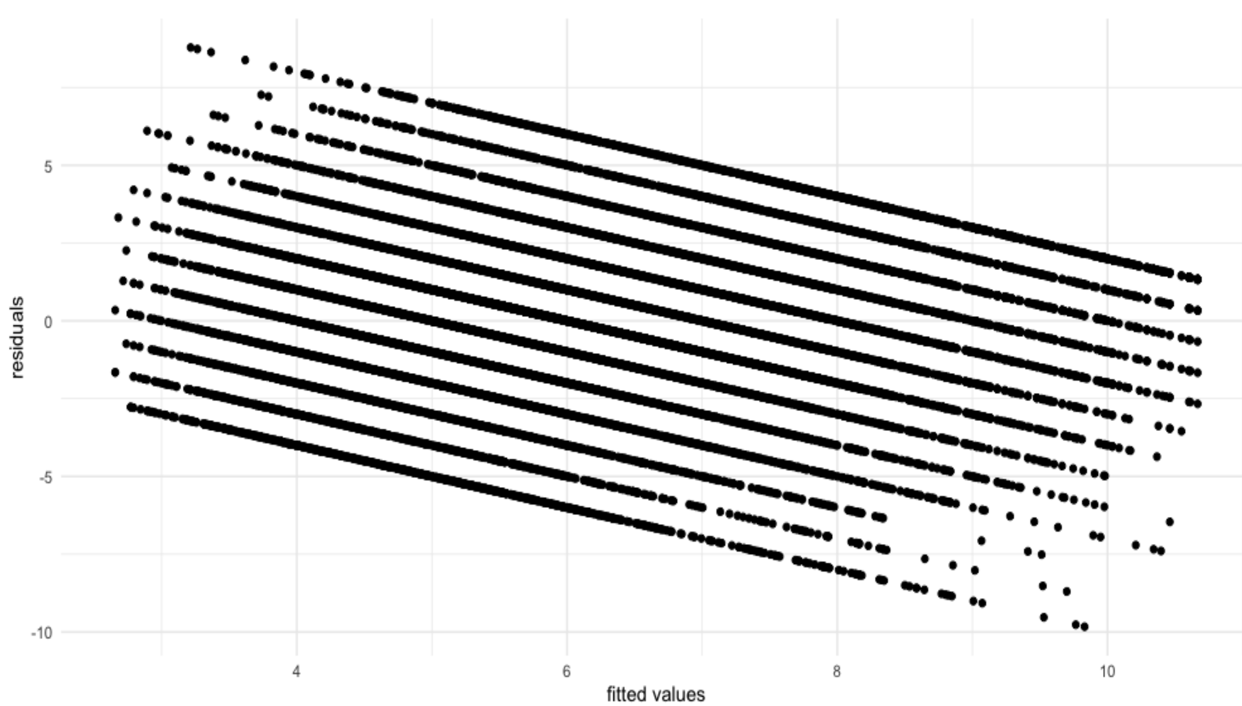


Figure 10: Residuals Vs Fitted Values



Intermediate Steps of Multilevel Model Building

In the second step, the unconditional means model is sequentially augmented with the individual-level predictors of political leanings, energy policy preferences and Russian policy preferences, and the control variables of age, gender and education, while the intercept is allowed to vary by country:

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}Polit\_Lean_{ij} + \beta_{2j}Energy\_Policy_{ij} + \beta_{3j}Rus\_Policy_{ij} + \beta_{4j}Age_{ij} + \beta_{5j}Gender_{ij} + \beta_{6j}Edu_{ij} + \varepsilon_{ij}$ (3)

Level 2: $\beta_{0j} = \gamma_{00} + u_{0j}$ (3.1)
$\beta_{1j} = \gamma_{10}$ (3.2)
...
$\beta_{6j} = \gamma_{60}$ (3.7)

Combined: $Y_{ij} = \gamma_{00} + \gamma_{10}Polit\_Lean_{ij} + \gamma_{20}Energy\_Policy_{ij} + \gamma_{30}Rus\_Policy_{ij} + \gamma_{40}Age_{ij} + \gamma_{50}Gender_{ij} + \gamma_{61}Edu_{ij} + \varepsilon_{ij} + u_{0j}$ (3.8)

In the above equations, γ_{00} is the mean disinformation resilience across all countries, controlling for all predictors. Var (ε_{ij}) = σ^2 is the individual respondents' differences around the mean of disinformation resilience, controlling for all predictors. Variance (u_{0j}) = τ_0^2 captures the variance of countries around the grand mean controlling for all predictors. The average effect of political identity (Polit_Lean) on disinformation resilience across all countries is γ_{10}; the average effect of energy policy preference (Energy_Policy) is γ_{20}, the average effect of Russian policy preference (Rus_Policy) is γ_{30}, and so on. Here, the intercept is randomized and the slope coefficients are fixed; in other words, this model accounts for a difference in the mean levels of disinformation resilience between the countries while assuming that the effects of individual-level predictors are the same across all the countries. Substantively, this random intercept model allows for the comparison between the group means while controlling for the individual-level factors within groups, i.e., "means-as-outcomes model" (Raudenbush & Bryk 2002, pp. 24-25).

In the next stage (4 – 4.8), I build a model with the random intercepts (β_{0j}) and random slopes ($\beta_{1j} - \beta_{3j}$), where every country has its own intercept and its own regression slopes. In identifying which slopes should be randomized, I consult not only theory (Barr et al. 2013) but also visualized relationships between disinformation resilience and energy policy, Russian policy preferences, and political leanings variables (see Figure 11 in the Appendix). Given the different trends identified in Figure 11 (Appendix), I opted for the use of random-effect slopes for these variables. As such, this model assumes that the average disinformation resilience level, as well as the relationships between disinformation resilience and political leanings, Russian and energy policy preferences vary across countries:

Level 1: $Y_{ij} = \beta_{0j} + \beta_{1j}Polit\_Lean_{ij} + \beta_{2j}Energy\_Policy_{ij} + \beta_{3j}Rus\_Policy_{ij} + \beta_{4j}Age_{ij} + \beta_{5j}Gender_{ij} + \beta_{6j}Edu_{ij} + \varepsilon_{ij}$ (4)

Level 2:

$\beta_{0j} = \gamma_{00} + u_{0j}$ (4.1)

$\beta_{1j} = \gamma_{10} + u_{1j}$ (4.2)

$\beta_{2j} = \gamma_{20} + u_{2j}$ (4.3)

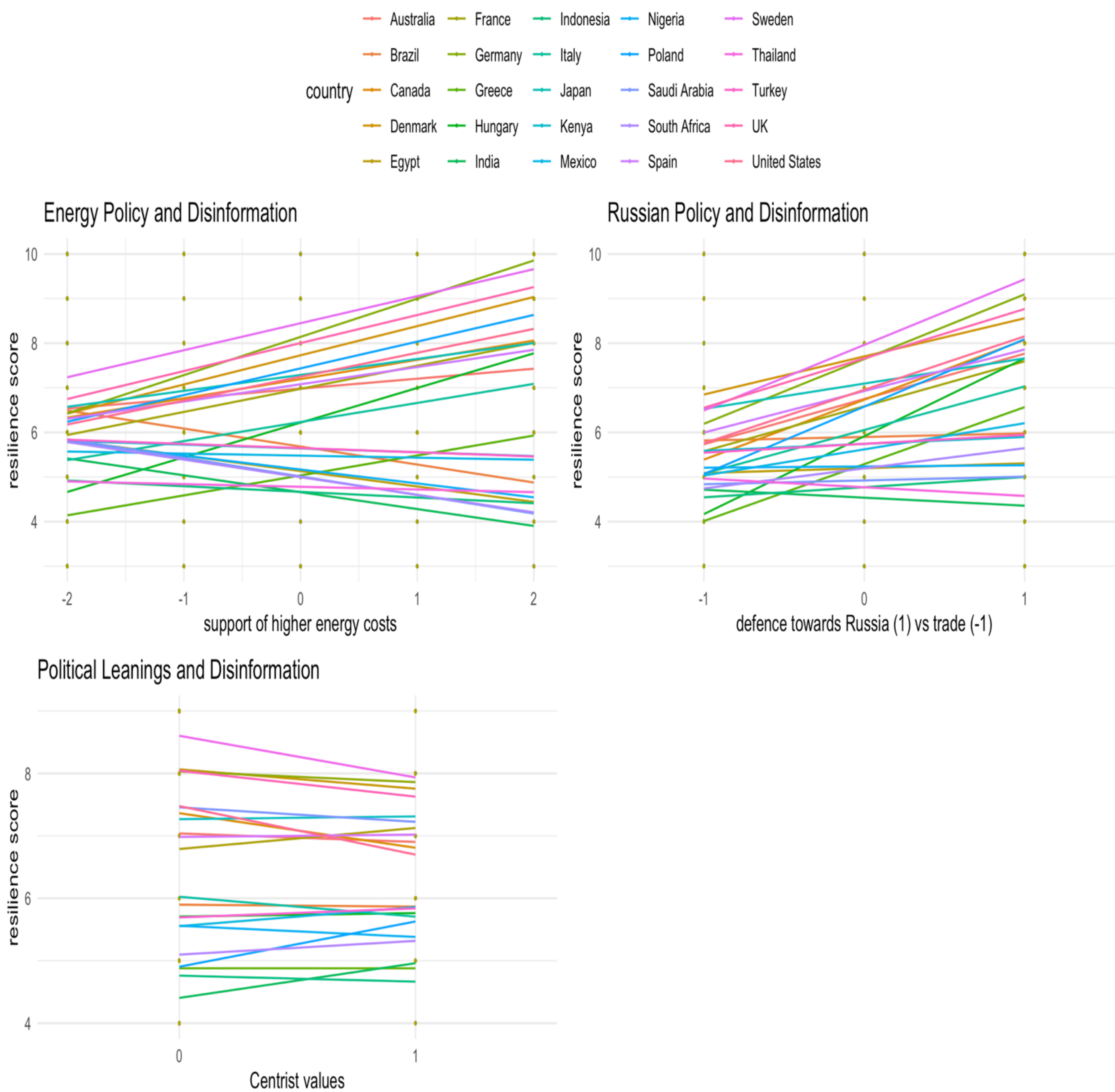
$\beta_{3j} = \gamma_{30} + u_{3j}$ (4.4)

…

$\beta_{6j} = \gamma_{60}$ (4.7)

Combined: $Y_{ij} = \gamma_{00} + \gamma_{10}Polit\_Lean_{ij} + \gamma_{20}Energy\_Policy_{ij} + \gamma_{30}Rus\_Policy_{ij} + \gamma_{40}Age_{ij} + \gamma_{50}Gender_{ij} + \gamma_{61}Edu_{ij} + u_{1j}Polit\_Lean_{ij} + u_{2j}Energy\_Policy_{ij} + u_{3j}Rus\_Policy_{ij} + u_{0j} + \varepsilon_{ij}$ (4.8)

Figure 11: Disinformation Resilience Distribution and Level-1 Predicting Variables



Alternate (Final) Model Estimation (Excluding Political Regime Variable)

$$\text{Level 1: } Y_{ij} = \beta_{0j} + \beta_{1j}Polit\_Lean_{ij} + \beta_{2j}Energy\_Policy_{ij} + \beta_{3j}Rus\_Policy_{ij} + \beta_{4j}Age_{ij} + \beta_{5j}Gender_{ij} + \beta_{6j}Edu_{ij} + \varepsilon_{ij} \text{ (6)}$$

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}Social\_Med_j + \gamma_{02}Media\_Freedom_j + \gamma_{03}Trade_j + \gamma_{04}Geo\_Dist_j + u_{0j}$ (6.1)

$$\beta_{1j} = \gamma_{10} + u_{1j} \text{ (6.2)}$$

$$\ldots$$

$$\beta_{6j} = \gamma_{60} \text{ (6.7)}$$

Combined: $Y_{ij} = \gamma_{00} + \gamma_{01}Social\_Med_j + \gamma_{02}Media\_Freedom_j + \gamma_{03}Trade_j + \gamma_{04}Geo\_Dist_j + \gamma_{10}Polit\_Lean_{ij} + \gamma_{20}Energy\_Policy_{ij} + \gamma_{30}Rus\_Policy_{ij} + \gamma_{40}Age_{ij} + \gamma_{50}Gender_{ij} + \gamma_{60}Edu_{ij} + u_{1j}Polit\_Lean_{ij} + u_{2j}Energy\_Policy_{ij} + u_{3j}Rus\_Policy_{ij} + u_{0j} + \varepsilon_{ij}$ (6.8)

Table 5: Coefficients of Multilevel Linear Regression Analysis (without Political Regime variable)

	Model 1	Model 1(Full)	Model 2	Model 2(Full)	Model 3	Model 3(Full)
Intercept	4.744***	4.853***	4.808***	4.875***	4.801***	4.756***
	[4.215, 5.273]	[4.322, 5.383]	[4.341, 5.274]	[4.383, 5.368]	[4.464, 5.139]	[4.352, 5.159]
Energy Policy	0.154***	0.190***	0.132*	0.177**	0.134*	0.177**
	[0.123, 0.185]	[0.157, 0.224]	[-0.029, 0.293]	[0.012, 0.342]	[-0.027, 0.295]	[0.012, 0.342]
Russian Policy Preference (No preference)	1.232***	1.204***	1.209***	1.204***	1.216***	1.205***
	[1.126, 1.338]	[1.088, 1.319]	[0.961, 1.457]	[0.935, 1.473]	[0.985, 1.447]	[0.942, 1.468]
Russian Policy Preference (Defence)	1.403***	1.530***	1.214***	1.394***	1.223***	1.394***
	[1.293, 1.512]	[1.413, 1.647]	[0.714, 1.714]	[0.891, 1.897]	[0.719, 1.728]	[0.898, 1.890]
Age	0.274***	0.289***	0.237***	0.258***	0.236***	0.257***
	[0.243, 0.305]	[0.256, 0.322]	[0.207, 0.267]	[0.226, 0.290]	[0.206, 0.267]	[0.225, 0.290]
Female	-0.125***	-0.133***	-0.052+	-0.071*	-0.055+	-0.072*
	[-0.208, -0.041]	[-0.223, -0.043]	[-0.134, 0.029]	[-0.158, 0.017]	[-0.136, 0.027]	[-0.159, 0.016]

	Model 1	Model 1(Full)	Model 2	Model 2(Full)	Model 3	Model 3(Full)
Edu (post-secondary)	0.532***	0.545***	0.422***	0.459***	0.423***	0.458***
	[0.405, 0.659]	[0.409, 0.681]	[0.298, 0.545]	[0.326, 0.591]	[0.299, 0.546]	[0.326, 0.590]
Edu (tertiary)	0.809***	0.831***	0.626***	0.683***	0.629***	0.683***
	[0.683, 0.935]	[0.693, 0.969]	[0.503, 0.749]	[0.548, 0.818]	[0.506, 0.753]	[0.548, 0.818]
Political Leanings (Centrist)		-0.018		-0.064		-0.063
		[-0.122, 0.086]		[-0.236, 0.109]		[-0.237, 0.112]
Social media					0.425	0.364
					[-2.930, 3.780]	[-3.455, 4.184]
Media freedom					0.472***	0.545**
					[0.170, 0.774]	[0.069, 1.021]
Trade with Russia					-2.575	-2.055
					[-30.022, 24.872]	[-35.143, 31.034]
Distance to Russia					-0.052	-0.039
					[-0.176, 0.073]	[-0.179, 0.100]
Variance between countries (Intercept)	0.990	0.863	0.753	0.733	0.364	0.396
Variance between countries (Energy policy)			0.094	0.086	0.094	0.087
Variance between countries (Rus policy (no preference))			0.189	0.194	0.158	0.184
Variance between countries (Rus policy (defence))			0.895	0.790	0.912	0.769

	Model 1	Model 1(Full)	Model 2	Model 2(Full)	Model 3	Model 3(Full)
Variance between countries (Polit leanings (centrist))				0.063		0.065
Variance between individuals within countries (Intercept))	6.641	6.765	6.210	6.366	6.215	6.366
Num.Obs.	25939	22919	25939	22919	25939	22919
R2 Marg.	0.082	0.096	0.065	0.086	0.135	0.151
R2 Cond.	0.201	0.198	0.239	0.181	0.183	0.201
AIC	122909.4	109026.9	121324.3	107834.9	121371.3	107827.7
BIC	122991.0	109115.4	121479.4	107963.6	121518.3	107988.5

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 6: Extreme Views vs Far-right and Far-left (Robust Standard Errors)

	M3(Full)	M3(Full)(w/0 Polit.Reg)	M3(Full)	M3(Full)(w/0 Polit.Reg)
Intercept	4.825***	4.816***	4.804***	4.799***
	[4.506 – 5.144]	[4.487 – 5.144]	[4.541 – 5.067]	[4.524 – 5.074]
Energy Policy	0.182**	0.182**	0.185**	0.185**
	[0.057 – 0.307]	[0.057 – 0.307]	[0.060 – 0.309]	[0.060 – 0.309]
Russian Policy Preference (No preference)	1.164***	1.164***	1.163***	1.163***
	[0.967 – 1.362]	[0.967 – 1.362]	[0.964 – 1.362]	[0.964 – 1.361]
Russian Policy Preference (Defence)	1.381***	1.381***	1.379***	1.379***
	[0.999 – 1.763]	[0.999 – 1.762]	[1.003 – 1.755]	[1.003 – 1.755]
Age	0.255***	0.255***	0.255***	0.255***
	[0.190 – 0.320]	[0.190 – 0.320]	[0.190 – 0.320]	[0.189 – 0.320]

	M3(Full)	M3(Full)(w/0 Polit.Reg)	M3(Full)	M3(Full)(w/0 Polit.Reg)
Female	-0.088	-0.088	-0.088	-0.088
	[-0.199 – 0.023]	[-0.199 – 0.023]	[-0.199 – 0.023]	[-0.199 – 0.023]
Edu (post-secondary)	0.453***	0.453***	0.449***	0.449***
	[0.312 – 0.593]	[0.313 – 0.593]	[0.311 – 0.588]	[0.311 – 0.588]
Edu (tertiary)	0.681***	0.682***	0.676***	0.676***
	[0.507 – 0.856]	[0.507 – 0.856]	[0.505 – 0.847]	[0.505 – 0.847]
Far left & right	-0.603***	-0.602***		-0.063
	[-0.882 – -0.323]	[-0.880 – -0.323]		[-0.200 – 0.074]
Far right			-0.359**	-0.359**
			[-0.614 – -0.104]	[-0.614 – -0.103]
Far left			-0.814***	-0.814***
			[-1.206 – -0.422]	[-1.206 – -0.422]
Social media	-0.450	-0.129	0.325	0.480
	[-5.262 – 4.362]	[-4.576 – 4.318]	[-3.121 – 3.771]	[-2.767 – 3.727]
Media freedom	0.775	0.664**	0.542	0.489***
	[-0.003 – 1.553]	[0.002 – 1.140]	[-0.008 – 1.092]	[0.190 – 0.788]
Political regime	-0.213		-0.102	
	[-1.581 – 1.155]		[-0.987 – 0.782]	
Trade with Russia	-6.297	-5.677	-2.115	-1.827
	[-39.452 – 26.859]	[-38.190 – 26.836]	[-21.754 – 17.523]	[-20.864 – 17.209]
Distance to Russia	-0.080	-0.082	-0.031	-0.032
	[-0.222 – 0.062]	[-0.219 – 0.055]	[-0.093 – 0.031]	[-0.091 – 0.028]
Variance between countries (Intercept)	0.78	0.78	0.36	0.36
Variance between countries (Energy policy)	0.08	0.08	0.08	0.08

	M3(Full)	M3(Full)(w/0 Polit.Reg)	M3(Full)	M3(Full)(w/0 Polit.Reg)
Variance between countries (Rus policy (no preference))	0.17	0.17	0.16	0.16
Variance between countries (Rus policy (defence))	0.75	0.75	0.73	0.72
Variance between countries (Extreme)	0.27	0.27		
Variance between countries (Right)			0.20	0.20
Variance between countries (Left)			0.66	0.66
Variance between individuals within countries (Intercept))	6.32	6.32	6.31	6.31
Num.Obs.	22919	22919	22919	22919
R2 Marg.	0.179	0.178	0.148	0.148
R2 Cond.	0.269	0.269	0.194	0.194
AIC	107686.283	107685.247	107647.332	107645.445

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Coefficients of Multilevel Linear Regression Analysis (Robust Standard Errors)

	M3	M3(Full)	M3(w/0 Polit.Reg)	M3(Full)(w/0 Polit.Reg)
Intercept	4.801***	4.758***	4.801***	4.756***
	[4.512 – 5.091]	[4.496 – 5.020]	[4.513 – 5.089]	[4.485 – 5.027]
Energy Policy	0.134*	0.177**	0.134*	0.177**
	[0.009 – 0.258]	[0.050 – 0.304]	[0.009 – 0.258]	[0.050 – 0.304]
Russian Policy Preference (No preference)	1.216***	1.205***	1.216***	1.205***

	M3	M3(Full)	M3(w/0 Polit.Reg)	M3(Full)(w/0 Polit.Reg)
	[1.030 – 1.402]	[0.994 – 1.416]	[1.031 – 1.402]	[0.994 – 1.416]
Russian Policy Preference (Defence)	1.223***	1.394***	1.223***	1.394***
	[0.831 – 1.616]	[1.009 – 1.780]	[0.831 – 1.615]	[1.009 – 1.779]
Age	0.236***	0.257***	0.236***	0.257***
	[0.173 – 0.300]	[0.192 – 0.323]	[0.173 – 0.300]	[0.192 – 0.323]
Female	-0.055	-0.072	-0.055	-0.072
	[-0.156 – 0.047]	[-0.187 – 0.043]	[-0.156 – 0.047]	[-0.187 – 0.043]
Edu (post-secondary)	0.423***	0.458***	0.423***	0.458***
	[0.297 – 0.549]	[0.319 – 0.598]	[0.297 – 0.549]	[0.319 – 0.598]
Edu (tertiary)	0.630***	0.683***	0.629***	0.683***
	[0.463 – 0.796]	[0.503 – 0.864]	[0.464 – 0.795]	[0.503 – 0.864]
Political Leanings (Centrist)		-0.063		-0.063
		[-0.200 – 0.074]		[-0.200 – 0.074]
Social media	0.557	0.288	0.425	0.364
	[-2.369 – 3.483]	[-3.332 – 3.909]	[-2.350 – 3.200]	[-3.080 – 3.809]
Media freedom	0.428	0.571*	0.472***	0.545***
	[-0.053 – 0.909]	[0.002 – 1.140]	[0.312 – 0.632]	[0.236 – 0.854]
Political regime	0.075	-0.050		
	[-0.716 – 0.867]	[-0.982 – 0.881]		
Trade with Russia	-2.585	-2.197	-2.575	-2.055
	[-19.690 – 14.521]	[-23.254 – 18.860]	[-19.078 – 13.927]	[-22.013 – 17.904]
Distance to Russia	-0.053	-0.039	-0.052	-0.039
	[-0.109 – 0.004]	[-0.107 – 0.030]	[-0.109 – 0.005]	[-0.106 – 0.028]
Variance between countries (Intercept)	0.36	0.39	0.36	0.40

	M3	M3(Full)	M3(w/0 Polit.Reg)	M3(Full)(w/0 Polit.Reg)
Variance between countries (Energy policy)	0.09	0.09	0.09	0.09
Variance between countries (Rus policy (no preference))	0.16	0.18	0.16	0.18
Variance between countries (Rus policy (defence))	0.91	0.77	0.91	0.77
Variance between countries (Polit leanings (centrist))		0.07		0.07
Variance between individuals within countries (Intercept))	6.21	6.37	6.21	6.37
Num.Obs.	25939	22919	25939	22919
R2 Marg.	0.088	0.158	0.144	0.158
R2 Cond.	0.234	0.206	0.191	0.206
AIC	121352.360	107807.896	121393.2	107806.054

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 8: Comparing Freedom House and V-Dem measurements (Robust Standard Errors)

	M3	M3(Full)	M3(Altern)	M3(Full)(Altern)
Intercept	4.801***	4.758***	4.805***	4.872***
	[4.512 – 5.091]	[4.496 – 5.020]	[4.531 – 5.078]	[4.492 – 5.251]
Energy Policy	0.134*	0.177**	0.134*	0.177**
	[0.009 – 0.258]	[0.050 – 0.304]	[0.009 – 0.258]	[0.050 – 0.304]
Russian Policy Preference (No preference)	1.216***	1.205***	1.216***	1.205***
	[1.030 – 1.402]	[0.994 – 1.416]	[1.031 – 1.402]	[0.994 – 1.416]

	M3	M3(Full)	M3(Altern)	M3(Full)(Altern)
Russian Policy Preference (Defence)	1.223***	1.394***	1.224***	1.394***
	[0.831 – 1.616]	[1.009 – 1.780]	[0.831 – 1.617]	[1.008– 1.781]
Age	0.236***	0.257***	0.236***	0.257***
	[0.173 – 0.300]	[0.192 – 0.323]	[0.173 – 0.299]	[0.192 – 0.322]
Female	-0.055	-0.072	-0.055	-0.072
	[-0.156 – 0.047]	[-0.187 – 0.043]	[-0.156 – 0.047]	[-0.187 – 0.043]
Edu (post-secondary)	0.423***	0.458***	0.423***	0.458***
	[0.297 – 0.549]	[0.319 – 0.598]	[0.297 – 0.550]	[0.319 – 0.598]
Edu (tertiary)	0.630***	0.683***	0.630***	0.683***
	[0.463 – 0.796]	[0.503 – 0.864]	[0.464 – 0.796]	[0.503 – 0.864]
Political Leanings (Centrist)		-0.063		-0.063
		[-0.200 – 0.074]		[-0.200 – 0.074]
Social media	0.557	0.288	0.966	0.684
	[-2.369 – 3.483]	[-3.332 – 3.909]	[-1.186 – 3.117]	[-2.012 – 3.379]
Media freedom	0.428	0.571*	-0.070	0.068
	[-0.053 – 0.909]	[0.002 – 1.140]	[-0.797 – 0.657]	[-0.794 – 0.930]
Political regime	0.075	-0.050	2.517	2.364
	[-0.716 – 0.867]	[-0.982 – 0.881]	[-0.750 – 5.783]	[-1.645 – 6.373]
Trade with Russia	-2.585	-2.197	1.371	3.765
	[-19.690 – 14.521]	[-23.254 – 18.860]	[-21.772 – 24.515]	[-25.721 – 33.251]
Distance to Russia	-0.053	-0.039	-0.055	-0.039
	[-0.109 – 0.004]	[-0.107 – 0.030]	[-0.114 – 0.004]	[-0.101 – 0.022]
Variance between countries (Intercept)	0.36	0.39	0.32	0.36

	M3	M3(Full)	M3(Altern)	M3(Full)(Altern)
Variance between countries (Energy policy)	0.09	0.09	0.09	0.09
Variance between countries (Rus policy (no preference))	0.16	0.18	0.16	0.19
Variance between countries (Rus policy (defence))	0.91	0.77	0.91	0.77
Variance between countries (Polit leanings (centrist))		0.07		0.07
Variance between individuals within countries (Intercept))	6.21	6.37	6.21	6.37
Num.Obs.	25939	22919	25939	22919
R2 Marg.	0.088	0.158	0.143	0.159
R2 Cond.	0.234	0.206	0.184	0.204
AIC	121352.360	107807.896	121368.2	107806.201

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 12: Marginal Effects for Energy Policy Preferences with Confidence Intervals

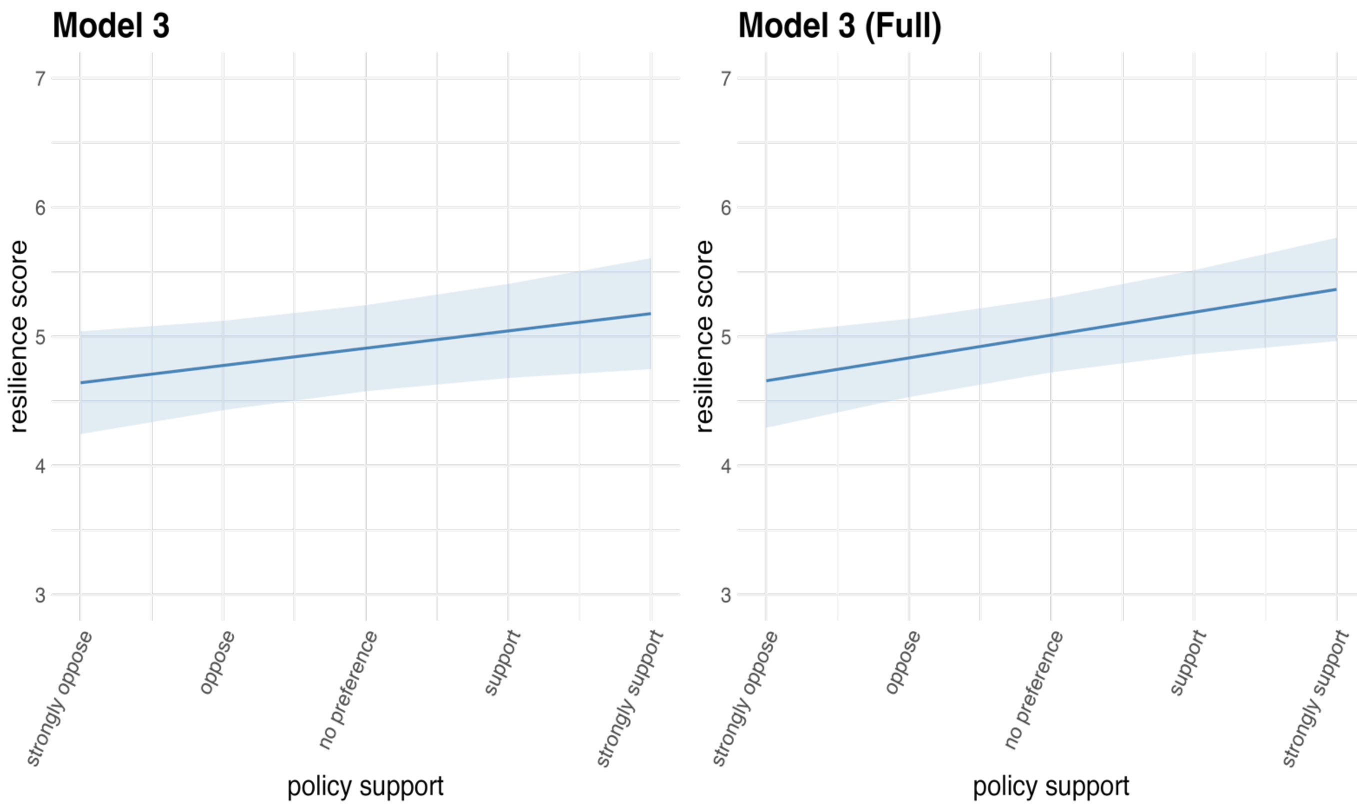


Figure 13: Marginal Effects for Attitudes Toward Russia with Confidence Intervals

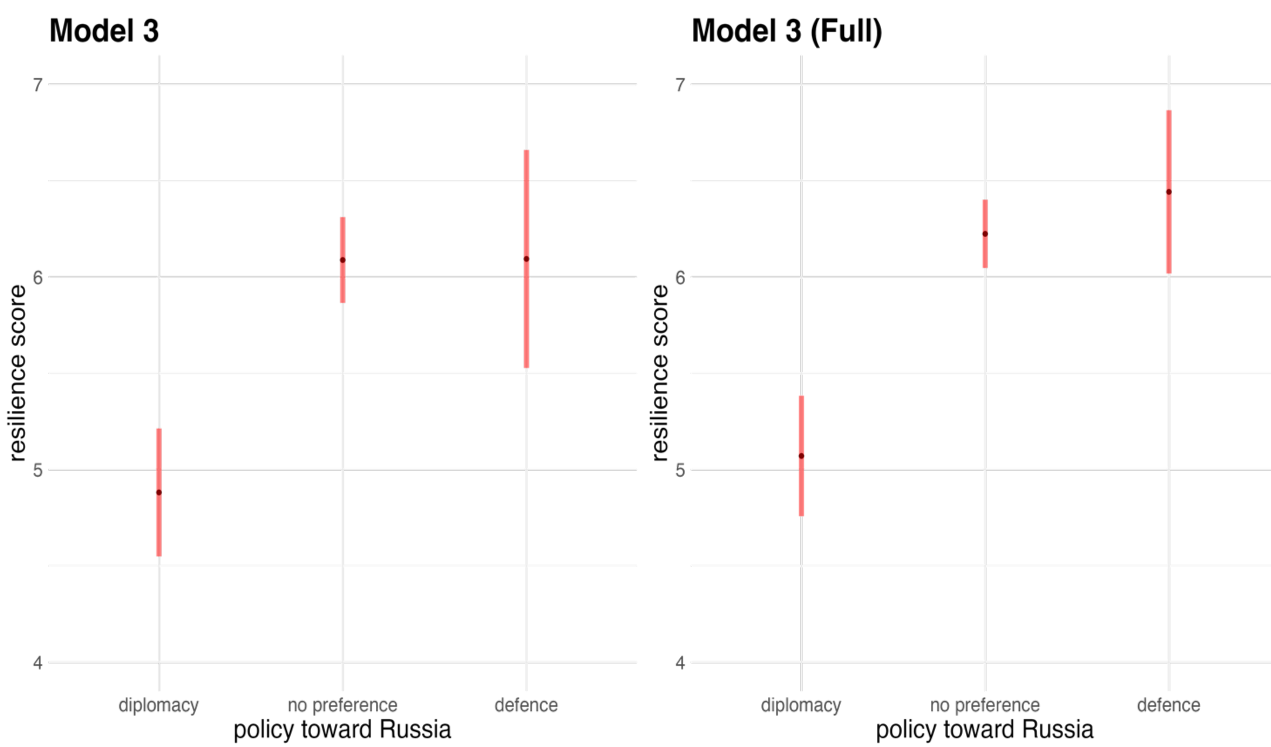
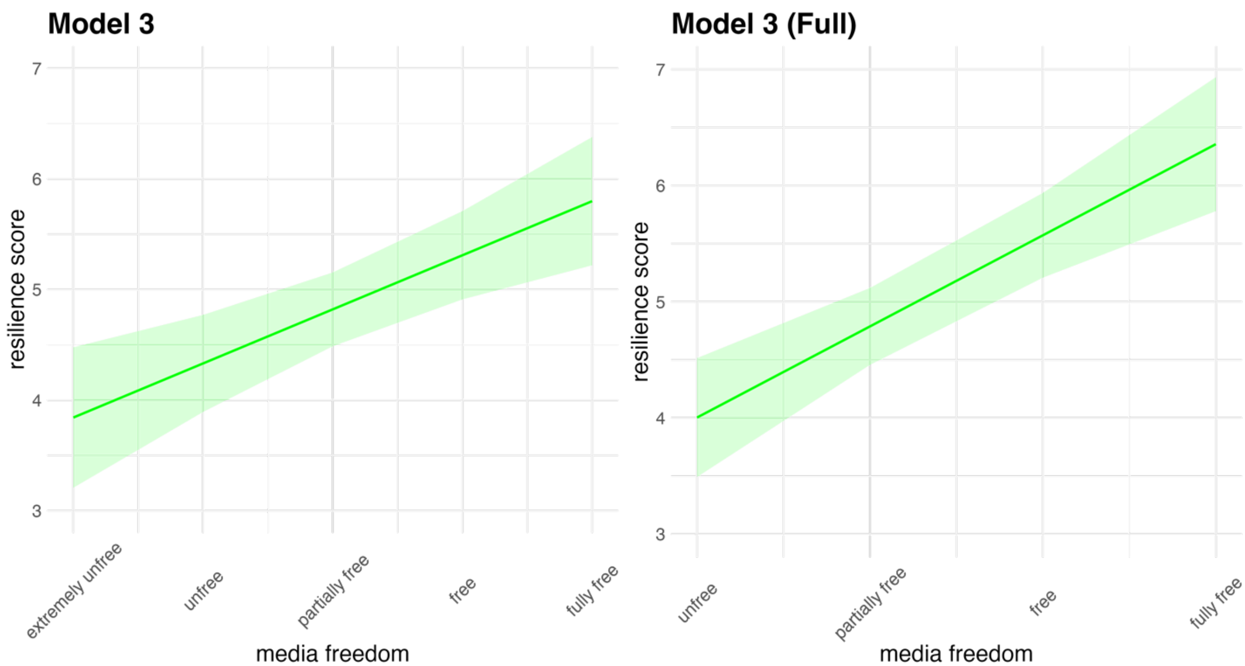


Figure 14: Marginal Effects for Media Freedom with Confidence Intervals



References

Aalberg, T., & Cushion, S. (2016). Public service broadcasting, hard news, and citizens' Knowledge of current affairs. In *Oxford Research Encyclopedia of Politics*.

Allcott, H., & Gentzkow, M. (2017). "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives*, 31 (2): 211-36. doi: 10.1257/jep.31.2.211.

Alyukov, M. (2022). Making sense of the news in an authoritarian regime: Russian television viewers' reception of the Russia–Ukraine conflict. *Europe-Asia Studies*, *74*(3), 337-359.

Apuke, O. D., & Omar, B. (2021). Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telematics and Informatics*, 56, 101475.

Arayankalam, J., & Krishnan, S. (2021). Relating foreign disinformation through social media, domestic online media fractionalization, government's control over cyberspace, and social media-induced offline violence: Insights from the agenda-building theoretical perspective. *Technological Forecasting and Social Change*, *166*, 120661.

Azarov, D., Koval, D., Nuridzhanian, G., & Venher, V. (2023). Understanding Russia's Actions in Ukraine as the Crime of Genocide. *Journal of International Criminal Justice*, mqad018.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255-278.

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European journal of communication*, *33*(2), 122-139.

Boulianne, S., Tenove, C., & Buffie, J. (2022). Complicating the resilience model: a four-country study about misinformation. *Media and Communication*, 10(3), 169-182.

Bueno, C. O. (2023). Russia's Role in the Far-Right Truck Convoy: An analysis of Russian state media activity related to the 2022 Freedom Convoy. *The Journal of Intelligence, Conflict, and Warfare*, *5*(3), 1-22.

Collins, G. (2017). Russia's use of the "energy weapon" in Europe. *Issue Brief,* 7. https://hdl.handle.net/1911/98850.

Digital News Report: Country and market data (2022). https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022/country-and-market-data.

Ecker, U.K.H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Panayiota Kendeou, P., Vraga, E.K. & Amazeen M.M. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology* 1, 13–29. https://doi.org/10.1038/s44159-021-00006-y.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological methods*, *12*(2), 121.

Erlich, A., & Garner, C. (2023). Is pro-Kremlin disinformation effective? Evidence from Ukraine. *The International Journal of Press/Politics*, 28(1), 5-28.

Erlich, A., Garner, C., Pennycook, G., & Rand, D. G. (2023). Does analytic thinking insulate against pro-Kremlin disinformation? Evidence from Ukraine. *Political Psychology*, *44*(1), 79-94.

Fallis, D. (2015). What is disinformation? *Library trends*, *63*(3), 401-426.

Flynn, D.J., Nyhan, B. & Reifler, J. (2017). The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs About Politics. *Advances in Political Psychology*, 38: 127-150. https://doi-org.proxy3.library.mcgill.ca/10.1111/pops.12394.

Freedom House. (2022). Country and Territory Ratings and Statuses, 1973-2023. https://freedomhouse.org/report/freedom-world.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.

Glatz, C., & Eder, A. (2020). Patterns of trust and subjective well-being across Europe: New insights from repeated cross-sectional analyses based on the European social survey 2002–2016. *Social Indicators Research*, *148*(2), 417-439.

globalEDGE. (2021). Market Potential Index (MPI). https://globaledge.msu.edu/.

Golan, G. (2006). Inter-media agenda setting and global news coverage: Assessing the influence of the New York Times on three network television evening news programs. *Journalism studies*, *7*(2), 323-333.

Gorwa, R. (2018). Poland: Unpacking the Ecosystem of Social Media Manipulation. In S. C. Woolley & P. N. Howard (Eds.), *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media* (online ed., pp. 86-103). Oxford Studies in Digital Politics, https://doi.org/10.1093/oso/9780190931407.003.0005.

Government of Canada: Countering disinformation with facts - Russian invasion of Ukraine (n.d.). Government of Canada. Retrieved October 20, 2023, from https://www.international.gc.ca/world-monde/issues_development enjeux_developpement/response_conflict-reponse_conflits/crisis-crises/ukraine-fact fait.aspx?lang=eng.

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science advances*, 5(1), eaau4586.

Guriev, S., & Treisman, D. (2019). Informational autocrats. *Journal of economic perspectives*, *33*(4), 100-127.

Hameleers, M. (2023). Disinformation as a context-bound phenomenon: toward a conceptual clarification integrating actors, intentions and techniques of creation and dissemination. *Communication Theory*, *33*(1), 1-10.

Hameleers, M., Brosius, A., & de Vreese, C. H. (2022). Whom to trust? Media exposure patterns of citizens with perceptions of misinformation and disinformation related to the news media. *European Journal of Communication*, 37(3), 237-268.

Headley, J. (2015). Challenging the EU's claim to moral authority: Russian talk of 'double standards'. *Asia Europe Journal*, *13*, 297-307.

Hinton, A. (2022, February). Putin's claims that Ukraine is committing genocide are baseless, but not unprecedented. *The Conversation*. https://theconversation.com/putins-claims-that-ukraine-is-committing-genocide-are-baseless-but-not-unprecedented-177511.

Holland, M. (2006). The Propagation and Power of Communist Security Services Dezinformatsiya. *International journal of intelligence and counterintelligence*, *19*(1), 1-31.

Humprecht, E., Esser, F., & Van Aelst, P. (2020). Resilience to Online Disinformation: A Framework for Cross-National Comparative Research. *The International Journal of Press/Politics*, 25(3), 493–516. https://doi.org/10.1177/1940161219900126.

Humprecht, E., Esser, F., Aelst, P. V., Staender, A., & Morosoli, S. (2021). The sharing of disinformation in cross-national comparison: Analyzing patterns of resilience. *Information, Communication & Society*, 1-21. https://doi.org/10.1080/1369118X.2021.2006744.

Henkel, I. (2021). Disinformation and Narratives. In: *Destructive Storytelling.* Palgrave Macmillan, Cham. https://doi-org.proxy3.library.mcgill.ca/10.1007/978-3-030-69503-3_6.

Hox J. J., Mirjam Moerbeek, & Rens van de Schoot. (2017). Multilevel Analysis:Techniques and Applications, Third Edition: Vol. Third edition. Routledge.

Kalathil, S. (2020). The Evolution of Authoritarian Digital Influence. *Prism*, *9*(1), 32-51.

Kilkenny, E. (2021). Russian Disinformation – The Technological Force Multiplier. *Global Insight: A Journal of Critical Human Science and Culture*, 2. https://doi.org/10.32855/globalinsight.2021.001

Kim, J.-N., & Gil de Zúñiga, H. (2021). Pseudo-Information, Media, Publics, and the Failing Marketplace of Ideas: Theory. *American Behavioral Scientist*, 65(2), 163–179. https://doi-org.proxy3.library.mcgill.ca/10.1177/0002764220950606.

Kim, M., & Cao, X. (2016). The impact of exposure to media messages promoting government conspiracy theories on distrust in the government: Evidence from a two-stage randomized experiment. I*nternational Journal of Communication*, 10, 20. Retrieved from https://ijoc.org/index.php/ijoc/article/view/5127.

Krickovic, A. (2015). When Interdependence Produces Conflict: EU–Russia Energy Relations as a Security Dilemma, *Contemporary Security Policy*, 36:1, 3-26, DOI: 10.1080/13523260.2015.1012350

Kuo, R., & Marwick, A. (2021). Critical disinformation studies: History, power, and politics. *Harvard Kennedy School Misinformation Review*, *2*(4), 1-11.

Kux, D. (1985). Soviet active measures and disinformation: Overview and assessment. *The US Army War College Quarterly: Parameters*, *15*(1), 17

La Cour, C. (2020). Theorising digital disinformation in international relations. *International Politics*, 57(4), 704-723.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., & Rothschild, D. (2018). The science of fake news. *Science*, 359 (6380), 1094–1096. https://doi.org/10.1126/science.aao2998.

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and Its Correction: Continued Influence and Successful Debiasing. *Psychological Science in the Public Interest*, 13(3), 106–131. https://doi.org/10.1177/1529100612451018.

Lewandowsky, S., Stritzke, W. G. K., Freund, A. M., Oberauer, K., & Krueger, J. I. (2013). Misinformation, disinformation, and violent conflict: From Iraq and the "War on Terror" to future threats to peace. *American Psychologist*, 68(7), 487–501. https://doi.org/10.1037/a0034515.

Lewandowsky, S. & Oberauer, K. (2016). Motivated rejection of science. *Current Directions in Psychological Science*, 25(4), 217-222.

Livingston, S. & Bennett, L. (2020). The Institutional Crisis at the Root of Our Political Disinformation and Division. *Social Science Research Council*, https://items.ssrc.org/disinformation-democracy-and-conflict-prevention/the-institutional-crisis-at-the-root-of-our-political-disinformation-and-division/.

Lyammouri, R., & Eddazi, Y. (2020). Russian interference in Africa: Disinformation and mercenaries. *Policy Centre for the New South*, PB-20(60), 1–6. https://www.policycenter.ma/sites/default/files/2021-01/PB_20-60_Lyammouri.pdf.

Lyndon, R., Tse, V., Moore, L. & May-Hobbs, M. (2022). Disinformation In Brazil: The 2019 Amazon fires on social media. In: Mair, M., Meckin, R. and Elliot, M. (Eds) *Investigative Methods: An NCRM Innovation Collection.* Southampton: National Centre for Research Methods, pp. 44-53. DOI: 10.5258/NCRM/NCRM.00004547.

Madrid-Morales, D., Wasserman, H., Gondwe, G., Ndlovu, K., Sikanku, E., Tully, M. & Uzuegbunam, C. (2021). Comparative approaches to mis/disinformation| motivations for sharing misinformation: A comparative study in six Sub-Saharan African countries. *International Journal of Communication*, 15, 20.

Maas, C. J., & Hox, J. J. (2004). The influence of violations of assumptions on multilevel parameter estimates and their standard errors. *Computational statistics & data analysis*, *46*(3), 427-440.

Mayer, T. & Zignago, S. (2011), Notes on CEPII's distances measures : the GeoDist Database*,* *CEPII Working Paper 2011-25*. Retrieved September 15, 2023, from http://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=6.

Mohan, V., & Wall, A. (2019). Foreign electoral interference. *Georgetown Journal of International Affairs*, 20, 110-119.

Nagasako, T. (2020). Global disinformation campaigns and legal challenges. *International Cybersecurity Law Review.* 1, 125–136. https://doi.org/10.1365/s43439-020-00010-7.

Nelson, A., & Stella Patino, S. (2021). COVID-19, A Tale of Two Pandemics: Novel Coronavirus and Fake News Messaging. *Health Promotion International*, 36(2), 524–534.https://doi.org/10.1093/heapro/daaa140.

Neudert, L.-M. N. (2018). Germany: A Cautionary Tale. In S. C. Woolley & P. N. Howard (Eds.), *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media* (online ed., pp. 153-184). Oxford Studies in Digital Politics, https://doi.org/10.1093/oso/9780190931407.003.0008.

Nyhan, B., & Reifler, J. (2010). When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior.* 32, 303–330 (2010). https://doi.org/10.1007/s11109-010-9112-2.

OECD: Disinformation and Russia's war of aggression against Ukraine (2022). Organisation for Economic Co-operation and Development. https://www.oecd.org/ukraine-hub/policy-responses/disinformation-and-russia-s-war-of-aggression-against-ukraine-37186bde/.

Paul, C., & Matthews, M. (2016). The Russian “firehose of falsehood” propaganda model. Rand Corporation, 2(7), 1-10.

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of school psychology*, *48*(1), 85-112.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). sage.

Reporters Without Borders. (2022). Press Freedom Index. https://rsf.org/en/index?year=2022.

Sanovich, S. (2018). Russia: The Origins of Digital Misinformation. In S. C. Woolley & P. N. Howard (Eds.), *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media* (online ed., pp. 21-40). Oxford Studies in Digital Politics, https://doi.org/10.1093/oso/9780190931407.003.0002.

Stanley J. (2018). What John Stuart Mill got wrong about freedom of speech. *Boston Review. https://www.bostonreview.net/articles/jason-stanley-what-mill-got-wrong-about-freedom-of-speech/.*

Statista (2022). Most used social media platforms in Saudi Arabia as of 3rd quarter of 2022. Statista. Retrieved September 15, 2023, from https://www.statista.com/statistics/1394073/saudi-arabia-most-used-social-media-platforms/#:~:text=According%20to%20a%20survey%20about,came%20second%20with%2076.8%20percent.

Shu, K., Bhattacharjee, A., Alatawi, F., Nazer, T. H., Ding, K., Karami, M., & Liu, H. (2020). Combating disinformation in a social media age. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(6), e1385.

Snegovaya, M. (2015). Putin’s information warfare in Ukraine. *Soviet Origins of Russia’s Hybrid Warfare’, Russia Report*, *1*, 133-135.

Subramanian, S. V., Jones, K., Kaddour, A., & Krieger, N. (2009). Revisiting Robinson: the perils of individualistic and ecologic fallacy. *International journal of epidemiology*, *38*(2), 342-360.

Tenove, C. (2020). Protecting Democracy from Disinformation: Normative Threats and Policy Responses. *The International Journal of Press/Politics*, 25(3), 517–537. https://doi-org.proxy3.library.mcgill.ca/10.1177/1940161220918740.

van der Linden, S., Panagopoulos, C., Azevedo, F. & Jost, J.T. (2021). The Paranoid Style in American Politics Revisited: An Ideological Asymmetry in Conspiratorial Thinking. *Political Psychology*, 42: 23-51. https://doi.org/10.1111/pops.12681.

Walker, C. (2016). The Authoritarian Threat: The Hijacking of" Soft Power". *Journal of democracy*, 27(1), 49-63.

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking* (Vol. 27, pp. 1-107). Strasbourg: Council of Europe.

Wintour, Patrick. (2024, February). US leading global alliance to counter foreign government disinformation. *The Guardian*. https://www.theguardian.com/technology/2024/feb/26/us-leading-global-alliance-to-counter-foreign-government-disinformation.

Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in quantitative methods for psychology*, *8*(1), 52-69.

Wood, T. & Porter, E. The Elusive Backfire Effect: Mass Attitudes' Steadfast Factual Adherence. *Political Behavior.* 41, 135–163 (2019). https://doi.org/10.1007/s11109-018-9443-y.

Woolley, S. C., & Howard, P. N. (Eds.). (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press.

Woolley, S. C., & Howard, P. N (b). (2018). Introduction: Computational Propaganda Worldwide. In S. C. Woolley & P. N. Howard (Eds.), *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media* (online ed., pp.3-18). Oxford Studies in Digital Politics, https://doi.org/10.1093/oso/9780190931407.003.0002.

Yablokov, I. (2022). Russian disinformation finds fertile ground in the West. Nature Human Behavior 6, 766–767. https://doi.org/10.1038/s41562-022-01399-3.

Yariv T., Boomgaarden, Strömbäck H. G., Vliegenthart R., Damstra A. & Lindgren E. (2020). Causes and Consequences of Mainstream Media Dissemination of Fake News: Literature Review and Synthesis. *Annals of the International Communication Association*, 44(2), 157-173, DOI: 10.1080/23808985.2020.1759443.

Zhdanova, M. & Orlova D. (2018). Ukraine: External Threats and Internal Challenges. In S. C. Woolley & P. N. Howard (Eds.), *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media* (online ed., pp. 41-63). Oxford Studies in Digital Politics, https://doi.org/10.1093/oso/9780190931407.003.0003.