Faculty of Business Studies and Economics

Dissertation for the attainment of the doctoral degree Dr. rer. pol. by the

Doctoral Commission Dr. rer. pol. of the University of Bremen

The Political Economy of Digital Repression

David Karpa

supervised by

Prof. Dr. Torben Klarl

and

Prof. Dr. Christian Cordes

Submission: August 27, 2024

Defense: December 12, 2024

Abstract

How can we explain why governments become more repressive and the lack of public resistance? With this dissertation, I offer several explanations. First, technological progress leads to new tools of repression that do not suffer from the same drawbacks as their predecessors but, on the contrary, have *positive* externalities that contribute to authoritarian stability. Second, media manipulation helps build public support for digital governance, which in turn increases the repressive capacity of regimes. Even without media manipulation, and regardless of the institutional context, digital and algorithmic governance faces very little resistance. Third, propaganda can be used strategically and is effective in shaping attitudes and behavior (beyond support for individual policies). This goes as far as persuading people to avoid sanctions on war-critical goods - in the case of Kazakhstanis trading with Russia. Fourth, new digital surveillance tools offer the dual functionality of monitoring society for and deterring overt dissent. Digital surveillance thus limits the potential for anti-government action while not requiring the same level of supporting bureaucracy, and having the side effects and costs as its analog predecessor. These findings are discussed in light of the development of contemporary societies. As such, this dissertation contributes to the literature on (the political economy of) digital authoritarianism, and many subfields to which the individual chapters relate.

Contents

List of Figures

List of Tables

1 Introduction

The advent of the digital age has ushered in a new era of political control, one that is far more subtle and insidious than the blunt instruments of the past. Whereas authoritarian regimes once relied on brute force and overt repression to maintain their grip on power, today's digital autocrats have found ways to subvert technology for their own ends. While the internet was once praised as a tool for liberation and democratization, it has also been co-opted and transformed into a powerful mechanism of surveillance and social control. This dissertation seeks to unravel the complex political economy that underpins this new age of digital repression, exploring how technological advancement and contemporary applications reinforce the power of autocratic rulers. Through a series of empirical investigations, I aim to shed light on the multifaceted drivers behind the rise of surveillance states and informational autocracies. I consider and discuss the implications for the future of democracy and freedom in the digital age.

Attempts to explain political phenomena have been documented at least since Plato and Aristotle formulated *Politeia* and *Politiká*, respectively. Authoritarian regimes have long utilized surveillance and repression as tools of control, a practice with deep historical roots. From the frumentarii and speculatores in ancient Rome to the Inquisition, from Vidocq under Napoleon to the numerous surveillance agencies of the 20th century – such as the Gestapo, NKVD, KGB, Stasi, and others. This dissertation shifts the focus from conventional uses of surveillance and repression to the emerging digital forms of surveillance and repression in the 21st century. But how has technological change transformed the face of repression in the digital age, and how has it affected existing power structures? To address these questions, this introduction reviews the relevant recent literature from economics and political science that describes authoritarian rule and informs this dissertation's research. In order to approach the recent literature on informational autocracies and digital authoritarianism, an overview of the influential literature describing authoritarian politics and state repression is first presented. This literature is strongly influenced by the analysis of totalitarian detours in the 20th century and developments since then – the transition from totalitarian to authoritarian rule.

Leaders, both in the latter half of the 20th century and today, have confronted two significant challenges: managing authoritarian control and navigating the complexities of power-sharing within authoritarian regimes (Svolik 2012). Put simply, the problem of au-

thoritarian control is how citizens are prevented from challenging or changing a government that doesn't meet their needs. In the absence of authoritarian control, the masses could revolt and overthrow a regime that does not satisfy their preferences for policies and governance, and that continues to produce institutions that are unsatisfactory to the majority. These efforts by citizens to coordinate and challenge those in power are met with either repression or co-optation. In the first case, dictators resort to repression in order to maintain power, which will hinder citizens from coordinating and challenging power structures. But, as we will see throughout the introduction, repression also comes at a cost: it can have adverse effects and create backlash from the domestic population, or lead to sanctions and international isolation. There are also economic costs that come in the form of lower efficiency and fewer incentives for innovation when intuitions are not inclusive (Acemoglu and Robinson 2012). Alternatively, dictators resort to co-opting their populations and eliminating the desire to change the status quo. Buying off elites and large groups can be costly, and efficiently managing the balance between repression and co-optation is a central challenge for contemporary autocrats (Xu 2021).

The problem of authoritarian power-sharing, essentially, concerns the balance of power between the dictator and her allies (Svolik 2012). Unlike democracies, where power is distributed through institutional checks and balances, authoritarian leaders must navigate a delicate balance of power among various elites – such as military officials, party leaders, and influential business figures – whose support is essential for maintaining control. The leaders must share enough power and resources to keep these elites loyal while ensuring that no single faction becomes strong enough to threaten their own authority. This fragile arrangement requires constant vigilance and strategic distribution of influence, as any imbalance can lead to internal conflict, coups, or even the collapse of the regime.

The two problems of authoritarian rule are also intertwined. The inherent challenge faced by authoritarian rulers in obtaining accurate and reliable information while maintaining control over their regime is referred to as the *informational dilemma* (Egorov and Sonin 2024). Autocrats rely heavily on information to make decisions, ensure loyalty, and suppress dissent. However, the very nature of authoritarianism creates a paradox: the more an autocrat centralizes power and suppresses opposition, the less likely they are to receive

truthful information. This dilemma arises because subordinates and bureaucrats within an authoritarian regime often have incentives to distort information to please the autocrat, avoid punishment, or gain favor. Fear of retribution for delivering bad news or critical reports can lead to a culture of sycophancy, where officials present only positive or skewed information. As a result, the autocrat may be insulated from the realities of the situation, leading to poor decision-making and potential instability within the regime. While surveillance and repression are tools to control the flow of information, they can exacerbate the dilemma by further discouraging transparency and honest communication. This informational asymmetry makes it difficult for autocrats to accurately assess threats, gauge public sentiment, or address problems effectively, ultimately undermining the stability and longevity of their rule.

This dissertation is mainly concerned with the first problem of authoritarian rule – that of authoritarian control. Furthermore, the focus is on *digital* authoritarian tools securing control. That makes this dissertation essentially one about digital authoritarianism, or, in other words, digital repression. One definition of digital repression is used by Feldstein (2021, p.25), who describes it "as the use of information and communications technology to surveil, coerce, or manipulate individuals or groups in order to deter specific activities or beliefs that challenge the state". While the problems of authoritarian rule have not fundamentally changed with the advent of the Information Age, the tools available to meet this challenge have.

This chapter serves as an introduction to the remainder of this dissertation. In order to guide the reader towards the following chapters which are essays engaging in distinct research questions, the relevant literature will be introduced in section 1.1. In addition, the introduction summarizes the remaining chapters, embeds them in the literature and highlights the contributions where appropriate. An outlook on further research and a discussion of the implications of the research concludes the introduction, in section 1.2.

The chapters 2 to 5 constitutive the main research done by the author and co-authors. Chapter 2 reviews the literature on political institutions and innovation processes, with a focus on the effect that authoritarian institutions have. A formal model is developed that reflects on the crucial role of surveillance and the gathered data in the innovation process of artificial intelligence. Chapter 3 uses experimental data to investigate the determinants of

public support for algorithmically enhanced governance. The chapter investigates whether public support diminishes when citizens are reminded of the repressive potential of this governance, in conjunction with an analysis on and discussion of the role that media plays in forming public opinion. Chapter 4 is more specifically concerned with the effects of propaganda on shaping public opinion and behaviour. Through a combination of various data sources, a link between Russian language proficiency, Russian media consumption, and sanction circumvention in Kazakhstan is demonstrated. Finally, chapter 5 draws on experimental data to examine the effects of increasing state surveillance capabilities. As hypothesized, state surveillance is found to increase self-censorship with strong heterogeneity in the effect, contributing to both inequality and authoritarian stability.

1.1 Digital Repression

Repression has many faces, but it's essence is the countering or elimination of a behavioural threat that challenges the status quo (Davenport 2007). Behavioural threats come in the form of protests, movements, elites, or coups, and the like. Whereas power must be shared with elites to prevents coups, the non-elite public does not get to participate in politics in an autocracy. When the public threatens to revolt, co-optation or repression is used to control the population (Svolik 2012). In the following, an overview of digital repression will be laid out. It will be guided along the analytical framework of Earl et al. (2022), who develop a typology that distinguishes digital repression by three different dimensions. First, they differentiate the kind of repressor: state or private actors and, for the case of state actors, whether those actors are directly controlled by national governments or are more decentralized (e.g., local police in the United States). Second, they distinguish between overt or covert repression. Overt repression is recognizable by others – sometimes on purpose in order to deter. Covert repression, conversely, works by repressors at least trying to conceal their actions; sometimes out of necessity, for example a secret service making use of informants, sometimes in order not to evoke a backlash by using visible repression. Third, Earl et al. (2022) distinguish between (i) attempts at control that involve force and coercion or (ii) instead seek to incentivize preferred forms of expression and behaviour. The first, *coercion*, refers to violence, arrests, and surveillance, whereas the latter, *channelling*, refers

to incentives given by laws or regulations, or in other words, carrots and sticks.

What makes this typology extremely useful for guiding this literature review is that it describes both digital repression *drawing* on traditional processes and digital repression *expanding* on traditional processes. In other words, the typology describes digital repression as evolving tools that suit new (digital) environments (Table 1.1), and as new tools stemming from new (digital) environments (Table 1.2). The chapters of this dissertation encompass phenomena from a multitude of categories from this typology. Conveniently, all of the individual chapters are describable by the typology at hand, which allows for an efficient embedding of the research into the landscape of the literature. By following this perspective, the necessary attention can be paid to both contemporary developments of digital repression – but also to its roots in analog repression. Due to "clear overlaps between traditional coercive strategies—detentions, torture, beatings, extrajudicial killings – and digital variants" (Feldstein 2021, p.13), a comprehensive overview of repression *before* the digital age is neglected in this review, in order to avoid redundancy. In the following, an overview of different forms of repression is presented along the typology by Earl et al. (2022), in combination with an embedding of the chapters 2 to 5 into the analytical framework and literature. Wherever possible, examples from the original article by Earl et al. (2022) will be either exchanged or complemented by examples from the region this dissertation is (mostly) concerned with – that is, Eurasia and Central-Asia in particular.

1.1.1 Digital repression drawing on traditional processes

In this section, the literature on digital repression drawing on traditional repression processes will be reviewed, guided by Table 1.1 drawing on the typology developed by Earl et al. (2022). Earl et al. (2022) describe how state repression that targets social movements, protests, and activism has evolved from analog to digital repression. An earlier review by Earl (2003) on analog repression established a theoretical framework that provided the basis for the typology in Earl et al. (2022).

Table 1.1: Digital repression drawing on traditional processes.

	Physical control			
	Physical coercion (e.g., violence, arrests, and surveillance)		Channeling (i.e., carrots for preferred behavior or overbroad sticks)	
	Overt	Covert	Overt	Covert
State agents tightly coupled with national political officials	Physical violence or legal action against digital activists by militaries or national police (e.g., arrest of bloggers) -1-	Digital surveillance by national authorities (e.g., NSA surveillance in the United States) -2-	State-sanctioned online grievance platforms (e.g., online petitions to the White House site) -3-	National laws or policies that limit online speech and/or activity (e.g., online morality and defamation laws), including but not limited to dissent -4-
State agents loosely connected with national political officials	Physical violence or legal action against digital activists by local police (e.g., arrests of Twitter account holders) -5-	Digital surveillance by local authorities (e.g., local U.S. police stingray use to monitor protesters' cellphones) -6-	Local government online grievance platforms (e.g., local government complaint sites) -7-	Regional or local social credit systems (e.g., local experimentation with social credit systems in Chinese cities) -8-
Private agents	Physical violence, harassment, or legal action by private actors (e.g., individuals and groups doxing and harassing protesters online; private lawsuits to harass online activists) -9-	Private surveillance (e.g., security contractors tracking protesters through online media) and surveillance capitalism -10-	Corporate online complaint forums and/or organizational social media policies (e.g., policies about candidate and/or employee social media usage) -11-	Platform community standards and/or platform reward structures (e.g., Facebook and Twitter) -12-

Note: Source: Earl et al. (2022). The Table was recreated without introducing changes to the original Table.

Overt physical coercion

The first column of Table 1.1 (cells 1, 5, 9) is concerned with overt physical coercion by state and private actors. Most research on this topic has focused on police or military violence. To this category belong the most obvious forms of repression – physical and legal coercion, or the threat thereof. A high degree of visibility is often not an unintended side-effect, but part of the design. The underlying logic is that visible repression can deter not only repressed individuals, but also others from engaging in undesired behaviour. This emphasis on deterrence is echoed by the literature on censorship, in which censorship is modelled as self-disciplining behaviour induced by fear of repression (Roberts 2018). This mechanism of "dissuasion" (Roberts 2018, p.45) – discouraging spreading or accessing information by explicitly articulating a corresponding expected punishment – is here understood as repression, not censorship.

Overt *physical* coercion in the form of imprisonment and torture of activists and dissidents or disappearances and political killings have been at the core of the repressive apparatus of the totalitarian regimes of the 20th century (Guriev and Treisman 2019). While

a decrease in political imprisonment and killings can be seen over the last decades (Guriev and Treisman 2019, 2022), autocrats still resort to violence, especially in more closed non-democracies. Recent examples are the cases Saudi bloggers who were arrested and tortured (Pan and Siegel 2020), the imprisonment and consequent murder of Alexei Navalny, or the imprisonment of Kazakh online activists (Anceschi 2015). The first study on Saudi bloggers discusses the different effects that the arrests had (Pan and Siegel 2020). One the one hand, arrested bloggers were silenced (direct effect), whereas their followers did not decrease their criticism of the regime (no indirect effect).[1] However, subsequent to the changed behaviour of bloggers who decreased their negative sentiment towards the regime, other influencers and bloggers consequently adopted the change in sentiment and were less critical of the regime (downstream effect).

Examples of *legal* deterrence include laws that prohibit the insulting of Islam in Saudi-Arabia or Russian laws that forbid spreading "fake news" or mentioning the Russo-Ukrainian war (Manea 2016). Similarly vague laws are in place in Turkey,[2] or Iran, where the spreading of "immoral" information is prohibited (Roberts 2018). The vagueness of these laws allows for flexibility in ruling, which is often used to the disadvantage of citizens, and also allows to fabricate cases. For instance, Alexei Navalny, who was killed in a Russian prison, was sentenced to 19 years on fabricated charges.[3] Non-state actors also contribute to digital repression – sometimes in accordance with the law, for example in Germany, where controversial content removals have been reported corresponding to the NetzDG.[4] In this particular case, posts of news outlets were censored by Twitter. In other cases private actors engage in repression in discordance with the law, for example, when actors receive threats on social media. Many individual cases are documented and sometimes uncoordinated individual efforts culminate in more coherent attacks, as in China, where nationalists attacked Taiwanese Facebook pages.[5] In other cases women were being harassed more or less systematically on

[1] For a review on backlash or adverse effects of digital repression, see Roberts (2020).

[2] https://freedomhouse.org/country/turkey/freedom-net/2023

[3] https://www.hrw.org/news/2024/02/16/russia-navalny-dies-prison

[4] https://freedomhouse.org/country/germany/freedom-net/2023

[5] https://www.ft.com/content/5ae7b358-ce3c-11e6-864f-20dcb35cede2

the internet, often for their public and/or political statements (Earl et al. 2022).[6]

Importantly, Earl et al. (2022) do not distinguish between repression and digital repression, since activists and movements operate in both physical *and* digital spaces. They suggest that digital repression research should attend to any actors who are threatening to a regime and often these actors are bloggers, activists or journalists who post online and have a large number of followers on social media. Another important point is that digital repression is more prevalent in, but not limited to non-democracies – a theme that will be emphasized more throughout the review (and has already become evident in the example corresponding to Germany and the NetzDG). Predictive policing, for example, is being used for chilling protests and collective action not only in autocracies, but also elsewhere (Walsh and O'Connor 2019). Other examples include the arresting and raiding of activists' homes based on Twitter activities in the U.S. (Earl et al. 2013).

One crucial aspect often under-emphasized in the literature is the *dual function* of digital surveillance, which is why it is allocated twice in the typology. Digital surveillance serves multiple purposes, or more precisely, has multidimensional effects. On the one hand (digital) surveillance enables the state or private agents to collect information and infer about individuals *in* the data but also *outside* of the data (Zuboff 2019). Given a sufficient amount of data about individuals with specific traits, inference on *other* individuals with these traits is also possible. This aspect of surveillance enables targeted repression (Roberts 2020; Xu 2021; Gohdes 2023) and also helps the authoritarian state to solve it's informational dilemma. This opaque form of surveillance belongs to the category of *covert* physical coercion and is hence subject to the next section 1.1.1.

On the other hand, when visible, digital surveillance functions as a *complement* to overt repression, in that it increases (enables) the deterrent effect. Citizens being aware of surveillance may self-discipline their behaviour due to the possibility of repression. Indeed, the whole idea of deterring implicitly necessitates some form of surveillance capacity. Deterrence would not occur if the perception of behaviour going unnoticed by the authorities prevailed. In Figure 1.1, this corresponds to the *increased expectation of negative outcomes*,

[6] https://blogs.worldbank.org/en/developmenttalk/protecting-women-and-girls-cyber-harassment-global-assessment

which arises due to the perception of being under surveillance. If individuals are aware of state or private actors having access to GPS data, text messages, browsing history, etc., the calculation changes and behaviour is adapted. In the literature, this deterrent component of surveillance has also been called "chilling effect" (Schauer 1978).

Chilling effects – the deterrence of lawful behaviour out of fear that it is suspect – have been studied explicitly by several scholars (Penney 2016, 2017, 2022; Stoycheff 2016; Stoycheff et al. 2019; Büchi et al. 2022; Kappeler et al. 2023) and implicitly by many others (Roberts 2018; Manokha 2018; Tannenberg 2022; Stoycheff 2022; Oz and Yanik 2022). Where others have focused on the preventive effects of targeted repression enabled by surveillance (Xu 2021; Pei 2024), here the idea is that individuals exercise *self*-restraint, because of surveillance and the possible implication of repression. In the words of Pan and Siegel (2020) discussed above, chilling effects refer to the *indirect* effect of repression.[7]

Büchi et al. (2022) develop a theoretical model of chilling effects and argue that *salience shocks* of digital surveillance lead to inhibited digital communication behaviour, see Figure 1.1. Importantly, the authors model *salience shocks* as accidents. Consider Edward Snowden's revelations about the NSA's continued surveillance of U.S. citizens as the prime example for involuntarily revelations. I argue this model might be incomplete, since it curtails the value of *strategic* signaling of surveillance capabilities – in order to invoke a deterrence effect. In other words, it must not necessarily be accidental when the public becomes aware of the surveillance capabilities and practices.

Irrespective of the voluntariness of the revelations, previous studies have shown that salience shocks with respect to digital surveillance create chilling effects. An experimental study has shown, for example, that citizen's *willingness to speak out* is decreased after having received a message that primed them to perceive their online activities were subject to surveillance by the U.S. government (Stoycheff 2016). Another study found that treatment stimuli that mimic Terms of Service which allow for government surveillance – again in the U.S. – decrease a range of measures that correspond to what Büchi et al. (2022) have termed "lower intention to engage in in digital communication", see Figure 1.1. Similarly,

[7] As mentioned earlier, no deterrence effect was found in the study by Pan and Siegel (2020), perhaps because the absence of Saudi surveillance capabilities may have rendered their deterrence efforts ineffective.

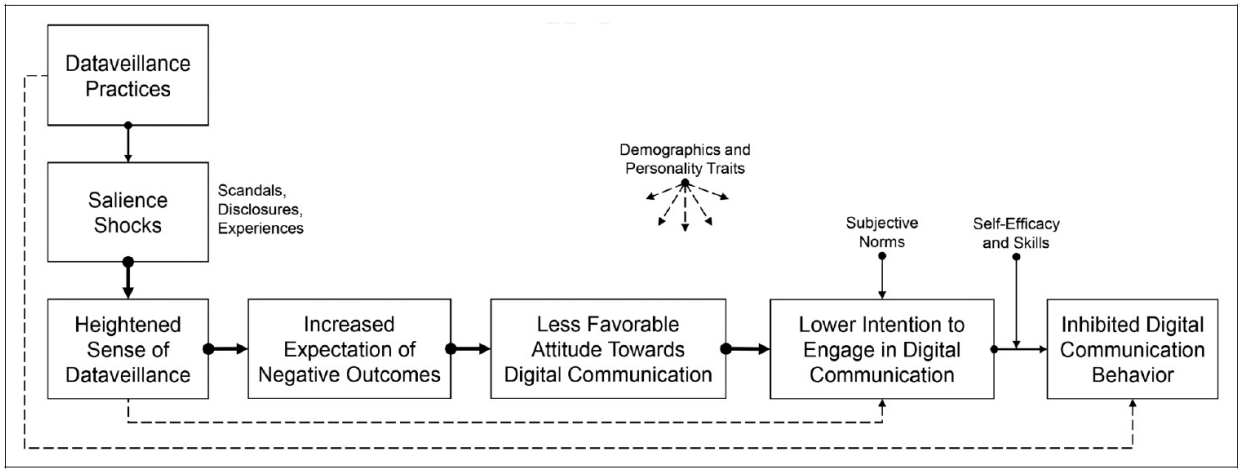
experimental data presented in Büchi et al. (2022) captures decreased *intentions* towards engaging in digital communication, or, in other words, an increase in the *intention* to self-censor. The role of emotions is also experimentally investigated by Stoycheff (2022) who finds that fear and anger mediate chilling effects.

Recent research on surveillance in autocracies has implicitly suggested a similar mechanism, albeit in other words: surveillance induces self-discipline (mostly self-censorship) due to the fear of repression (Roberts 2018; Manokha 2018; Tannenberg 2022; Stoycheff 2022; Oz and Yanik 2022). Other research has started to investigate the underlying (cognitive) mechanisms of coping with surveillance (Festic 2022; Kappeler et al. 2023). In another study, cross-country panel data revealed decreasing political participation and democratization due to online surveillance in 21 countries (Stoycheff et al. 2020). There might be strong heterogeneity with respect to differences between cultures and generations, and the fine-grained attitudes towards institutions which have been formed by individual experiences (Kalmus et al. 2022).

I contribute to the general literature on digital authoritarianism and the specific literature on digital surveillance and chilling effects by showing how surveillance reduces digital communication behaviour in an experimental study in chapter 5 of this dissertation. In this paragraph, I highlight key contributions of this study to the literature. First, I make use of a randomized controlled trial, in which participants of the study are randomly assigned in either of two treatment groups, or a control group.[8] This allows me to causally infer the effect of each treatment condition compared to the control group. Second, I employ a design that allows me to directly measure self-censorship. In terms of Büchi et al. (2022) and Figure 1.1, I am able to directly measure *inhibited digital communication behaviour*, or, self-censorship. The study thus methodologically advances the field in going beyond measuring *intentions*, as has been done before by others described above. This contributes to a high degree of external validity, that is, the degree to which the effects I find in the study seem to be corresponding to phenomena in the real-world. Third, I conduct this study online with citizens from Kazakhstan, which again lends my study a high degree of external validity for two

[8]The surveillance condition reminds participants of the possibility of state surveillance, whereas the privacy condition reminds them of the encryption of internet traffic.

Figure 1.1: Chilling effects of digital surveillance on the individual level



Note: *Source:* Taken from Büchi et al. (2022). Dataveillance refers to "the automated, continuous, and unspecific collection, retention, and analysis of digital traces by state and corporate actors" and is thus a subcategory of digital surveillance and a technical way of referring to what is commonly known as *mass surveillance*.

reasons: (i) Kazakhstan is a country with a large-scale mass surveillance program of which the population is aware (Raman et al. 2020), so the *salience shock* which I experimentally create can evoke *real* chilling effects, and (ii), because the context the study generalizes to is *digital* surveillance and inhibited *digital* communication behaviour, and I conduct my study in the same *digital* space to which I am trying to generalize to. The study thus advances the field methodologically, and tests existing theories in new and suitable contexts.

Indeed, I do find chilling effects due to digital surveillance in Kazakhstan. The main results of the study in chapter 5 are that participants in the surveillance condition self-censor, 4% on average on items concerning domestic politics and between 2.5 and 3.2% on geopolitical topics, whereas exposure to the privacy treatment does not yield an average effect that is statistically significant. Strong heterogeneity in the surveillance treatment effects was detected, with effect sizes increasing up to three times the size (9%), or diminishing entirely, for some demographic groups. The implications are discussed in detail in the chapter itself, to which the reader is referred. In the typology in Figure 1.1, chapter 5 is best described by cell 1, as it deals with nation-wide (actor level) deterrence (overt physical coercion).

Digital surveillance – when exposed – is thus to be theoretically handled as overt repres-

sion. Complementing other more direct repressive actions, digital surveillance enables, so I argue, large parts of the deterrent effect. In sum, the full range of overt physical repression contributes to authoritarian control, as in Svolik (2012), by removing the behavioural threat either directly or indirectly by deterring others from challenging the status quo.

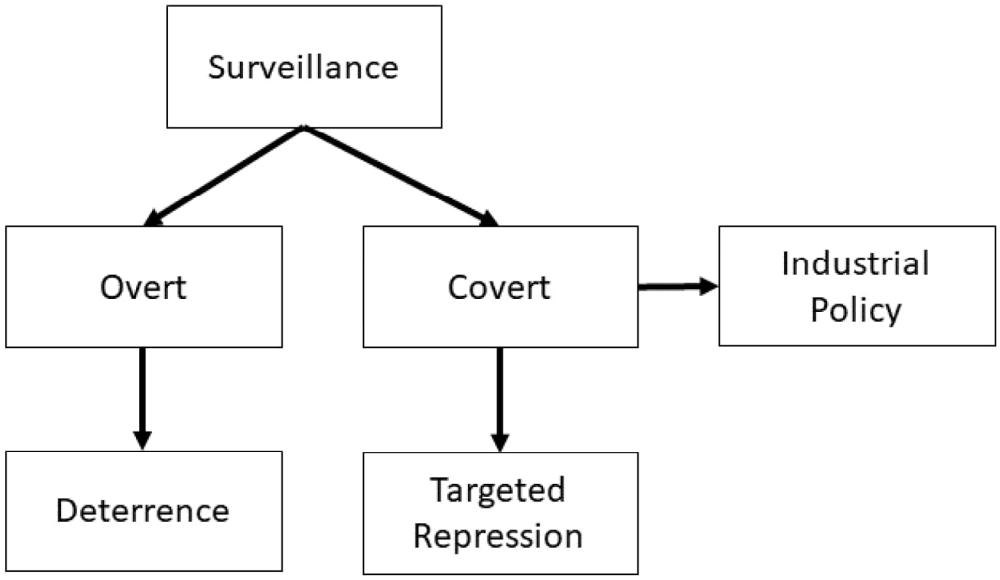
Covert physical coercion

The second column of Table 1.1 (cells 2, 6, 10) deals with *covert* physical coercion by state and private actors. While the explicit visibility (or: performative nature (Earl et al. 2022)) of repression can also serve a purpose, as described above, the non-performative nature of covert physical coercion also serves a function. Indeed, the techniques described here are more effective when they remain unobserved by their targets or the general public. Figure 1.2 highlights the functions of digital surveillance in a stylized manner. This section continues with a discussion of covert physical coercion, of which covert surveillance is an important part.

Large-scale (or mass) digital surveillance is both longstanding and well-documented. It serves the purpose of identifying subjects for targeted repression. Often, there are specific groups that are being targeted, for example Muslims in the aftermath of the terrorist attacks on 9/11 in the U.S. (Stoycheff et al. 2019), political dissidents in China (Pan 2020; Xu 2021), or ethnic minorities in China (Chin and Lin 2022) and Turkey (Çelik 2013). While the goal is usually to surveil specific groups, from a technical perspective there are different scales at which surveillance is done. First, there is all-encompassing mass surveillance done at the very root of the internet, at the level of the internet Service Provider (ISP). To maintain this level of control, regimes have to spend significant resources and build a long-term infrastructure. Some argue that China is really the only country that operates with a Leninist-like bureaucratic apparatus build for operations like this (Pei 2024). However, Russia with its SORM-program has similar access to all telecommunications like calls and internet traffic, with perhaps slightly fewer capabilities in analyzing the resulting data.[9] Other countries with similar access are Syria (Gohdes 2020), many post-Soviet states (Roberts 2018; Earl et al. 2022) or the U.S. and Britain (Earl et al. 2022).

[9] https://www.hrw.org/news/2016/07/12/russia-big-brother-law-harms-security-rights

Figure 1.2: Functions and externalities of digital surveillance



Beyond outright access to *all* telecommunications, there are also other techniques for mass surveillance – but more limited with respect to access and continuity. Kazakhstan, for example, implemented a Man-in-the-middle (MITM) attack in 2019.[10] With this technique, citizens are forced to install a custom root certificate which allows the government to decrypt Hypertext Transfer Protocol Secure (HTTPS) internet traffic. In other words, the government was able to inspect intercepted internet traffic that is usually end-to-end encrypted and thus private. This attack targeted mostly social media and communications running through Kazakhstan's largest internet Service Provider and was active for 21 days (Raman et al. 2020).

Of course, monitoring of social media is a routine for many police agencies and secret services world-wide. This is often done by lower-level state actors like police agencies in the wake of current events, or data is funneled down to local police offices from the national level (Xu 2021). Documented examples are not scarce and include instances of arrests of journalists and activists in Saudi Arabia (Pan and Siegel 2020), the U.S. (Borradaile et al. 2020), China (Roberts 2018), and practically most of the rest of the world, but to a greatly varying extent.

[10] https://www.f5.com/labs/articles/threat-intelligence/kazakhstan-attempts-to-mitm-itscitizens

Digital surveillance also has economic implications. A burgeoning literature examines the political economy of surveillance states, with a particular focus on the role of data for and beyond political repression. But what is it that makes data so valuable? First, Jones and Tonetti (2020) argue for the *non-rivalry* of data as they can be used multiple times by multiple firms. Hence, so the argument goes, data offer the potential for increasing economic returns corresponding to their economic exploitation. In this vein, Beraja et al. (2023c) argue that data are shareable between firms when it comes to the development of innovations in the realm of artificial intelligence (AI). Additionally, firms that receive government data as a form of a subsidy included in government contracts, in the example of China, are not only more effective in developing facial recognition AI technology, but *also* in developing more commercial AI applications. This *spill-over* is documented by Beraja et al. (2023b), who emphasize the mutually reinforcing relationship between AI development and authoritarian institutions. China's authoritarian institutions and politics do not deter it from being the world leader regarding AI frontier technology – an empirical fact that would not have been predicted by the main-stream literature on economic growth or scientific development. In fact, a consensus seems to exist that states the opposite: it is *democracy* that causes growth (Acemoglu and Robinson 2012; Acemoglu et al. 2019), and technological development should also be linked to democratic institutions (Aghion et al. 2007; Wang et al. 2021). This view receives some resistance (for example, Gao et al. (2017)), but is generally accepted. Here, however, there seems to be exception. As Beraja et al. (2023b) argue, China's innovation in AI capabilities have not risen despite, but *because of* China's autocratic politics. Due to the economics of data mentioned above and (i) the nature of authoritarian institutions and the ability to undermine privacy rights and other civil liberties, (ii) the lack of checks and balances enabling large-scale industrial policies unrivaled in democratic market economies, and (iii) large domestic demand for security and commercial AI applications, China leads in facial recognition AI technology development. Importantly, they argue, "that continued frontier innovation and economic development in China may not be associated with more inclusive political institutions; rather, such innovation may further entrench the autocratic regime" (Beraja et al. 2023b, p. 1396). This constitutes a vicious cycle between the distortion of the direction of innovation in AI towards autocratic use cases, increasingly authoritarian gov-

ernments, and thus even more demand for surveillance technology (Acemoglu and Johnson 2023).

Data come in many forms besides pictures of faces which can be used for facial recognition models. GPS location data, for example, can be used to predict protests and identify protesters[11] or Twitter meta-data can be used to predict protests before they actually happen (Bahrami et al. 2018). Collecting different types of data allows for different use cases, and in general, more data is often better. Accordingly, chapter 3 contains a study that experimentally examines support for *digital governance solutions* (DGS), for which data play a crucial role. DGS are defined as "systems that use digital technologies, artificial intelligence and big data to enforce existing legal rules and provide various services to the population, but also hold the potential for authoritarian surveillance and control." Here, again, it is important to emphasize the potential dual use of data: (i) they power DGS that *channel* behaviour, similar to social credit systems, and (ii), can also be used somewhere else, for example, as subsidies in the development of machine learning tools. The same data can be used multiple times and for multiple use cases. Chapter 3 will be introduced more thoroughly below, in conjunction with the literature on channeling (i.e., cells 4, 8, and 12 in Table 1.1).

Chapter 2 also engages with this literature. Within the chapter, the literature on innovation with respect to institutions is reviewed. Based on this literature a formal model is developed that takes into account the particularities of data as a commodity that can be used as an input by the R&D sector. Within the model, the trade-off between the negative effect of surveillance on research and creativity, and the positive effect of the availability of large amounts of data are the main points of emphasis. The model predicts that – on average – the effect of authoritarian institutions on innovation remain negative. However, in fields such as artificial intelligence where large amounts of data are important, authoritarian surveillance states can outperform competitive democracies. The model is consistent with empirical findings, for example Beraja et al. (2023b), and contributes in explaining how this anomaly emerges.

As laid out in more detail in the chapter, this study contributes to multiple disciplines and

[11] https://www.bbc.com/news/world-asia-china-61793149 https://privacyinternational.org/explainer/4503/how-police-can-determine-your-location-protest

strands of literature. First, the study engages in the literature on growth by presenting an endogenous growth model that speaks to Jones and Tonetti (2020); Cong et al. (2021, 2022) and Beraja et al. (2023c). By introducing a myopic government that produces governmental data as a by-product of surveillance and assuming that households derive positive utility from surveillance in the form of perceived security, the study advances models in this field. Consequently, the model is able to represent governmental as well as private data in the economy, which adds to external validity and realism of the model. Second, contrasting Beraja et al. (2023c), the government is characterized as self-interested in its attempt to secure power. This emphasizes the political perspective inherent in autocratic governance, even in economic policies and processes (Egorov and Sonin 2024). This has also theoretic implications for the R&D sector further discussed in the chapter, to which the reader is referred. Finally, the potential productivity impact of governmental surveillance relative to output on the growth rate of algorithms is considered transitory and is not seen to have long-run effects on the growth rate and labor market allocations. Despite creativity being reduced by more surveillance (Karpa et al. 2022), new data are generated, thereby increasing output via new algorithms that, in turn, contribute to productivity in the R&D sector. It is shown that, in the long run, both effects cancel each other out, as aggregate governmental surveillance grows at the same rate as aggregate output. In sum, chapter 2 focuses on the economic implications of exploiting the data generated by – initially politically motivated – surveillance. In the study presented, as in the studies elaborated on in this section, the mutually beneficial role of authoritarian institutions and innovation in AI on the one hand, and private and state actors on the other hand, are emphasized.

Not only can AI security technology be used to police political unrest in the domestic sphere, but this technology is also exported to other countries who use it for political repression, too. Crucially, whether or not a technology is used for political repression – or bought in the first place – depends on the institutional safeguards in place. Beraja et al. (2023a) report that only weak democracies and autocracies institutionally deteriorate after having bought AI tech during years of increased political unrest, whereas mature democracies seem to resist this development, arguably due to existing checks and balances. A similar development – as in private and state interests aligning – is documented for the case of the

development of tech giants in the U.S. under "surveillance capitalism" (Zuboff 2019), forming a new American "surveillance state" (Tau 2024). Zuboff (2019) documents how Facebook and Google's rise is directly connected to increasing security interests of the U.S. government after 9/11, neo-liberal skepticism towards regulation in general, and the entrenchment of the Obama administration with the tech giants. In essence, the Obama campaign used targeted political advertisement which built upon big techs surveillance capabilities, which dramatically enhanced the efficiency of the campaign. This has been described as "computational politics" (Tufekci 2014).

Overt Channeling

Cells 3, 7, and 11 describe overt channeling. Channeling itself refers to incentivizing behaviour preferred by the repressor with measures such as rules, policies, laws or specific governance structures. Channeling does not only apply to modifying behaviour, but to voicing, too. Earl et al. (2022) explicitly note that channeling does not entail direct pressure or force, and in that, is more of an indirect approach of limiting behavioural threats to the regime. In essence, channelling tries to remove the behavioural threat by steering it in another direction, or, by redirecting behaviour into less challenging or threatening forms. Channeling is conceptually close to *Nudging*, in which the decision-making process is influenced by altering the choice architecture (e.g., framing, defaults) (Thaler and Sunstein 2008). The main difference is that nudging is concerned with preserving the freedom of choice (Thaler and Sunstein 2003). While both nudging and channeling aim to influence behaviour, nudging does so by subtly altering choices within an existing environment, whereas channeling reshapes the environment to guide behaviour more directly.

Overt channeling is done by grievance or complaint platforms by different actors whose level of affiliation with the state varies. In general, government websites that enable citizens to lodge complaints are common and can benefit regimes by exposing grassroots conflicts, potentially demobilizing citizens who might otherwise have been more active. First, there are *national* initiatives where authorities channel public expression by providing forums for complaints and commentary, which subtly restricts the manner of dissent while appearing to encourage openness. For example, both the Obama and Trump administrations in the

United States offered petitioning platforms on their White House websites (Earl et al. 2022). In Russia, "letters to the president" are an important institution for channeling citizen complaints upward to state officials. As in the U.S., citizens can use the Kremlin's Web site to report a problem or express a concern.[12] Government officials can respond accordingly, contributing to the regime's (more specifically, the president's) legitimacy.

Lower-level government officials also utilize official complaint platforms to manage dissent. China, for example, offers complaint platforms on regional, city, and often at even county level (Earl et al. 2022). By doing so, the Chinese government channels dissent and discontent into controlled bays. Moreover, wherever necessary, authorities can react to the grievances in order to preempt collective action and protests, which is the priority in China (King et al. 2013; Chen et al. 2016). The potential for larger protests or the risk of exposing local failures to higher authorities can force local officials to respond. Local Chinese officials are skilled at blocking grassroots grievances from escalating to higher levels (Pan and Chen 2018), not least because of career considerations (Libman and Rochlitz 2019).

Covert Channeling

Covert channeling in cells 4, 8, and 12 in the last column of Table 1.1 refers to overly broad incentives or controls that influence contention. These incentives or controls can extend their impact so widely that the effects on contention may become obscured to casual observers. Often, these institutions claim to be used for a specific purpose, while at the same time fulfilling some or many other functions.

One example is the legal framework surrounding the Russian internet. This framework has developed over the past decade, ostensibly to protect children and the broader Russian way of life, but while promoted as beneficial for the well-being of Russians, these restrictions also curb dissent along with various other types of online information and activity (Kravchenko 2019). Similar laws exist in Turkey,[13] China (Roberts 2018), or Uzbekistan.[14] These laws are officially to protect "morality" (Turkey), or "constitutional order" (Uzbek-

[12] http://en.letters.kremlin.ru/

[13] https://freedomhouse.org/country/turkey/freedom-net/2023

[14] https://freedomhouse.org/country/uzbekistan/freedom-net/2023

istan), but they also limit the expression of dissent online, as well as the coordination between opposition members and activists. This hidden functionality is the reason these are classified as *covert*, whereas other laws are clearly targeting (political) dissent and are thus classified in the first column of Table 1.1 under overt legal coercion. This distinction might be subtle, but it is meaningful.

Another example are social credit systems (SCS), which are developed and employed in China. Citizens have a social credit score and gain and lose points according to their behaviour. On the one hand, SCS help to close *institutional gaps*, for example overcoming notoriously low social trust in the Chinese society and economy, or policing tax evasion and petty crimes (Kostka 2019). On the other hand, they are used to channel speech and behaviour much more broadly into pro-governmental directions. Government-defined pro-social behaviour is incentivized and anti-social behaviour is punished. Because of the usefulness of the system for increasing social trust these systems enjoy a high degree of public approval (Kostka 2019).[15] The politically motivated steering of behaviour is hidden behind the fact that large parts of social credit systems are operating around regulating economic relations and behaviours (Liang et al. 2018). Due to the tightly controlled media landscape and information environment in China, citizens might simply not be aware of the repressive potential of these systems (Xu et al. 2022; Kostka et al. 2023).

Chapter 3 reflects on this literature and empirically examines the transferability of systems, such as the social credit systems in China, to other contexts. The study presented in chapter 3 uses data from two surveys collected in (i) Russia and (ii) Estonia, Germany, the United States, and Turkey, respectively. The surveys contained an experimental setup, in which the question *how information about the repressive potential of a digital governance solution affects approval for the system* is investigated. 70% of respondents in Russia approve of the introduction of a digital governance solution that increases bureaucratic efficiency and contains some punitive legal capabilities, a number slightly lower than the 80% identified by (Kostka 2019) for China. Once Russian citizens are specifically reminded that the system can be used to identify and prosecute political dissent, support drops by 24% to

[15] Although survey responses in this context always have to be taken with caution (Robinson and Tannenberg 2019).

46%. Citizens who generally approve of the government and consume information mainly from state-controlled media are more likely to favor introducing a digital governance solution, linking the results to the literature on media effects in autocracies (Adena et al. 2015; Peisakhin and Rozenas 2018; Enikolopov et al. 2022). The second study replicates the results from Russia for Estonia, Germany, the United States, and Turkey mostly, allowing the hypothesis that political regime type has a significant effect on approval to be rejected. Approval rates are highest in Estonia (74.8%), followed by Turkey (66.7%), Germany (65.9%), and the U.S. (45.9%). Once citizens are reminded that the system can be used to prosecute political dissent, support drops significantly in all of the countries (with the exception of the US, where the drop is also visible, but not statistically significant). In all five countries, satisfaction with public services significantly increases approval of a DGS, refuting the hypothesis that gaps in the quality of public services create a demand for digital governance solutions, unlike in China (Kostka 2019).

The results hint at several important implications. First, as in the Chinese context (Xu et al. 2022), knowledge about the repressive potential of a digital governance solution can significantly reduce public support for the system. This is important, as even in autocracies, public support remains essential for the proper functioning of a DGS. Second, regime legitimacy matters. If citizens trust their government and receive information about the world mainly through state-approved sources, they are significantly more likely to support the introduction of a digital governance solution by the state. This finding is related to the recent literature on "informational autocrats" (Guriev and Treisman 2019, 2020, 2022), in which the role of the media is essential in convincing the public of the autocrats competence, but as we show, also of the effectiveness of policies or new institutions. These then enjoy more approval, which in turn makes them more effective because of the economies of scale of data and other properties of data, as discussed above. The theory of informational autocrats will be discussed in length in the respective chapter on information channeling (right side of Table 1.2) below. Third, contrary to expectations, frustration with the quality of public services does not increase support for a DGS. On the contrary, citizens who are satisfied with the state and the quality of public services are also more supportive of introducing a DGS. Finally, no conclusive evidence that regime type matters for the approval of digital gover-

nance solutions is found. Approval rates were highest in Estonia (a democracy), followed by Russia, Turkey, and Germany (an autocracy, a hybrid regime, and a democracy), with the U.S. (another democracy) being somewhat of an outlier with much lower approval than in the other countries. In this regard, the study thus introduces a comparative perspective to the literature which has been scarce with limited recent exceptions (for example, Rabe and Kostka (2024)).

This study adds to the ongoing debate on democratic backsliding, the resilience of authoritarian regimes, and broader issues of institutional development. The governance technologies powered by big data discussed here could significantly change perceptions of government surveillance and political control. Over the past few years, various social credit systems have been piloted in different regions of China, demonstrating this shift (Kostka 2019; Kostka and Antoine 2020; Strittmatter 2020; Li and Kostka 2022; Liu 2022). The COVID-19 pandemic, coupled with advancements in artificial intelligence, big data, and facial recognition, has further enhanced the effectiveness of digital and algorithmic governance (Knight and Creemers 2021; Katzenbach and Ulbricht 2019; Feldstein 2021). Surveillance data is now being leveraged to support Chinese firms, bolster domestic industry, and reinforce authoritarian institutions (Beraja et al. 2023b). As a result, China has become a global leader in digital governance, big data, and facial recognition technologies (Feldstein 2023), offering these technologies in integrated packages, such as "smart cities" (Große-Bley and Kostka 2021). Additional implications are explored in chapter 3.

As noted earlier, China is increasingly exporting these technologies, especially to other authoritarian regimes (Beraja et al. 2023a). When an authoritarian government faces domestic political protests, there is a particularly high likelihood that it will import a digital governance system (DGS) with policing features from China (Beraja et al. 2023a). These surveillance technologies are often bundled with infrastructure and other technologies as part of global initiatives like the "Belt and Road Initiative". Just as trade with democratic countries can encourage democratization (Tabellini and Magistretti 2022), trade with China can reinforce authoritarian regimes by exporting digital governance solutions. Russia is especially susceptible, but many other nations, particularly in the Global South, are also engaged in trade and infrastructure partnerships with China. This trend is further explored in section

Table 1.2: Digital repression expanding on traditional processes

	Information control			
	Information coercion (i.e., controlling information by limiting access or content)		Information channeling (i.e., influencing production and consumption of information)	
	Overt	Covert	Overt	Covert
State agents tightly coupled with national political officials	Limited national Internet connectivity (e.g., North Korea), temporary Internet blackouts, and state-based content filtering -1-	National content filtering where that filtering is not clear to users (e.g., returning 404 errors for filtered material) -2-	Government accounts posting distracting information and/or flooding online spaces or hashtags with irrelevant material -3-	Government disinformation and/or misrepresentations that influence contention -4-
State agents loosely connected with national political officials	Regional Internet blackouts and/or content filtering -5-	Regional content filtering where that filtering is not clear to users -6-	Local government or police information posting distracting information and/or flooding online spaces or hashtags with irrelevant material -7-	Local government and/or police disinformation and/or misrepresentations that influence contention -8-
Private agents	Deplatforming activists or organizations and/or moderating activist or organizational content -9-	Down-ranking, search filtering, shadow banning, throttling the spread of, or otherwise making protest-related material more obscure -10-	Private actors posting distracting information and/or flooding online spaces or hashtags with irrelevant material -11-	Private disinformation and/or misrepresentations that influence contention -12-

Note: Source: Earl et al. (2022). The Table was recreated without introducing changes to the original Table.

1.2.1.

1.1.2 Digital repression expanding on traditional processes

Digital repression can largely be understood through traditional repressive processes like censorship and state-media systems. However, when scaled up with digital tools, these forms of control influence behaviour in new ways and may gain new qualities. Extensive research on authoritarian politics highlights the importance of information control and propaganda (Zhuravskaya et al. 2020). The seminal work of Guriev and Treisman (2019, 2020, 2022) emphasizes the role of information, media, and persuasion as the central means of control. This very influential theory has been discussed widely and serves as a reference point in large parts of the literature on digital authoritarianism. Hence, I provide a short review on the *informational autocracy* proposed by Guriev and Treisman (2019), before following the structure of Earl et al. (2022).

The theory of informational autocracy by Guriev and Treisman (2019) explains how contemporary authoritarian regimes maintain power predominantly through sophisticated information control rather than relying extensively on overt repression. These regimes fo-

cus on shaping public perception and manipulating media narratives to cultivate a facade of legitimacy, stability, and most importantly, economic success. Informational autocracies utilize a combination of propaganda, censorship, and strategic dissemination of biased information to influence public opinion. By controlling the flow of information, they can present a curated version of reality that downplays regime shortcomings and emphasizes successes. This strategic manipulation allows autocrats to maintain control with a lower profile, reducing the likelihood of international condemnation because of violent repression, which could lead to isolation. A key aspect of this theory is that by shaping the information landscape, informational autocracies can avoid the backlash and instability that often accompany overt repression, as highlighted above. In this, informational autocrats differ from totalitarian rulers of the 20th century, that made much more use of overt and bloody violence. Overall, the theory of Guriev and Treisman (2019) highlights the evolution of authoritarian strategies in the digital age, emphasizing the critical role of information control in sustaining autocratic rule while minimizing the visible use of force and coercion. Informational autocracies are related to what others called electoral autocracies (Morse 2012) or hybrid regimes (Geddes et al. 2014).

Because of the importance of information control to contemporary autocrats, Earl et al. (2022) develop a typology that captures digital repression that *expands* on traditional processes, as shown in Table 1.2. They extend the typology shown earlier in Table 1.1 by mapping strategies for controlling information (e.g., censorship) and channeling information (e.g., disinformation campaigns) through the introduction of Table 1.2. Digital-era information control, while rooted in pre-digital practices, has evolved significantly. Despite some scholars noting continuities between traditional and modern censorship methods, there's clear evidence that digital technologies have transformed both the goals and reach of information control strategies (Feldstein 2021). According to Earl et al. (2022), historically, censorship primarily targeted knowledge suppression through tactics like book bans and media ownership control. This approach shaped public discourse by restricting unfavorable information and allowing those in power to craft narratives. While digital censorship incorporates these methods, it goes further by curtailing individual expression, and hence limiting people's ability to communicate and be visible online. This expansion has profound implications for social

coordination and collective action. Modern information control and channeling techniques now have the power to directly shape behaviour and alter perceptions of the information ecosystem itself.

Overt information coercion

In cells 1,5, and 9 of Table 1.2, overt information coercion is mapped. In general, this section is about limiting access to information on the internet, or access to the internet itself. On a technical level, restricting access is done by internet shutdowns or by introducing "firewalls". In the context of Roberts (2018), this is equivalent to *friction*, that is, censoring information by increasing the cost of it, thus making it less accessible. In this particular case, the added cost is introduced by "friction", which tries to metaphorically convey that the process of accessing information is more difficult, because it takes more time, technical skills, or money.

The first access restriction, internet shutdowns, can be done on the national or, more often, regional level. The literature shows that there are some events in which internet shutdowns are much more frequent, that is, during large protests and riots (Hassanpour 2014), and during elections (Lutscher et al. 2020; Garbe 2023). Apparently, the need to reduce the potential of collective action warrants the use of drastic measures like shutting down, or drastically throttling the speed of, the internet. By shutting the internet down, the possibility of communication between protesters and activists is substantially impeded (Gohdes 2020). Another reason for regimes to resort to shutting down the internet is contain the spread of information during critical events (Gohdes 2023). When regimes commit atrocities during civil wars or mass protests, there is a risk of (i) a backlash from the domestic population and (ii) international attention and its consequences, such as sanctions. Hence they try to minimize this danger by removing access to digital communication between citizens, but also between citizens and the outside world. Many scholars emphasize the inherent trade-off and resulting strategic implications that come with internet shutdowns: not only does the public loose its access to information and communication, but state actors also loose sight of anti-regime sentiment (Garbe 2023) and their ability to identify targets for individual-level repression. In fact, during shutdowns in Syria mass repression increased

and targeted repression decreased, precisely because the government blinded it's own forces by shutting the internet down in some regions (Gohdes 2020). This is also why nationwide shutdowns are very unlikely to be frequent events, and covert filtering of information is much better suited for long-term use, see section 1.1.2.

The second overt access restriction, firewalls or the filtering of content and websites, happens due to domestic laws and legal pressure that might compel companies to restrict access or delete content. For-profit considerations, or maintaining a "healthy" platform environment according to their standards, may also be reasons for companies to remove content or to deplatform activists or politicians. In some instances however, deplatforming activists occurred for political reasons, for example when a successful campaign account for Kamala Harris was deleted on the platform X owned by an outspoken billionaire supporting the opposing political candidate.[16] Additionally, because ISPs are increasingly state-owned in autocracies, data flows within and from the country to the global internet can be controlled by state actors (Keremoğlu et al. 2024). This development only contributes to further strengthening autocratic governance, with international coalitions between autocracies lending each other technological tools. Because of these ownership structures, large-scale censorship initiatives at the backbone of the internet are possible. In China, for example, government restrictions on content lead to a multilayered system of filtering from automated to human intervention (King et al. 2017; Roberts 2018; Pei 2024). Any search terms that could remotely hint towards the Tiananmen massacre in 1989 are blocked, such as "protest", "candle", or even numbers like 4, 8, or 35.[17] Some posts never appear, or are deleted shortly after being posted (King et al. 2014). Beyond filtering for sensitive terms in messaging and blogging services, approximately 13% of *all* social media posts are censored in China – often with the motive of stopping mobilization or any form of collective action (King et al. 2013). While the censorship efforts of China are certainly unrivaled in scale and scope, many other countries also engage in extensive information access restrictions (Freyburg and Garbe 2018; Stoycheff et al. 2020).

---

[16] https://www.thewrap.com/white-dudes-for-harris-x-account-suspended-elon-musk-trump/, or https://www.washingtonpost.com/opinions/2024/08/02/white-men-kamala-harris-x-twitter/

[17] https://chinadigitaltimes.net/2012/06/sensitive-words-the-tiananmen-edition/

Covert information coercion

From the dictator's perspective, covert censorship has several advantages. First, covert censorship avoids international backlash and isolation, and second, it avoids domestic backlash, including protests and circumvention attempts (Roberts 2020). Examples of less obvious censorship include denial of service attacks (DoS), manipulation of search results, or filtering and slowing access to information from certain sources. Denial of service attacks are a type of cyber attack in which traffic to a particular domain is overloaded, making the site unavailable to others. Reported cases include many authoritarian elections, where foreign media are attacked to censor negative coverage (Lutscher et al. 2020), and dissident or critical websites or newspapers, such as the Chinese version of the New York Times and Wikipedia (Gohdes 2023). In addition to blocking these sites in China, DoS attacks have temporarily taken them down entirely. While the inability to access an entire website may be very noticeable, it is often difficult to determine why a site is inaccessible, let alone who is responsible, which is why Earl et al. (2022) refers to DoS attacks as covert.

Search engines can be architecturally modified to serve regimes by systematically displaying some content while hiding others, thus helping to alter social realities (Jiang 2014). Kazakhstan's government used selective filtering to restrict access to political websites and posts, while increasing the availability of non-political ones (Anceschi 2015). Again, when content is made unavailable, it may not be clear to the user who receives a "404 Not Found" error why the content is inaccessible. In fact, it could be a broken link, a connection problem, human error, or any number of reasons other than government censorship.

Overt information channeling

Digital and social media have opened new avenues for repressive actions based on influence, using a technique termed as information channeling by Earl et al. (2022). Information channeling differs from the defensive, access-based form of digital repression discussed in the last two subsections 1.1.2 and 1.1.2. Instead of restricting or removing access to information, it focuses on redirecting or influencing attention and thus influencing behaviour. Informational autocrats, introduced above, primarily rely on these forms of repression, since distracting from content or overwhelming with other information is typically easier than blocking access

to information or posts entirely. Due to the less intrusive nature, it is also potentially less provoking with respect to backlashes (Acquisti et al. 2022). Unlike traditional propaganda, which primarily aims to alter attitudes, opinions, and beliefs, information channeling focuses on influencing behaviour. Although downstream effects may influence beliefs and second-order beliefs (beliefs about what others believe), indoctrination is not the primary goal of information channeling.[18] Repressors make preferred information easier or more appealing to access, similar to using incentives to encourage desired behaviours in traditional channeling, as shown in Table 1.1. By changing the information environment, citizens might not encounter information on issues they would have supported or activities, like protests, they might have participated in. Shaping the overall information environment is similar to how broader laws and practices alter the context for decisions and behaviours related to protests in traditional channeling.

More specifically, overt forms of information channeling (cells 3, 7, and 11 in Table 1.2) can occur in the form of state actors producing and spreading information in order to redirect attention. One example is the manipulation of the information environment with "bots" or click-workers. In many countries, the domestic information environment is spammed with pro-regime posts and messages. This is also known as *flooding*, as in flooding the information space with distracting information to make real information less accessible (Roberts 2018). For instance in the Gulf countries and Saudi Arabia, where government actors contributed to mass production of online statements via automated accounts on Twitter (Leber and Abrahams 2019). This manipulation through bots in the Gulf is also aimed at securing organic participation from supportive citizens, mobilizing them for pro-regime purposes. In Kazakhstan, so called "Nurbots" flooded the internet with praise for the first president Nursultan Nazarbayev.[19] Besides pro-regime or pro-president sentiment, there is also the approach of non-political distraction in Kazakhstan. Social networks are filled with sports personalities, celebrities, and pop stars whose accounts are managed by state-run media companies (Anceschi 2015). Crucially, this approach pursues distraction, not convincing,

[18] Indoctrination is still a part of the propaganda efforts, but here the focus is on the direct change of behaviour.

[19] https://factcheck.kz/en/glavnoe-en/the-nation-leaders-farm-who-manages-the-bot-and-troll-networks-in-kazakhstan/

and is thus distinct from traditional propaganda.

Covert information channeling

Covert information channeling (cells 4,8, and 12 in Table 1.2) involves spreading disinformation to manipulate discussions or hide the true identity of the source, regardless of the information's accuracy. These efforts are targeted at shifting the focus, altering beliefs, expressions, or behaviours, and even impacting perceptions of others' opinions and actions (second-order beliefs). Perhaps the least clear distinction between categories in the typology Earl et al. (2022) is between overt and covert information channeling. They classify channeling efforts as covert contingent on the attempt to deceive, not whether citizens actually get aware of the actors strategies.

For example, after the massive and violent protests in Kazakhstan in 2022, the government spread disinformation to conceal (or partially justify) police violence and torture (Laruelle and Kudaibergen 2024). To this day, the exact events and motives of the protesters remain unclear to many, as the regime continues to withhold information from the investigative process while ambiguous and conflicting narratives circulate. All of this happened after the internet was shut down during the protests, which only added to the chaos and confusion, and perhaps allowed the government to untruthfully claim the "invasion of 20,000 foreign terrorists".[20] In Russia, government-affiliated actors spread false information about the main opposition leader Alexei Navalny.[21] Tactics similar to those described in section 1.1.2 are used by click-workers in China who spread disruptive information during critical political events by posing as ordinary citizens (King et al. 2017). Jones (2022) shows how technologically savvy regimes in the Arab Gulf region use automated supply chains for disinformation. These false narratives take hold domestically and are later exported globally.

In addition to targeting domestic populations, there are also influence campaigns that specifically target other populations abroad. Russian bots and user accounts have interfered in both the UK's Brexit referendum and the 2016 U.S. presidential election (Alizadeh et al. 2020; Earl et al. 2021). Recorded instances also include Russian misinformation campaigns

[20] https://www.rferl.org/a/kazakhstan-january-cheryl-reed-journalists/32217364.html

[21] https://citizenlab.ca/2017/05/tainted-leaks-disinformation-phish/

about the "Black Lives Matter" movement in an attempt to polarize the public and create conflict (Earl et al. 2022). Russian efforts with disinformation campaigns have culminated in what some have called "post-truth" or "pseudo-realities" (Jones 2022).

In the context of the full-scale Russian invasion of Ukraine in early 2022, Russian campaigns were launched to sell an alternative narrative. In fact, Russia is so successful and dominant in producing an alternate reality that many Russian-originated narratives have been adopted around the globe, for example in China (Hanley et al. 2024) or in English-speaking political discussions on the internet platform Reddit (Hanley et al. 2023). Chapter 4 picks up on this development and examines how Russian media influence citizens abroad.[22] First, the chapter examines the reasons why citizens choose to consume media from Russia. In Kazakhstan, large parts of the population *speak* Russian, while *identifying* as Kazakh. This study argues that it is the ability to speak Russian that opens the door to consuming media *in* Russian, but also media *from* Russia. Second, using original data from a survey experiment in Kazakhstan, the study shows how attitudes toward political issues change and correlate with Russian media consumption. Most importantly, attitudes toward sanctions against Russia are also affected. In a next step, the relationship between Russian speakers in regions and increased trade with Russia in those same regions is approached econometrically. The central finding of the study is that trade with sanctioned products in regions with more Russian speakers increases, as compared to regions with fewer Russian speakers.

The paper contributes to the growing body of research on how media in authoritarian regimes can heavily shape public beliefs and behaviors (Enikolopov et al. 2011; Adena et al. 2015; Anceschi 2015; Zhuravskaya et al. 2020; Pan et al. 2022). Exporting media and narratives can also influence populations abroad. Over the past decade, Russian disinformation has influenced events like Brexit and the 2016 U.S. presidential election (Alizadeh et al. 2020; Earl et al. 2021), and has spread divisive narratives around movements like Black Lives Matter (Earl et al. 2022). Following Russia's 2022 invasion of Ukraine, similar strategies were employed to sway global opinion against Western support for Ukraine and to undermine sanctions. This study is the first to show how these tactics are successfully altering public

[22] A comprehensive review of the literature on media effects can be found in chapter 4 and is omitted here to avoid redundancy.

perceptions and easing concerns about sanctions evasion. The research demonstrates how the Russian diaspora has become both a target and a tool in Putin's effort to rebuild the Russian empire. Language is identified as a key factor that makes populations vulnerable to Russian propaganda. The findings align with existing studies on media effects in autocracies, emphasizing the role of language and online media. The spread of Russian propaganda to other contexts suggests that factors beyond identity and language may also be at play, highlighting the importance of studying these phenomena and understanding how propaganda narratives spread from context to context. The following section discusses these findings in a broader context and concludes with an outlook for further research.

1.2 Discussion and Outlook

One of the foci of this dissertation is surveillance technology. In chapter 2, my co-authors and I presented a model that explains the rise of contemporary surveillance states by showing that the negative externalities of autocratic institutions can be outweighed by their positive impact on innovation in this specific area – thus opening a pocket of innovation. We acknowledged the crucial role of data available to researchers, which is echoed in recent literature. Acemoglu and Johnson (2023, p.355), for example, write that the "one advantage" China has in developing AI tools are "much larger quantities of data". Moreover, the beneficial relationship between authoritarian institutions and innovation in AI is not unidirectional, but mutually reinforcing. Technological progress is shaped by autocratic leaders and their demand for repressive technology (Acemoglu and Johnson 2023; Beraja et al. 2023a,b). In China, the direction of innovation is shaped specifically by government issued contracts, which come with demand for surveillance technology, funding, and include the most important ingredient: data (Beraja et al. 2023c). These companies can't help but adapt to this enormously incentivized environment, shifting their focus to facial recognition and other dual-use surveillance technologies. The more autocrats demand these technologies for the prevention and suppression of dissent, the greater the increase in the quality and quantity of the development of these tools can be expected. As a result, more repressive technology is demanded because the increased functionality and supply is so attractive to governments worldwide. This vicious cycle results in an increasing digital authoritarian grip that is driven

by it's internal logic. The loss of civil liberties does not generate the public resistance one might expect. Clearly, the expectation that democracies would shy away from demanding and using this technology is wrong and anachronistic at this point (Zuboff 2019; Earl et al. 2022).

In the following chapter, 3, my co-author and I show that public approval for digital governance with "negative [...] externalities (i.e., lost civil liberties and political rights)" (Beraja et al. 2023a, p.23) is substantial across multiple institutional contexts. While approval diminishes when DGS are explicitly framed as repression technology, support is still around or above 50% in most countries in our sample. The public seems to be unaware of, or unwilling to adapt to, the complex risks of dual surveillance technologies, in this case facial recognition-based digital government. Can we hope, then, that "responsible citizens" will act individually to protect their liberties, culminating in sufficient organization of civil society to counterbalance the growing capacity for repression? Or could the balance of power between states and civil society be left in favor of despotic states? If this delicate balance of power – "the narrow corridor" – is to be maintained, civil society must certainly keep moving along with technological development (Acemoglu and Robinson 2019). In this chapter, we also discuss how a technological "ratchet effect" could come into play – allowing a country to move from softer to harder authoritarianism, but preventing it from eventually switching back and democratizing again. In hybrid regimes and countries that oscillate between more and less authoritarian forms of government, such as Turkey or Hungary, but also the United States, a well-functioning surveillance infrastructure could play a key role in stabilizing authoritarianism after a period of democratic backsliding. This thesis is thus relevant beyond contemporary autocracies, because in the long run many democracies are at risk of leaving the narrow corridor – and not being able to return.

These digital governance solutions, which include big data analytics tools, may not offer the same level of liberation potential for civil society as the internet and digital communications did (Diamond 2015; Gohdes 2023). Instead, it seems that it is mainly state power that increases over time, increasing the likelihood of more "despotic states" such as China (Acemoglu and Robinson 2019). Some argue for trade regulation along the lines of other dual-use technologies (military-civilian) (Beraja et al. 2023a), a recommendation that seems

appropriate in light of the discussion above. The EU has recently taken initiatives regarding consumer protection[23] and the regulation of facial recognition technology in public spaces.[24] Because of the points raised above – and in light of the rise of right-wing populism in Europe and other parts of the world – it seems even more appropriate to have safeguards in place that limit the power of the state. Democratic institutions help, at least in part, to prevent repression and to deal with protests and grievances by means other than repression (Beraja et al. 2023b).

Even if a simple ban on a technology suffices in one case, the developments of misinformation campaigns and consequent post-truth societies are much more delicate and complex, as has been shown in chapter 4. It is highly doubtful whether a top-down regulatory approach will be able to contain this development, especially in the face of many (even democratic) leaders abusing and further accelerating the development towards polarization and a multitude of conflicting realities. Germany's political landscape, for example, has been invaded by "Polarisierungsunternehmer" (Mau et al. 2023) – or *polarization entrepreneurs*. These are people who try to anchor themselves in the political discourse by adopting a strongly conflictual position, exploiting an increasingly volatile and susceptible media landscape in the 21st century. This also includes right-wing populist actors and anyone who manages to politicize a social issue in such a way that it contributes to the formation of camps and affective polarization (Wagner 2021).

Some argue that to counter this, a free and critical press is essential to hold liberal leaders accountable and separate them from populist and illiberal ones (Jones 2022). While a free press seems to be a necessary condition to curb false news, whether a free press alone suffices remains questionable. Some argue that this trend – fake news and alternative realities – marks a development from politics of inevitability to politics of eternity (Snyder 2018). The politics of inevitability is a sense that the future is just more of the present, that we know the laws of progress, that we have no choice, and thus no real action.[25] This inevitability of

[23] https://www.euronews.com/next/2024/03/07/here-are-the-changes-tech-companies-are-implementing-to-comply-with-eu-law

[24] https://www.amnesty.org/en/latest/news/2023/06/eu-european-parliament-adopts-ban-on-facial-recognition-but-leaves-migrants-refugees-and-asylum-seekers-at-risk/

[25] Similarly, Blühdorn (2013) argues that Western societies merely "simulate" democratic societies to keep

historic development proved to be wrong, however, as it became more and more evident that not everything will get better on it's own. The experience of polycrisis in the 21st century,[26] coupled with the growing disillusionment with the idea that individual freedoms would simply unfold and not need to be curtailed by these very crises (Blühdorn 2024), eventually made the narrative of inevitability collapse. The collapse of the politics of inevitability, as Snyder (2018) argues, ushers in another experience of time: the politics of eternity. Here, democratic pluralism and the right to choose are rejected in favor of an ordered, one-dimensional vision of the world, with a clear view of a (constructed) enemy. Whereas inevitability promised a better future for everyone, eternity places one nation at the centre of a cyclical story of victimhood. Many of the polarization entrepreneurs and right-wing populists follow this line of reasoning. Mau et al. (2023) describe how in politics and the public discourse around it, the *arenas of inequalities* get (strategically) confused such that strong emotions are triggered and the discussion tips towards partisan fights and ultimately to affective polarization. Instead of having substantial discussions on real policies, for example how to deal with poverty or climate change, these actors systematically answers question about one topic with answers from another. What "triggers" individuals is, for example, that *they* have to bear the costs of climate change, too, not only some abstract society. The unwillingness to pay these real economic costs can be channeled by populists into the plea for closed borders, climate change denial, general "anti-wokeness", and other stories of constructed victimhood. Depending on the context, an arbitrary enemy can be constructed to fit the narrative. In the case of Putin's Russia, Neo-Nazis residing in Kiev were constructed (Marples 2022), in the case of Donald Trump's campaigning it's an allegedly evil establishment (Enders and Uscinski 2021). Whether or not contemporary societies will be able to contain these post-democratic (Crouch 2019; Manow 2020), post-liberal (Blühdorn 2024), and post-truth (Jones 2022) developments – enabled by new forms of political repression – remains unclear. Blühdorn (2024) argues, that analogous to the climate crisis, clinging to the idea of containing this development may not only not be possible, but, conversely, is doomed to failure and anachronistic. Instead of

---

up with their normative ideas of democracy but that, ultimately, this democracy is performative, serving self-assurance.

[26] https://www.weforum.org/agenda/2023/03/polycrisis-adam-tooze-historian-explains/

wasting resources trying to avoid this next modernity, it might be more appropriate to think about how to live in it.

But how will our modernity look like? And how can we know whether contemporary societies will evolve more in one direction or the other? Will we see the development of the surveillance state model, or will the right-wing populist model be more compatible with the modern world? Different forms of digital repression seem to be substitutes and the exact trade-offs are contingent on a multitude of factors. The repressive capacity, technological infrastructure, and institutional context matter. Not all governments have the capacity to engage in sophisticated methods of digital repression and they tend to react to specific threats, as described by Davenport (2007) in the "law of coercive responsiveness". From the perspective of a repressive regime, the appropriate response to VPNs, for instance, might be information channeling instead of blocking of websites (Roberts 2018; Earl et al. 2022), whereas the response to protests might be violence and internet shutdowns, followed by spreading of rumours and "fake news"(Laruelle and Kudaibergen 2024). Where the technological capacity for repression is underdeveloped, governments can outsource the development of technology to private companies and buy ready-made solutions.

Precisely these technological capacities are sold by Chinese companies and, by extension, the Chinese state. Beraja et al. (2023a, p.23) argue buying AI-powered repression technology from China can lead to "negative global externalities (i.e., lost civil liberties and political rights)". Contributing to this dangerous development is that, in many countries, public approval for digital governance solutions is substantial, despite significant capabilities for (political) repression of these systems (see chapter 3). Assuming increasing proliferation of such systems world-wide and the lack of public resistance, does this imply a shift in power between states and civil society? Will tyranny favoured by technology and its ruthless exploitation prevail, as predicted by some?[27] It is unclear whether economic or political considerations by the Chinese leadership drive their exporting efforts and contended in the literature (Feldstein 2021). The consequences, however, can be problematic for civil societies all over the world. The argument of Davenport (2007) would be that governments *react*

[27] https://www.theatlantic.com/magazine/archive/2018/10/yuval-noah-harari-technology-tyranny/568330/

to behavioural threats, but as was discussed extensively, technological progress regarding repression technology may *create* new demand, too. Chapter 4 has shown that Russia is an exporter of right-wing and authoritarian ideology with profound influence on post-Soviet, but also other countries. This could perhaps further increase the demand for repression technology, for which China is a technological frontrunner and export leader. The following section will engage with the export of repression technology from China and consult data in order to add to the discussion of it, while ending the introduction to the chapters 2 to 5 with an outlook for further research.

1.2.1 Research outlook

Export of repressive capacities

In this section I develop the discussion about the export of authoritarianism introduced in section 1.1.1 further. To this end, I lay out avenues for further research that directly build upon findings from this dissertation's chapters and the corresponding literature. In particular, I suggest to combine data about trade between China and African countries, and data about the capacity for digital repression. I present a preliminary analysis and thereby hope to guide the reader from the state of the current affairs and the respective literature towards an outlook for further research. This data-driven approach reflects on the general procedure of the individual chapters, that is, to empirically investigate the questions at hand.

In Feldstein (2021), a Digital Repression Capacity Index is developed using four variables taken from the v-dem data.[28] This index encompasses capacities on internet filtering, shutting down the internet, cybersecurity staff, and staff and resources for regulating content on the internet.[29]. Combining the the data from v-dem and data about *exports from China*[30] to African countries yields Figure 1.3.

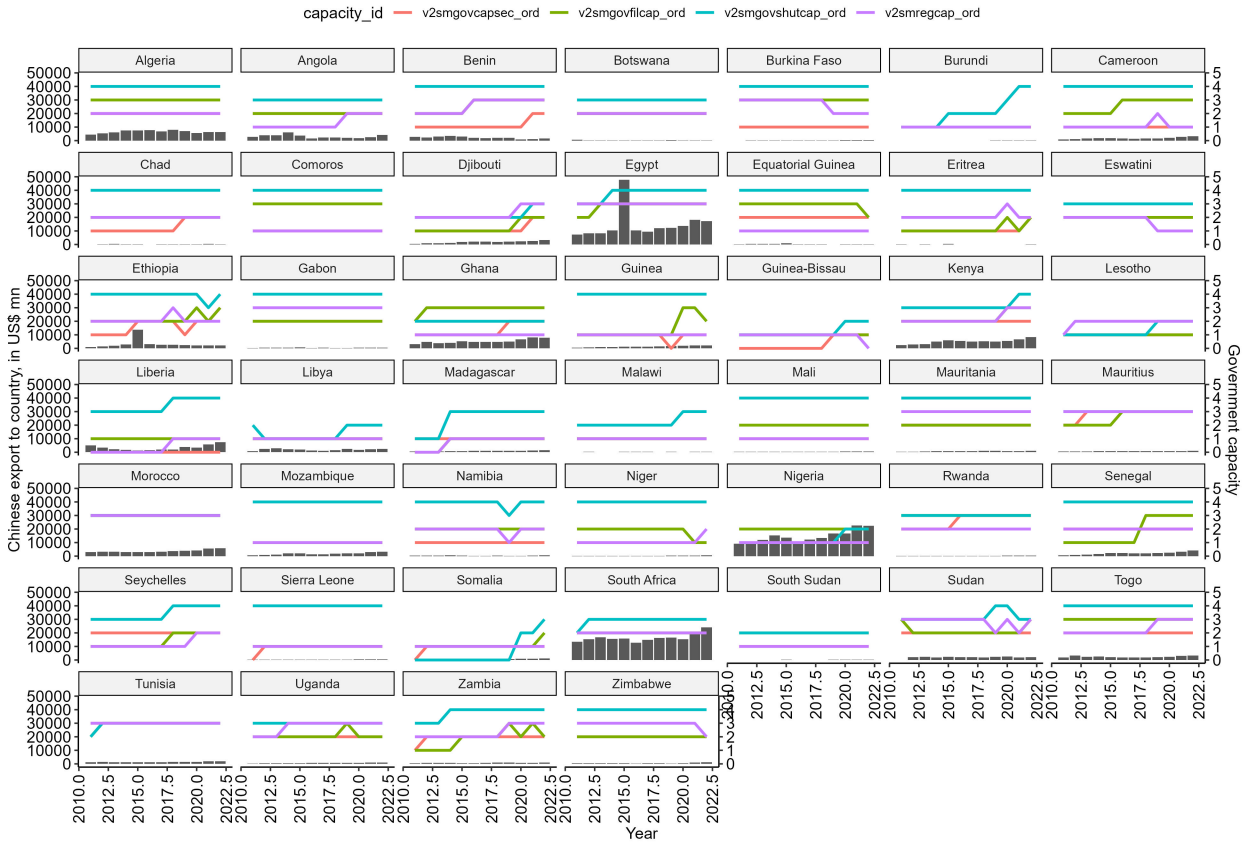
Beraja et al. (2023a) show that China exports AI surveillance technology by utilizing

[28] https://v-dem.net/data/the-v-dem-dataset/

[29] These capacities correspond to *information control* in Table 1.2. They do not capture surveillance and other *physical control* capacities. Moreover, cybersecurity captures generally relevant capacities which do not have to be limited to digital repression (Feldstein 2021).

[30] https://www.sais-cari.org/data-china-africa-trade

Figure 1.3: Import from China and repressive capacities



Note: *Source:* Data from v-dem and China Africa Research Initiative. Own presentation.

trade data on the level of individual deals. Here, I use trade data on the *overall volume of imports* from China. One implication of this choice is lower precision with respect to capturing specifically AI surveillance imports. At the same time, however, I am able to make an inference on the effect of the *overall impact* of importing at volume from China on digital repression capacities. Furthermore, in following Feldstein (2021) and choosing the subjective expert measures of repression *capacities* from v-dem, I distinguish this exercise from the perspective followed by Beraja et al. (2023a), that is, institutional development. Beraja et al. (2023a), instead, use the commonly used polity2 indicator,[31] which places a country on a scale between autocracy (-10) and democracy (10). The two approaches, while similar in scope, differ fundamentally in the scale of the perspective. Based on the literature,

[31] https://www.systemicpeace.org/inscrdata.html

Table 1.3: Trade with China and digital repression capacities

	Dependent variable:							
	Filtering (1)	Shutdown (2)	Cybersecurity (3)	Regulation (4)	Filtering (5)	Shutdown (6)	Cybersecurity (7)	Regulation (8)
GDP P CAP	0.0001***	0.0001***	0.00002	0.00004**	0.0001***	0.0001**	0.00001	0.00003
	(0.00002)	(0.00003)	(0.00002)	(0.00002)	(0.00003)	(0.00004)	(0.00002)	(0.00002)
trade vol CN	0.00001	0.00004*	-0.00001	-0.00001	0.00002***	0.00004**	-0.00001	-0.00001
	(0.00001)	(0.00002)	(0.00002)	(0.00001)	(0.00001)	(0.00001)	(0.00002)	(0.00001)
trade vol US	-0.00004	-0.0001*	0.00000	-0.00004*	0.00002	-0.0001*	0.0001	-0.00001
	(0.00003)	(0.0001)	(0.00004)	(0.00002)	(0.00004)	(0.0001)	(0.00005)	(0.00004)
Constant	2.324***	3.359***	1.855***	1.734***	2.688***	4.008***	2.150***	2.077***
	(0.110)	(0.159)	(0.073)	(0.082)	(0.195)	(0.236)	(0.152)	(0.178)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Robust SE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dep. var. lag 1 year	No	No	No	No	Yes	Yes	Yes	Yes

Note: *p<0.1; **p<0.05; ***p<0.01

I would expect to find positive correlations between the trade volume with China and the capacities to filter and shut down the internet. I might not expect the same effect for the capacities for cybersecurity and regulation, since these are based on human capital – and less on technology – which is clearly much less related to the trade volume.

Table 1.3 shows the results of regression models I ran in order to evaluate the effects trading with China has on digital repression capacities. My independent variables in the specifications (1) to (4) and (5) to (8), are the capacities to filter the internet, shut down the internet, cybersecurity, and regulate content on the internet, as suggested by Feldstein (2021). All models include fixed effects on the country-year level and standard errors that are robust to heteroscedasticity. Naturally, I control for GDP per capita taken from the world bank database.[32] I do also control for the trade volume with the United States, following the econometric approach of Beraja et al. (2023a). Finally, in the specifications (5) to (8), the dependent variables are lagged 1 year. This is done to account for the lag it takes for the trade to have an effect on capacities (and expert ratings).

It appears that the trade volume with China does have a positive effect on the capac-

[32] https://databank.worldbank.org/source/sustainable-development-goals-(sdgs)/Series/NY.GDP.PCAP.KD

ities to *shut down the internet*, as well as *filter the internet*. Whereas the coefficients for internet filtering are only corresponding to a significant effect after a year-long lag, the coefficients for internet shutdowns capacities being increased by importing from China are significant throughout the specifications. This is intuitively reasonable, since internet shutdown are comparatively easy to detect. Both cybersecurity and content regulation seem to be unaffected according to the present model specifications.[33] These results are in line with theoretical expectations and previous empirical findings, and they speak to the questions raised in the introduction to this dissertation. More specifically, it does look like increasing the trade volume with China leads to increasing digital repression capacities, at least to some extent. Indeed, it does appear that China exports repressive capacities to other (more or less vulnerable) states. These imports of repressive infrastructure will "likely also matter immensely for where, how, and how much of different forms of digital repression are deployed" (Earl et al. 2022, p.9). Some countries may "leapfrog" to certain levels of repression capacity without having to develop them on their own, giving rise to new modes of governance that are able to neglect this otherwise so crucial part. It is important to study these issues in order not to lose sight of developments in the balance of power between states and civil society. There may be a lock-in of an institutional arrangement with fewer civil liberties under authoritarian rule – introduced by technological change and those who drive it.

This dissertation has contributed in two main ways: first, it investigates the political economy of surveillance technology, and second, it measures the consequences of digital repression. As Earl et al. (2022) point out, many facets of digital repression have been under investigation of the broader social sciences – but consequences remain underexplored. Chapter 5 showed how digital repression changes communication behaviour and chapter 4 showed how state-led propaganda infects individuals and consequently affects their attitudes and economic behaviour. Trade with sanctioned products has been shown to be systematically higher among those affected by the propaganda in Kazakhstan. There are numerous implications and avenues for further research. This consumption of propaganda is likely to

[33] Of course, I am aware that these models are not yet fully developed and that the econometric approach itself lacks some robustness. Nevertheless, due to the overlap between results presented in this section, the literature, and the results of chapters 2 to 5, I here discuss possible implications.

have different "entering points" into people's life just as different outcomes. Which parts of the package that is sold by (mostly Russian) right-wing propagandists fits which context and individuals? This packages consists of many intertwined stories and narratives, mostly conspiracy theories and skepticism towards institutions. Consequently one could follow calls from the literature and investigate how media consumption influences vaccination attitudes (Jones 2022). Indeed, in Kazakhstan a sudden increase of measles cases has been reported (Akilbekova et al. 2024) – the hypothesis that this development correlates with, or is even causally connected to, vaccination hesitancy due to propaganda is both theoretically and practically important. But also in the U.S. vaccination hesitancy increases dramatically, with a very clear group driving this development: Donald Trump and his Republicans.[34]

The whole research agenda of *affective polarization* has identified the crucial role emotions play in shaping outcomes of propaganda and polarization entrepreneurs (Mau et al. 2023). In other areas of digital repression, this role remains underexplored. With respect to digital surveillance, for example, affective involvement possibly leads to effects of different sign. First, surveillance creates anger and leads to deterrence (Stoycheff 2022), but slowly increasing anti-regime sentiment, spurred precisely by this smoldering anger over time, might contribute to creating backlash (Roberts 2020; Earl et al. 2022). These affective and long-term effects of digital surveillance are not yet sufficiently explored (Büchi et al. 2022). Beyond deterrence, self-censorship, or backlash, digital repression may have other corrosive effects on civil society, such as the gradual loss of a sense of autonomy (Karpa et al. 2022). This loss of autonomy could further contribute to the erosion of the Kantian ideal of the autonomous subject, both of which (the idea and its erosion) characterize contemporary societies (Blühdorn 2024).

1.3 Bibliography

Acemoglu, D. and Johnson, S. (2023). *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity*. PublicAffairs.

Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2019). Democracy Does Cause Growth. *Journal of Political Economy*, 127(1):47–100.

[34] https://news.gallup.com/poll/648308/far-fewer-regard-childhood-vaccinations-important.aspx

Acemoglu, D. and Robinson, J. (2019). *The Narrow Corridor: States, Societies, and the Fate of Liberty*. Penguin Publishing Group.

Acemoglu, D. and Robinson, J. A. (2012). *Why Nations Fail: The Origins of Power, Prosperity and Poverty*. Crown Publishers, New York, 1st ed. edition.

Acquisti, A., Brandimarte, L., and Hancock, J. (2022). How privacy's past may shape its future. *Science*, 375(6578):270–272.

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., and Zhuravskaya, E. (2015). Radio and the Rise of the Nazis in Prewar Germany. *The Quarterly Journal of Economics*, 130(4):1885–1939.

Aghion, P., Alesina, A., and Trebbi, F. (2007). Democracy, Technology and Growth. *NBER Working Paper*, No. 13180.

Akilbekova, D., Yerdessov, S., and Gaipov, A. (2024). Urgent health update: rising measles cases in kazakhstan. *The Lancet Regional Health–Europe*, 37.

Alizadeh, M., Shapiro, J. N., Buntain, C., and Tucker, J. A. (2020). Content-based features predict social media influence operations. *Science Advances*, 6(30):eabb5824.

Anceschi, L. (2015). The persistence of media control under consolidated authoritarianism: Containing kazakhstan's digital media. *Demokratizatsiya: The Journal of Post-Soviet Democratization*, 23:277–295.

Bahrami, M., Findik, Y., Bozkaya, B., and Balcisoy, S. (2018). Twitter reveals: using twitter analytics to predict public protests. *arXiv preprint arXiv:1805.00358*.

Beraja, M., Kao, A., Yang, D., and Yuchtman, N. (2023a). Exporting the Surveillance State via Trade in AI. *NBER Working Paper*.

Beraja, M., Kao, A., Yang, D. Y., and Yuchtman, N. (2023b). AI-tocracy. *The Quarterly Journal of Economics*, 138(3):1349–1402.

Beraja, M., Yang, D. Y., and Yuchtman, N. (2023c). Data-intensive innovation and the state: evidence from ai firms in china. *The Review of Economic Studies*, 90(4):1701–1723.

Blühdorn, I. (2013). *Simulative Demokratie: Neue Politik nach der postdemokratischen Wende*. Suhrkamp Verlag.

Blühdorn, I. (2024). *Unhaltbarkeit: Auf dem Weg in eine andere Moderne*. Suhrkamp Verlag.

Borradaile, G., Burkhardt, B., and LeClerc, A. (2020). Whose tweets are surveilled for the police: an audit of a social-media monitoring tool via log files. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, FAT* '20. ACM.

Büchi, M., Festic, N., and Latzer, M. (2022). The Chilling Effects of Digital Dataveillance: A Theoretical Model and an Empirical Research Agenda. *Big Data & Society*, 9(1):205395172110653.

Çelik, B. (2013). The politics of the digital technoscape in turkey: Surveillance and resistance of kurds. *New perspectives on Turkey*, 49:31–56.

Chen, J., Pan, J., and Xu, Y. (2016). Sources of authoritarian responsiveness: A field experiment in china. *American Journal of Political Science*, 60(2):383–400.

Chin, J. and Lin, L. (2022). *Surveillance State: Inside China's Quest to Launch a New Era of Social Control*. St. Martin's Publishing Group.

Cong, L. W., Wei, W., Xie, D., and Zhang, L. (2022). Endogenous growth under multiple uses of data. *Journal of Economic Dynamics and Control*, 141:104395. Markets and Economies with Information Frictions.

Cong, L. W., Xie, D., and Zhang, L. (2021). Knowledge accumulation, privacy, and growth in a data economy. *Management Science*, 67(10):6480–6492.

Crouch, C. (2019). 10. post-democracy and populism. *The Political Quarterly*, 90(S1):124–137.

Davenport, C. (2007). State repression and political order. *Annual Review of Political Science*, 10(Volume 10, 2007):1–23.

Diamond, L. (2015). Liberation technology 1. In *In search of democracy*, pages 132–146. Routledge.

Earl, J. (2003). Tanks, tear gas, and taxes: Toward a theory of movement repression. *Sociological Theory*, 21(1):44–68.

Earl, J., James, R., Ramo, E., and Scovill, S. (2021). Protest, activism, and false information. In *The Routledge companion to media disinformation and populism*, pages 290–301. Routledge.

Earl, J., Maher, T., and Pan, J. (2022). The Digital Repression of Social Movements, Protest, and Activism: A Synthetic Review. *Science advances*, 8.

Earl, J., McKee Hurwitz, H., Mejia Mesinas, A., Tolan, M., and Arlotti, A. (2013). This protest will be tweeted: Twitter and protest policing during the pittsburgh g20. *Information Communication and Society*, 16(4):459–478.

Egorov, G. and Sonin, K. (2024). The political economics of non-democracy. *Journal of Economic Literature*, 62(2):594–636.

Enders, A. M. and Uscinski, J. E. (2021). The role of anti-establishment orientations during the trump presidency. *The Forum*, 19(1):47–76.

Enikolopov, R., Petrova, M., and Zhuravskaya, E. (2011). Media and Political Persuasion: Evidence from Russia. *American Economic Review*, 101(7):3253–85.

Enikolopov, R., Rochlitz, M., Schoors, K., and Zakharov, N. (2022). The effect of independent online media in an autocracy. *SSRN Electronic Journal*.

Feldstein, S. (2021). *The Rise of Digital Repression: How Technology Is Reshaping Power, Politics, and Resistance*. Oxford: Oxford University Press.

Feldstein, S. (2023). Surveillance in the Illiberal State. In Sajó, A., Uitz, R., and Holmes, S., editors, *Routledge Handbook of Illiberalism*, chapter 24, pages 351–364. Routledge.

Festic, N. (2022). Same, same, but different! qualitative evidence on how algorithmic selection applications govern different life domains. *Regulation & Governance*, 16(1):85–101.

Freyburg, T. and Garbe, L. (2018). Authoritarian practices in the digital age| blocking the bottleneck: Internet shutdowns and ownership at election times in sub-saharan africa. *International Journal of Communication*, 12(0).

Gao, Y., Zang, L., Roth, A., and Wang, P. (2017). Does democracy cause innovation? An empirical test of the popper hypothesis. *Research Policy*, 46(7):1272–1283.

Garbe, L. (2023). Pulling through elections by pulling the plug: Internet disruptions and electoral violence in uganda. *Journal of Peace Research*, 0(0):00223433231168190.

Geddes, B., Wright, J., and Frantz, E. (2014). Autocratic breakdown and regime transitions: A new data set. *Perspectives on Politics*, 12(2):313–331.

Gohdes, A. (2023). *Repression in the Digital Age: Surveillance, Censorship, and the Dynamics of State Violence*. Disruptive Technology and International Security Series. Oxford University Press, Incorporated.

Gohdes, A. R. (2020). Repression technology: Internet accessibility and state violence. *American Journal of Political Science*, 64(3):488–503.

Große-Bley, J. and Kostka, G. (2021). Big Data Dreams and Reality in Shenzhen: An Investigation of Smart City Implementation in China. *Big Data & Society*, 8(2):1–14.

Guriev, S. and Treisman, D. (2019). Informational Autocrats. *Journal of Economic Perspectives*, 33(4):100–127.

Guriev, S. and Treisman, D. (2020). A Theory of Informational Autocracy. *Journal of Public Economics*, 186(104158).

Guriev, S. and Treisman, D. (2022). *Spin Dictators: The Changing Face of Tyranny in the 21st Century*. Princeton University Press.

Hanley, H., Lu, Y., and Pan, J. (2024). Narratives of foreign media ecosystems in chinese social media discussions of the russo-ukrainian war. *Working paper*.

Hanley, H. W. A., Kumar, D., and Durumeric, Z. (2023). Happenstance: Utilizing semantic search to track russian state media narratives about the russo-ukrainian war on reddit. *Proceedings of the International AAAI Conference on Web and Social Media*, 17(1):327–338.

Hassanpour, N. (2014). Media disruption and revolutionary unrest: Evidence from mubarak's quasi-experiment. *Political Communication*, 31(1):1–24.

Jiang, M. (2014). The business and politics of search engines: A comparative study of baidu and google's search results of internet events in china. *New Media & Society*, 16(2):212–233.

Jones, C. I. and Tonetti, C. (2020). Nonrivalry and the economics of data. *American Economic Review*, 110(9):2819–58.

Jones, M. (2022). *Digital Authoritarianism in the Middle East: Deception, Disinformation and Social Media*. Hurst.

Kalmus, V., Bolin, G., and Figueiras, R. (2022). Who is Afraid of Dataveillance? Attitudes Toward Online Surveillance in a Cross-cultural and Generational Perspective. *New Media & Society*.

Kappeler, K., Festic, N., and Latzer, M. (2023). Dataveillance imaginaries and their role in chilling effects online. *International Journal of Human-Computer Studies*, 179:103120.

Karpa, D., Klarl, T., and Rochlitz, M. (2022). *Artificial Intelligence, Surveillance, and Big Data*, pages 145–172. Springer International Publishing, Cham.

Katzenbach, C. and Ulbricht, L. (2019). Algorithmic governance. *Internet Policy Review*, 8(4):1–18.

Keremoğlu, E., Weidmann, N. B., Gamero-Garrido, A., Carisimo, E., Dainotti, A., and Snoeren, A. C. (2024). Network topology facilitates internet traffic control in autocracies. *PNAS Nexus*, 3(3):pgae069.

King, G., Pan, J., and Roberts, M. E. (2013). How Censorship in China Allows Government Criticism but Silences Collective Expression. *American Political Science Review*, 107(2):326–343.

King, G., Pan, J., and Roberts, M. E. (2014). Reverse-engineering Censorship in China: Randomized Experimentation and Participant Observation. *Science*, 345(6199):1251722.

King, G., Pan, J., and Roberts, M. E. (2017). How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, not Engaged Argument. *American Political Science Review*, 111(3):484–501.

Knight, A. and Creemers, R. (2021). Going Viral: The Social Credit System and COVID-19. *Working Paper, Leiden University*.

Kostka, G. (2019). China's Social Credit Systems And Public Opinion: Explaining High Levels of Approval. *New Media & Society*, 21(7):1565–1593.

Kostka, G. and Antoine, L. (2020). Fostering Model Citizenship: Behavioral Responses to China's Emerging Social Credit Systems. *Policy & Internet*, 12(3):256–289.

Kostka, G., Steinacker, L., and Meckel, M. (2023). Under Big Brother's Watchful Eye: Cross-country Attitudes Toward Facial Recognition Technology. *Government Information Quarterly*, 40(1):101761.

Kravchenko, M. (2019). Russian anti-extremism legislation and internet censorship. *The Soviet and Post-Soviet Review*, 46:158–186.

Laruelle, M. and Kudaibergen, D. T. (2024). *Making Sense of the Bloody January 2022 Mass Protests*, page 244–274. Cambridge University Press.

Leber, A. and Abrahams, A. S. (2019). A storm of tweets: Social media manipulation during the gulf crisis. *Review of Middle East Studies*, 53:241–258.

Li, H. and Kostka, G. (2022). Accepting but not Engaging with it: Digital Participation in Local Government-run Social Credit Systems in China. *Policy & Internet*, 14(4):845–874.

Liang, F., Das, V., Kostyuk, N., and Hussain, M. M. (2018). Constructing a Data-Driven Society: China's Social Credit System as a State Surveillance Infrastructure. *Policy & Internet*, 10(4):415–453.

Libman, A. and Rochlitz, M. (2019). *Federalism in China and Russia*. Edward Elgar Publishing.

Liu, C. (2022). Who Supports Expanding Surveillance? Exploring Public Opinion of Chinese Social Credit Systems. *International Sociology*, 37(3):391–412.

Lutscher, P. M., Weidmann, N. B., Roberts, M. E., Jonker, M., King, A., and Dainotti, A. (2020). At home and abroad: The use of denial-of-service attacks during elections in nondemocratic regimes. *Journal of Conflict Resolution*, 64(2-3):373–401.

Manea, E. (2016). In the name of culture and religion: The political function of blasphemy in islamic states. *Islam and Christian–Muslim Relations*, 27(1):117–127.

Manokha, I. (2018). Surveillance, Panopticism, and Self-Discipline in the Digital Age. *Surveillance & Society*, 16:219–237.

Manow, P. (2020). *(Ent-)Demokratisierung der Demokratie*. Suhrkamp Verlag.

Marples, D. R. (2022). Russia's war goals in ukraine. *Canadian Slavonic Papers*, 64(2-3):207–219.

Mau, S., Lux, T., Westheuser, L., and Verlag, S. (2023). *Triggerpunkte: Konsens und Konflikt in der Gegenwartsgesellschaft*. edition suhrkamp : Sonderdruck. Suhrkamp.

Morse, Y. L. (2012). The era of electoral authoritarianism. *World Politics*, 64(1):161–198.

Oz, M. and Yanik, A. (2022). Fear of surveillance: Examining turkish social media users' perception of surveillance and willingness to express opinions on social media. *Mediterranean Politics*, 0(0):1–25.

Pan, J. (2020). *Welfare for Autocrats: How Social Assistance in China Cares for its Rulers*. Oxford University Press.

Pan, J. and Chen, K. (2018). Concealing corruption: How chinese officials distort upward reporting of online grievances. *American Political Science Review*, 112(3):602–620.

Pan, J., Shao, Z., and Xu, Y. (2022). How government-controlled media shifts policy attitudes through framing. *Political Science Research and Methods*, 10(2):317–332.

Pan, J. and Siegel, A. A. (2020). How saudi crackdowns fail to silence online dissent. *American Political Science Review*, 114(1):109–125.

Pei, M. (2024). *The Sentinel State: Surveillance and the Survival of Dictatorship in China*. Harvard University Press.

Peisakhin, L. and Rozenas, A. (2018). Electoral Effects of Biased Media: Russian Television in Ukraine. *American Journal of Political Science*, 62(3):535–550.

Penney, J. W. (2016). Chilling Effects: Online Surveillance and Wikipedia Use. *Berkeley Technology Law Journal*, 31(1):117–182.

Penney, J. W. (2017). Internet Surveillance, Regulation, and Chilling Effects Online: a Comparative Case Study. *Internet Policy Review*, 2(6).

Penney, J. W. (2022). Understanding chilling effects. *Minnesota Law Review*, 106:1451.

Rabe, W. and Kostka, G. (2024). Perceptions of social credit systems in southeast asia: An external technology acceptance model. *Global Policy*.

Raman, R. S., Evdokimov, L., Wurstrow, E., Halderman, J. A., and Ensafi, R. (2020). Investigating large scale https interception in kazakhstan. In *Proceedings of the ACM Internet Measurement Conference*, IMC '20, page 125–132, New York, NY, USA. Association for Computing Machinery.

Roberts, M. E. (2018). *Censored: Distraction and Diversion Inside China's Great Firewall*. Princeton: Princeton University Press.

Roberts, M. E. (2020). Resilience to online censorship. *Annual Review of Political Science*, 23(1):401–419.

Robinson, D. and Tannenberg, M. (2019). Self-censorship of Regime Support in Authoritarian States: Evidence From List Experiments in China. *Research & Politics*, 6(3):2053168019856449.

Schauer, F. (1978). Fear, Risk, and the First Amendment: Unraveling the "Chilling Effect.". *Boston University Law Review*, 58(0):685—732.

Snyder, T. (2018). *The Road To Unfreedom*. Crown, New York.

Stoycheff, E. (2016). Under Surveillance. *Journalism & Mass Communication Quarterly*, 93(2):296–311.

Stoycheff, E. (2022). Cookies and Content Moderation: Affective Chilling Effects of Internet Surveillance and Censorship. *Journal of Information Technology & Politics*, pages 1–12.

Stoycheff, E., Burgess, G. S., and Martucci, M. C. (2020). Online Censorship and Digital Surveillance: The Relationship Between Suppression Technologies and Democratization Across Countries. *Information, Communication & Society*, 23(4):474–490.

Stoycheff, E., Liu, J., Xu, K., and Wibowo, K. (2019). Privacy and the Panopticon: Online Mass Surveillance's Deterrence and Chilling Effects. *New Media & Society*, 21(3):602–619.

Strittmatter, K. (2020). *We Have Been Harmonized: Life in China's Surveillance State*. New York: Custom House.

Svolik, M. (2012). *The Politics of Authoritarian Rule*. Cambridge: Cambridge University Press.

Tabellini, M. and Magistretti, G. (2022). Economic Integration and the Transmission of Democracy. *NBER Working Paper No. 30055*.

Tannenberg, M. (2022). The Autocratic Bias: Self-censorship of Regime Support. *Democratization*, 29(4):591–610.

Tau, B. (2024). *Means of Control: How the Hidden Alliance of Tech and Government Is Creating a New American Surveillance State*. Crown.

Thaler, R. and Sunstein, C. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press.

Thaler, R. H. and Sunstein, C. R. (2003). Libertarian paternalism. *The American Economic Review*, 93(2):175–179.

Tufekci, Z. (2014). Engineering the public: Big data, surveillance and computational politics. *First Monday*, 19(7).

Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral Studies*, 69:102199.

Walsh, J. P. and O'Connor, C. (2019). Social media and policing: A review of recent research. *Sociology Compass*, 13(1):e12648.

Wang, Q.-J., Feng, G.-F., Wang, H.-J., and Chang, C.-P. (2021). The Impacts of Democracy on Innovation: Revisited Evidence. *Technovation*, 108:102333.

Xu, X. (2021). To Repress or to Co-opt? Authoritarian Control in the Age of Digital Surveillance. *American Journal of Political Science*, 65(2):309–325.

Xu, X., Kostka, G., and Cao, X. (2022). Information Control and Public Support for Social Credit Systems in China. *Journal of Politics*, 84(2).

Zhuravskaya, E., Petrova, M., and Enikolopov, R. (2020). Political effects of the internet and social media. *Annual Review of Economics*, 12(1):415–438.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Public Affairs.

2 Surveillance, Innovation, and Growth

Essay title: Authoritarian Surveillance, Innovation and Growth

Authors: Torben Klarl, David Karpa, Matheus Leusin, Michael Rochlitz

Abstract

How do authoritarian political institutions influence the ability of an economy to innovate? The existing literature identifies a mostly negative effect of autocracy on innovation. In this paper, we build a theoretical model to investigate if this premise still holds in autocracies that rely on digital surveillance for political control, and that use the data obtained through surveillance as a subsidy for innovation in fields such as artificial intelligence. Our model illustrates the trade-off between the negative effect of surveillance on research and creativity, and the positive effect of the availability of large amounts of data. We find that while on average the effect of authoritarian institutions on innovation remains negative, in fields such as artificial intelligence where large amounts of data are important, autocracies can – under specific circumstances – achieve better results than competitive democracies.

Keywords: endogenous growth, authoritarian innovation, surveillance, artificial intelligence, China

JEL Classification: E02, L52, O11, O38, O43

Publication: Under Review at European Economic Review (since 02. January 2024). Preprint avaible: Klarl, Torben and Karpa, David F. and Leusin, Matheus Eduardo and Rochlitz, Michael, Authoritarian Surveillance, Innovation and Growth (October 6, 2023). Available at SSRN: https://ssrn.com/abstract=4594849 or http://dx.doi.org/10.2139/ssrn.4594849

2.1 Introduction

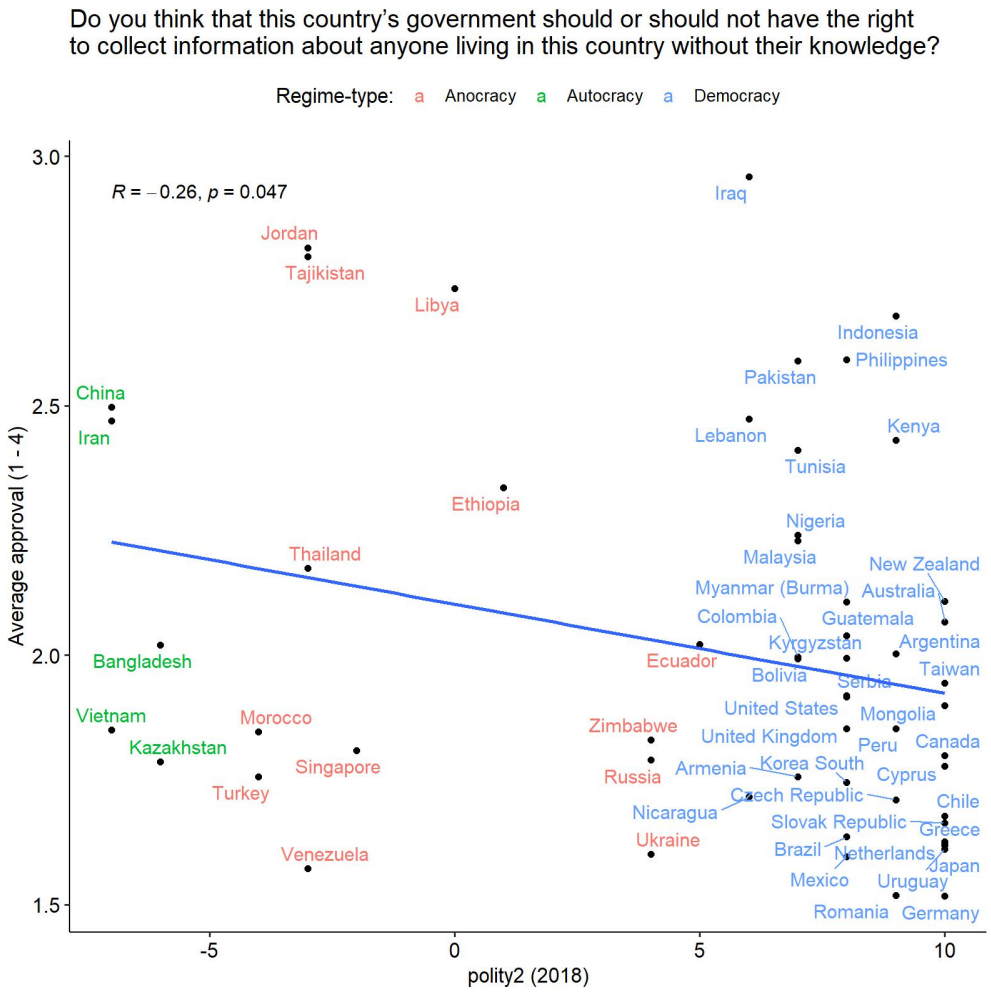
It is generally assumed that democratic institutions are more likely to encourage innovative activities that lead to higher productivity and, hence, higher economic growth than authoritarian institutions (Acemoglu and Robinson 2012; Acemoglu et al. 2019; Wang et al. 2021). However, does this conclusion still hold if we focus on data-intensive industries that heavily rely on large amounts of data produced by an autocratic state?

In this paper, we develop an endogenous growth model with a myopic government to illustrate the trade-off between surveillance and data availability in an autocracy. In traditional autocracies, surveillance and repression harm innovation, as these activities render scientists and researchers less creative and productive. In modern informational autocracies (see Guriev and Treisman 2022 for a conceptual discussion), however, surveillance can also be used to gather and bundle large amounts of data, which can then be used as an input – provided by the state – to accelerate innovation in data-intensive fields, such as artificial intelligence (Beraja et al. 2023c).

Recent contributions such as Cong et al. (2021) that focus on the growth aspects of a data economy point out that consumers might suffer from data misuse or privacy violations that come as a consequence of surveillance by the state. However, as we argue in this paper, higher levels of government surveillance might also entail potential benefits for households, such as lower crime rates, fewer terrorist attacks, or smoother government services. There is evidence that in particular in autocracies, people accept a certain level of government surveillance in exchange for more security and better government services (Kostka 2019; Habich-Sobiegalla and Kostka 2022; see also Figure 2.1).

A general finding of the recent literature on the data economy and economic growth is that data are either underused due to their non-rivalry and property rights owned by consumers (Jones and Tonetti 2020) or overused as a result of an inefficiently small R&D sector (Cong et al. 2021). We complement this literature by showing that R&D misallocation tends to zero, if households that own their data put a large weight on governmental surveillance relative to private data misuse, i.e. the costs of surveillance are sufficiently low. We argue that in the race for primacy in data-intense technologies between large democracies such as

Figure 2.1: Approval for government surveillance and institutional environment



Note: Data on average approval is from wave 7 of the World Value Survey; to measure institutional quality we take the polity2 indicator from https://www.systemicpeace.org/polityproject.html

the US and large autocracies such as China (see Lee 2018 for a discussion), the ability of a state to gather and bundle data at a sufficiently low social cost and use it as an input for R&D could well play the role of a game changer, providing autocracies with an advantage over democratic political systems.

Our paper and model speak to several different strands of literature. First, we present an endogenous growth model that builds directly on Veldkamp (2005); Jones and Tonetti (2020); Cong et al. (2021, 2022) and Beraja et al. (2023c). Like in Cong et al. (2021) we let innovator firms develop and supply differentiated varieties of data-intensive goods, such as algorithms. These goods are used to produce the final good. In contrast to Veldkamp (2005); Jones

and Tonetti (2020); Cong et al. (2021) and Cong et al. (2022), we also introduce a myopic government that produces governmental data as a by-product of surveillance and assume that households derive positive utility from surveillance. Hence, we have governmental as well as private data in the economy.

Second, in contrast to Beraja et al. (2023c), the government is self-interested in that it tries to secure its power. For this reason, it raises taxes to finance surveillance. One major point this paper makes is that from a socially optimal point of view, as in Cong et al. (2021), data are overused at the expense of R&D labor. The misallocation of R&D labor is particularly pronounced if the degree of knowledge spillover is large and/or the importance of data for the development of new algorithms is low. However, we show that governmental surveillance can moderate this distortion towards zero. This finding complements the work of Cong et al. (2021). Another finding is that rent-seeking governments tend to set a tax above zero. For a reasonable parameter calibration, this accelerates the negative impact of underemployment in the R&D sector, which is still present due to data overuse.

Third, a potential productivity impact of governmental surveillance relative to output on the growth rate of algorithms is only transitory, but has no long-run effects on the growth rate and labor market allocations. More surveillance reduces creativity, but generates new data, thereby increasing output via new algorithms that in turn contribute to productivity in the R&D sector. We show that in the long run, both effects exactly cancel each other out, as aggregate governmental surveillance grows at the same rate as aggregate output.

Our paper is organized as follows. Section 2.2 motivates the paper, by providing a brief overview of the existing literature on the effect of democratic and authoritarian political institutions on innovation, and by outlining how recent breakthroughs in big data technologies such as artificial intelligence may have changed the trade-off between authoritarian control and innovation. Section 2.3 presents our baseline model. Section 2.4 introduces the social planner's problem, and section 2.5 the action of the rent-seeking government. Section 2.6 performs a calibration exercise to illustrate the steady-stage labor market allocations, and section 4.7 concludes.

2.2 Motivation

2.2.1 Democracy, Autocracy and Innovation

Theoretically and empirically, the literature on the effect of political institutions on economic growth remains divided. While there is growing support for the idea that democracy is good for growth (Jamali et al. 2007; Acemoglu and Robinson 2012; Knutsen 2013, 2015; Acemoglu et al. 2019), other studies find ambiguous or no effects (Olson 1982; Libman 2012; Piatek et al. 2013; Murtin and Wacziarg 2014; Pozuelo et al. 2016; Truex 2017; Ghardallou and Sridi 2020).

A subset of this literature studies innovation as one possible mechanism why democracies might outperform autocracies (Huang and Xu 1999; Carayannis and Campbell 2014; Knutsen 2015; Silve and Plekhanov 2018; Tang and Tang 2018; see Gao et al. 2017 for a dissenting view). One potential channel is the effect of democracy on human capital, which in turn can positively affect innovation (Tebaldi and Elmslie 2008; Klomp and de Haan 2012). Other studies find a positive effect of political freedom on innovation in high-tech sectors, while the effect remains ambiguous or negative for low-tech sectors (Aghion et al. 2007; Zuazu 2019). With respect to natural resources, the effect seems clearer, as they seems to have a negative effect on innovation in autocracies, but not in democracies (Rosenberg and Tarasenko 2020). Finally, a literature based on case studies argues that, even though on average authoritarian regimes lag behind democracies when it comes to generating innovation, sometimes pockets of competence exist where autocracies were able to successfully compete with liberal democracies (Graham 1987, 1993; Stokes 2000; Josephson 2005; Gomez and Canales 2015). Overall, the existing literature suggests the long-term effect of authoritarian institutions on innovation to be either negative or ambiguous.

Our paper introduces two innovations to this literature. First, we focus on highly digitized autocracies, i.e. authoritarian regimes that use sophisticated methods to censor, monitor, and control the internet and other sources of information (King et al. 2013, 2014; Shadmehr and Bernhardt 2015; Roberts 2018; Strittmatter 2020; Guriev and Treisman 2022), rather than relying on repression and more traditional methods of control. We show that surveillance in such autocracies still introduces distortions for both scientific research and

the economy in general, and thus comes at an economic cost. However, the fact that large amounts of data are gathered, bundled and made available in a centralized way can also offer potential advantages, in particular with respect to research in data-intense fields such as deep learning. Building on this trade-off, our second innovation is to investigate how informational autocracies fare with respect to research and innovation in technologies that rely on large amounts of data. Section 2.2.2 briefly introduces the specific features of data-intense technologies, and explains why they might have the potential to change the way authoritarian institutions affect innovation.

2.2.2 Artificial Intelligence and Deep Learning

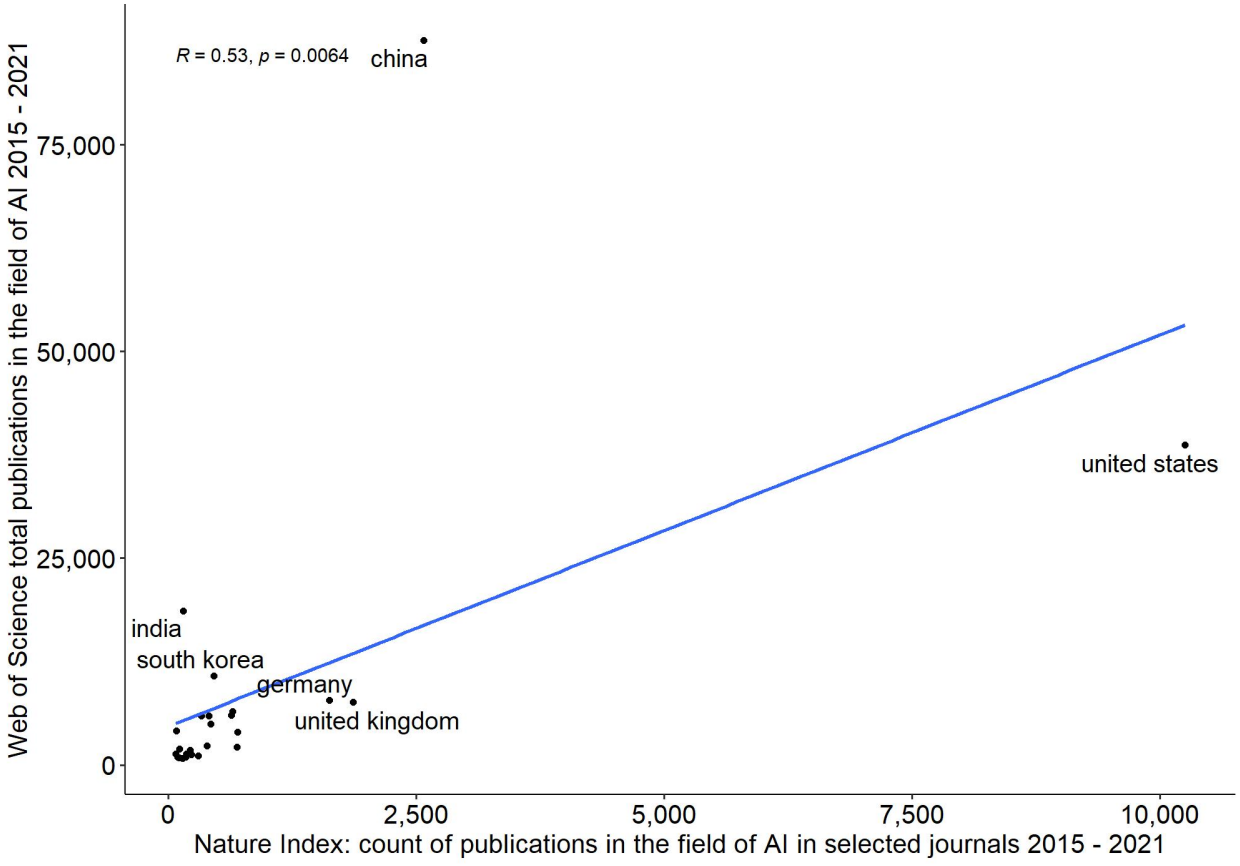
In 2006, two publications on recursive learning in many-layered neuronal networks gave a boost to research in the field of machine learning (Hinton and Salakhutdinov 2006; Hinton et al. 2006). Building on this breakthrough, researchers concentrated on developing algorithms that were able to learn from large amounts of data, without giving the algorithm any detailed previous instructions. Instead, the algorithm relies solely on artificial multi-layered neural networks, which in their functioning resemble the neural networks of the human brain. After some time, these algorithms were able to classify and learn from large amounts of data with high levels of precision (Krizhevsky et al. 2012; He et al. 2016). Importantly, the amount of data available and the depth of the network (measured by the number of layers) are positively correlated with the ability of the algorithm to learn and self-improve, hence *deep* learning (Hey 2009; Domingos 2015). It is this feature which lead some researchers to describe the quantity of data available for researchers as the "new oil" of the knowledge economies of the 21st century (Spitz 2017; Lee 2018; Taffel 2023).

Once the initial breakthrough in machine learning had been made, the new technology was rapidly applied to various fields, such as image recognition, natural language processing, toxicology, medical image analysis, management, bioinformatics, financial fraud detection, as well as surveillance and military technologies. For our argument it is important that in the *application* phase of these technologies, the type of cutting-edge research abilities where US elite-universities still have an absolute advantage have become relatively less important. Instead, the training and incremental improvement of existing algorithms with large amounts

of data has become central to advance innovation in these sectors (Agrawal et al. 2018; Cockburn et al. 2018; Beraja et al. 2023b). Such processes can also be carried out by well-qualified specialists, even if they are not competing at the absolute frontier of global knowledge production.

Figure 2.2 illustrates this point, and shows that with respect to the overall amount of AI-related publications (as measured by the total number of publications in the field of AI that are indexed by Web of Science), China has already an advantage over all other countries in the world. We take this number as an indication that already today, China is leading when it comes to the overall amount of researchers that are able to do applied work in the field. When it comes to top-level publications (proxied by the Nature index of high-quality publications in AI), however, the United States are still ahead (Wagner et al. 2020, 2022).

Figure 2.2: Quality vs Quantity of Publications in the Field of AI



It is thus with respect to the training of existing technology with large amounts of data that autocracies might have an advantage. First, authoritarian states are using surveillance

systems that can gather larger amounts of data about the behavior and characteristics of the population than most democracies (Qiang 2019; Strittmatter 2020), despite surveillance also playing a role in many democratic countries (Zuboff 2019). Second, as a result of less stringent privacy laws, this data can be shared by the government with private firms, similar to a subsidy, providing them with an advantage over firms in democratic political contexts (Jones and Tonetti 2020; Beraja et al. 2023b,c,a). Third, population size does matter, as in states with a larger population the amount of data that can be provided as input is also larger. Finally, digital technology penetration through a society is playing a role as well. The larger is the number of digital services and devices that are used, the more data is generated, and the higher the amount of data that can be collected by the government. In the next section, we introduce a simple model to illustrate how these specific features can influence the speed of innovation in a digital autocracy.

2.3 Baseline Model

2.3.1 Households

Our economy is populated with a fixed number of infinitely-lived, homogeneous representative households. Each household has $L(t) = L(0)\exp[nt]$ members, with $n \geq 0$ denoting the exogenous growth rate of the population. Further, $L(0) = L_0 > 0$. In every period, each member inelastically supplies one unit of labor per time unit. We normalize the number of households to one. We further follow Veldkamp (2005) and Jones and Tonetti (2020) and assume that consumers produce data as a by-product of consumption.

These data can be sold to the research sector (Jones and Tonetti 2020). However, data commonly comprises personal information and, thus, the potential misuse of data leads to a disutility that households consider when they sell their data.

The instantaneous utility function introduced below captures in a stylized manner the utility costs and benefits of state surveillance. The household's instantaneous utility function $u(t)$ is given by

$$u(t) = \left[\frac{[G(t)^{\epsilon}c(t)^{1-\epsilon}]^{1-\theta}}{1-\theta} - \iota[d_c(t)^{\kappa}G(t)^{1-\kappa}]^{\chi}\right], \tag{2.1}$$

with $\theta > 0$, $\kappa \in (0, 1]$ and $\theta \neq 1$ as the magnitude of the elasticity of marginal utility of

consumption. χ parameterizes the weighted average of disutility of data misuse (with weight κ) or privacy violation due to governmental surveillance $G(t)$ weighted with the parameter $1-\kappa$. $\epsilon \in [0,1)$ weights the utility derived from governmental surveillance, $G(t)$ relative to private consumption, $c(t)$. For the special case that $\kappa = \epsilon = 1$, a household does not derive utility and disutility from governmental surveillance. While Cong et al. (2021) focus on this special case ($\kappa = \epsilon = 1$), our paper goes further by focusing explicitly on governmental surveillance activities.

Beyond being consistent with a balanced growth path, (2.1) captures in a stylized manner the utility costs and benefits of surveillance. More surveillance leads to more (subjective) security and stability, such as lower crime rates and fewer terrorists attacks, but, on the other hand, restricts civil liberties. Thus, as long as $[G(t)^{\epsilon}c^{1-\epsilon}]^{1-\theta}\epsilon < \iota\chi[d_c(t)^{\kappa}G(t)^{1-\kappa}]^{\chi}$, the marginal utility of G turns out to be negative, as the marginal costs of surveillance exceeds the marginal benefits in utility terms. In the following, we impose that $\chi > 1$ in order to guarantee the convexity of disutility in (2.1).

Given $\theta > 0$ and $\rho > 0$, the representative household's problem is to choose a plan $\{c(t), d_c(t)\}_{t=0}^{\infty}$ so as to

$$\max_{\{c(t),d_c(t)\}} \int_0^{\infty} \exp[-(\rho-n)t] \left[\frac{[G(t)^{\epsilon}c(t)^{1-\epsilon}]^{1-\theta}}{1-\theta} - \iota[d_c(t)^{\kappa}G(t)^{1-\kappa}]^{\chi} \right] dt \tag{2.2}$$

subject to

$$c(t) \geq 0 \tag{2.3}$$

$$d_c(t) \geq 0 \tag{2.4}$$

$$\dot{a}(t) = (r(t)-n)a(t) + w(t) + p_{dc}(t)d_c(t) - c(t) \tag{2.5}$$

$$\frac{\dot{d}_c(t)}{d_c(t)} \leq \frac{\dot{c}(t)}{c(t)}, \frac{\dot{d}_g(t)}{d_g(t)} \leq \frac{\dot{G}(t)}{G(t)} \tag{2.6}$$

and the No-Ponzi game condition

$$\lim_{t \Rightarrow \infty} a(t) \exp[-\int_0^t (r(s)-n)ds] \geq 0. \tag{2.7}$$

Here, $r(t)$ is the risk-free interest rate, $a(t)$ is the per capita financial wealth that consists

of raw capital and perpetual patents as will become clearer below. Moreover, the constraint (2.38) requires that the growth rate of data is bounded by the growth rate of consumption. Restriction (2.38) also implies that data is a by-product of consumption: $d_c(t) \leq \zeta c(t)$ for some arbitrary constant $\zeta \in (0,1)$ (see Cong et al. (2021)). As usual, ρ is the consumer's subjective discount rate, while $p_{dc}(t)$ shows the price per data unit $d_c(t)$ that can be realized by selling data (as a by-product of consumption) to intermediate good producers (see Veldkamp (2005) or Cong et al. (2021)). $w(t)$ is the wage rate for labor supply.

Solving the optimization problem delivers the Euler equations for consumption (2.8) and data (2.9), respectively as

$$\frac{\dot{c}(t)}{c(t)} = \frac{1}{\tilde{\theta}}\left(r - n - \rho + (1-\theta)\epsilon\frac{\dot{G}(t)}{G(t)}\right) \tag{2.8}$$

$$\frac{\dot{p_c}(t)}{p_c(t)} + (1-\chi\kappa)\frac{\dot{d}_c(t)}{d_c(t)} - (1-\kappa)\chi\frac{\dot{G}(t)}{G(t)} = r - n - \rho, \tag{2.9}$$

and the transversality condition (tvc)

$$\lim_{t\Rightarrow\infty} a(t)\exp[-\int_0^t (r(s)-n)ds] = 0. \tag{2.10}$$

This follows directly from applying Pontryagin's maximum principle to the problem. Note further that $\tilde{\theta} \equiv [1-(1-\theta)(1-\epsilon)]$ denotes the effective rate of the intertemporal elasticity of substitution. Note that for $\epsilon = 0$, we have $\tilde{\theta} = \theta$.

2.3.2 Surveillance and the Government

In order to keep the model analytically tractable, we model the governmental sector as simple as possible. We assume that a fraction $m(t) \in (0,1)$ of governmental surveillance activities $G(t)$ generates aggregate governmental data $D_g(t) = m(t)G(t) = d_g(t)L(t)$. We allow $m(t)$ to decrease (or $\frac{1}{m(t)}$ to increase) over time to include the possibility that governmental surveillance becomes more efficient over time, in the sense that more data can be generated with a given level of surveillance activities due to technological improvements. Governmental surveillance activities are financed via two sources. First, the government imposes a profit tax on final goods producers with a constant tax rate $\tau \in (0,1)$. Second, the

government exclusively sells data-sets $d_g(t)$ to the research sector (where data firms operate) at price $p_{dg}(t)$. For simplicity, we do not allow data sharing across data firms. Moreover, the government cannot lend or borrow. The budget is thus always balanced. Taken together, the governmental constraint reads as

$$G(t) = \tau Y(t) + p_{dg}(t) d_g(t) L(t). \tag{2.11}$$

From equation (2.11), an implicit assumption is that governmental data are not shareable across firms. The reasons for this assumption are twofold. First, we focus on the normative and positive implications of data shareability within a firm.[1] Second, as shown by (Beraja et al. (2023c)), from an empirical perspective this case is more relevant for authoritarian regimes. Like in Beraja et al. (2023c)), governments collect their own data and sell them to a specific firm for analysis, while simultaneously excluding other firms from using the same data.

2.3.3 Production Side of the Economy

Final Goods Sector

The production side of our economy borrows elements from Romer (1990) and Jones (1995). The final goods sector produces the consumption aggregate with labor and intermediate goods as factor inputs in an environment with perfect competition. The production of the final goods sector is written as

$$Y(t) = L_Y(t)^{1-\alpha} \int_0^{N(t)} x_{i,t}^{\alpha} di, \tag{2.12}$$

where $Y(t)$ indicates the output of the consumption aggregate. $L_Y(t)$ represents labor used in the final goods production, and $N(t)$ is the technological frontier. $x_{i,t}$ is the amount of a specific, i indexed intermediate good $x_{i,t}$ (e.g. a machine for instance) that is used in final goods production at time t. $\alpha \in (0, 1)$ indicates the share of intermediate inputs.

[1] The implications of non-rival data have already been studied by Jones and Tonetti (2020), although we have to point out that data shareability across firms would strengthen the importance of governmental surveillance data in our model.

Let $p_i(t)$ be the price paid for the ith intermediate good. Profit maximization together with the assumption of perfect competition implies that factors are paid their marginal products:

$$w_Y(t) = (1-\tau)(1-\alpha)\frac{Y(t)}{L_Y(t)}, \tag{2.13}$$
$$p_i(t) = (1-\tau)\alpha L_Y(t)^{1-\alpha} x_{i,t}^{\alpha-1}, \tag{2.14}$$

where w_i stands for the wage rate paid in the final goods sector.

Intermediate Goods Sector

From (2.14), we obtain the downward-sloping demand function of intermediate goods as:

$$x_{i,t} = L_Y(t)\left(\frac{\alpha(1-\tau)}{p_i(t)}\right)^{\frac{1}{1-\alpha}} \tag{2.15}$$

In the spirit of Dixit and Stiglitz (1977), we assume that the intermediate goods sector is monopolistically competitive. This implies that each firm produces exactly one of the differentiated intermediate goods. For production, each firm producing intermediate goods has to purchase one intermediate goods-specific blueprint from the R&D sector that will be introduced below. After the intermediate goods producer has purchased the blueprint, she is able to convert one unit of capital into one unit of intermediate good: $k_{i,t} = x_{i,t}$. The costs of the blueprint are the fixed costs for each firm. Moreover, the assumption of free entry ensures that fixed costs equal operating profits. This, in turn, implies that overall profits are zero. We assume that the marginal and average costs of production are constant. Hence, the operating profit's flow is given by

$$\begin{aligned}\pi_{i,t} &= (p_{i,t} - r)k_{i,t} \\ &= ((1-\tau)\alpha L_Y(t)^{1-\alpha} x_{i,t}^{\alpha-1} - r(t))k_{i,t}\end{aligned} \tag{2.16}$$

Profit maximization yields the usual markup pricing:

$$p_{i,t} = p(t) = \frac{r(t)}{\alpha}, \tag{2.17}$$

where $\frac{1}{\alpha} > 1$ is the markup over marginal costs. Using (2.17) in (2.14), we find that

$$x_{i,t} = \left[\frac{(1-\tau)\alpha^2}{r(t)}\right]^{\left(\frac{1}{1-\alpha}\right)} L_Y(t) \equiv x(t), \tag{2.18}$$

which implies that the quantity of intermediate goods production is independent of the variety. Thus, using (2.17) in (2.15), the aggregate capital stock is simply $K(t) = \int_0^N x_{i,t} di = N(t)x(t)$, so that (2.12) can be written as

$$Y(t) = L_Y(t)N(t)\left[\frac{(1-\tau)\alpha^2}{r(t)}\right]^{\left(\frac{\alpha}{1-\alpha}\right)}. \tag{2.19}$$

Inspecting (2.19) reveals that final output decreases with an increasing profit tax rate τ.

Data Firms in the Research Sector

The novelty of our paper is that private as well as governmental data enter R&D when new blueprints are developed. This distinguishes our contribution from Jones and Tonetti (2020) or Beraja et al. (2023c), where data only enter directly into the production of final goods. These contributions implicitly focus on data-driven industries by assuming that exclusively data intermediate goods are used for final good production. While this seems reasonable for industries such as Google or Facebook, more traditional industries still use non-data intermediate goods, where data can be used to improve the quality of such intermediate goods. As we focus on surveillance as well, our paper is also different from Cong et al. (2021, 2022), who neglect data from government surveillance as a potential driver of innovation.

In our paper, data firms operate in the R&D sector that employs scientists, $L_N(t)$ as well as a weighted average of private and governmental data. Hence, total data employed in the research sector at time t to discover new blueprints (or algorithms) is given by

$$D(t) \equiv [D_c(t)]^{\beta}[D_g(t)]^{1-\beta} = (d_c)^{\beta}(d_g)^{1-\beta}L(t), \tag{2.20}$$

with weighting factor $\beta \in [0,1]$. Two points in (2.20) are worth mentioning. First, if $\beta = 1$, the research sector only employs private data, i.e. $D(t) = D_c(t)$ (see Cong et al. (2021) for instance). Second, there is a scale effect. The size of aggregate data $D(t)$ depends on the size of the population $L(t)$, which means that more people (larger countries) produce more private as well as governmental data.

The aggregate technological frontier evolves according to:

$$\dot{N}(t) = \bar{\eta} N(t)^{\phi} (D(t))^{\xi} L_N(t)^{1-\xi}, \tag{2.21}$$

where $\bar{\eta} > 0$ is an efficiency term of innovation. $\xi \in (0,1)$ represents the relative contribution of data $D(t)$ and R&D scientists $L_N(t)$ in the production process of new varieties (or algorithms) $N(t)$, while $0 < \phi < 1$ captures the "standing on the shoulders of giants" effect of technology on the change in technology which can be interpreted as knowledge spillovers.[2]

We further assume that surveillance reduces the efficiency of innovation. This can be justified by the underlying psychological pressure of the government on researchers. Authoritarian surveillance entails targeted repression and makes citizens adhere to social or legal norms (Roberts 2018). Repression, however, also disincentives innovation activities on the entrepreneurial level (Acemoglu and Robinson 2012) and on the level of the individual, where it hinders the creativity that is crucial in the innovation process (Karpa et al. 2022). On the other hand, more governmental surveillance also implies more data and more varieties that can in turn enhance research productivity. To capture this idea in a parsimonious way, in contrast to Cong et al. (2021) or Cong et al. (2022), the research productivity term is endogenously explained by governmental surveillance activities $G(t)$ relative to output $Y(t)$:

$$\eta(t) = \bar{\eta} \left(\frac{G(t)}{Y(t)} \right)^{-\omega} = \bar{\eta} \left(\frac{d_g(t) L(t)}{m(t) Y(t)} \right)^{-\omega}, \tag{2.22}$$

with $\omega \in [0,1)$ that reflects the strength of negative influence of surveillance activities on research productivity. $m(t)$ is an exogenously growing (with constant rate) efficiency parameter. Because researchers are socialized in a given regime, they take the value of η as

[2]For $\phi < 0$, we have the "fishing out effect" effect, i.e. it is harder to find a new blueprint if the number of already discovered blueprints $N(t)$ is very large. For $\phi = 0$, both effects offset each other.

given by neglecting the negative influence of surveillance activities. For $\omega = 0$, (5) collapse to $\eta(t) = \bar{\eta}$, i.e. an exogenously given and time-independent efficiency term of innovation.

We assume that data firms develop new blueprints for new varieties of capital goods under conditions of free market entry and perfect competition. Hence, data firms enter as long as marginal benefits equals marginal costs of production. In other words, data firms maximize their profits $\pi_N(t)$ according to

$$\begin{aligned} \max_{\{L_N(t),d_c(t),d_g(t)\}} \pi_N(t) &= p_N(t)\eta(t)N(t)^{\phi}(D(t))^{\xi}L_N(t)^{1-\xi} - w_N(t)L_N(t) \\ &\quad -p_{dc}(t)d_c(t)L(t) - p_{dg}(t)d_g(t)L(t), \end{aligned} \tag{2.23}$$

where $p_N(t)$ represents the price of a blueprint. The first order conditions pin down the R&D sector's demand function of data from private households and government as well as the wages in the research sector. The free-entry conditions are:

$$p_N(t)(1-\xi)\eta N(t)^{\phi}(D(t))^{\xi}L_N(t)^{-\xi} = w_N(t), \tag{2.24}$$

$$p_N(t)\beta\xi\eta(t)N^{\phi}(t)(D(t))^{\xi-1}\frac{D(t)}{d_c(t)}L_N(t)^{1-\xi} = p_{dc}(t)L(t), \tag{2.25}$$

$$p_N(t)(1-\beta)\xi\eta(t)N(t)^{\phi}(D(t))^{\xi-1}\frac{D(t)}{d_g(t)}L_N(t)^{1-\xi} = p_{dg}(t)L(t). \tag{2.26}$$

2.3.4 Equilibrium

A competitive equilibrium is a set of allocations $\{c(t), Y_t, a(t), \{x_{i,t}\}_{i\in[0,N_t]}, d_c(t), d_g(t), L_Y(t), L_N(t), L(t), N(t), G(t), D_c(t), D_g(t)\}_{t=0}^{\infty}$, a price system $\{w(t), r_t, \{p_{i,t}\}_{i\in[0,N_t]}, p_{cd}(t), p_{gd}(t), P_N(t)\}$ and an imposed policy scalar $\{\tau\}$ such that for all t:

(i) $\{c(t)\}$ and $\{a(t)\}$ solve the household problem (2.37)-(2.7), $\{x_{i,t}\}_{i\in[0,N_t]}$ and $\{L_Y(t)\}$ solve the final goods producer problem, $\{p_{i,t}\}_{i\in[0,N_t]}$ and $\{\pi_{i,t}\}_{i\in[0,N_t]}$ solve the intermediate goods producers problem for all $i \in [0, N_t]$, $\{L_N(t), d_c(t), d_g(t)\}$ solve the data firm's R&D problem (2.23).

(ii) $\{w(t)\}$ clears the labor market $L_N(t)+L_Y(t) = L(t)$, $\{\mathrm{r(t)}\}$ clears the asset market with

$a(t)L(t) = N(t)p_N(t)$, $\{p_{gd}(t)\}$ clears the surveillance data market $d_g(t)L(t) = D_g(t)$, $\{p_{cd}(t)\}$ clears the consumption data market $d_c(t)L(t) = D_c(t)$, {N(t)} follows from the R&D production function (2.21). $G(t)$ follows from $\{\tau, Y(t), p_{dg}(t), d_g(t), L(t)\}$. In every point in time, the governmental budget (2.11) is balanced.

2.3.5 Governmental Surveillance and Growth: A Balanced Growth Path Characterisation

In this section, we solve the model along the balanced growth path (bgp). A bgp is a trajectory such that all variables grow at a constant exponential (but not necessarily equal) rate forever. In Appendix 1, we derive the growth rate for the decentralized economy on the bgp. This growth rate is different for private and governmental data provision. We summarize this finding in the following proposition:

Proposition 1. *As in Jones (1995), the decentralized economy does not exhibit a scale effet. As shown in Appendix 2.A.2, the bgp growth rates for* $y(t) \equiv \frac{Y(t)}{L(t)}$, $N(t)$, $c(t)$ *and* $g(t) \equiv \frac{G(t)}{L(t)}$ *are given by*

$$\hat{g} = \left\{ \frac{\xi[(1-\theta)\epsilon - (1-\kappa)\chi] + \kappa\chi}{\xi[\theta - 1 + \chi(1-\kappa)] + (1-\phi)\chi\kappa} \right\} n. \tag{2.27}$$

The bgp growth rates for private and public data sets $d_c(t)$ *and* $d_g(t)$*, respectively, are given by*

$$\begin{aligned} \hat{g}_d &= \frac{1-\phi}{\xi}\hat{g} - \frac{n}{\xi} \\ &= \left\{ \frac{[1-\theta][1+\epsilon(1-\phi)] - (1-\kappa)(2-\phi)]}{\xi[\theta - 1 + \chi(1-\kappa)] + (1-\phi)\chi\kappa} \right\} n. \end{aligned} \tag{2.28}$$

Remark 1. Cong et al. (2021): For the special case that the government is absent, i.e. $\epsilon = \omega = 0$ and $\kappa = 1$, the bgp growth rate for the decentralized economy collapses to $\hat{g}|_{\epsilon=\omega=0,\kappa=1} = \frac{\chi}{\xi(\theta-1)+\chi(1-\phi)}$, *while the bgp growth rate of data-sets reads as* $\hat{g}|_{\epsilon=\omega=0,\kappa=1} = \frac{1-\theta}{\xi(\theta-1)+\chi(1-\phi)}$. For the special case that the government is absent and, additionally, preferences are logarithmic ($\theta = 1$), the bgp growth rate collapses to $\hat{g} = \frac{n}{1-\theta}$. This growth rate is larger compared to Jones (1995) as in our setting the diminishing returns of research labor

$(1-\xi<1)$ are directly offset by data usage within the R&D sector.

Remark 2. Appendix 2.A.2 shows that for the parameter restriction

$$\chi > \hat{\chi} \equiv \frac{\epsilon(\theta-1)\xi}{\kappa-(1-\kappa)\xi} \tag{2.29}$$

the bgp growth rate (2.27) exists and is positive, while the bgp growth rate for data sets, (2.28), exists but is negative.

In other words, (2.29) ensures the convexity of the disutility term χ in the representative household's utility function originating from private data production and governmental surveillance. In contrast to Cong et al. (2021), the restriction is more severe due to the presence of government surveillance activities affecting directly positively ($\xi\epsilon$) as well as negatively ($(1-\kappa)\xi$) the representative household's utility function.

Proposition 2. *As shown in Appendix 2.A.3, for the decentralized economy, on the bgp, the share of employed R&D is constant and reads as*

$$\hat{l}_n(t) \equiv \frac{\hat{L_N}(t)}{L(t)} = 1-\hat{l}_y(t) = \hat{l}_n = \frac{\alpha(1-\xi)}{\left[\frac{g_d^*+n-g^*}{g^*}\right]\epsilon(\theta-1)+\theta+\alpha(1-\xi)+\frac{\rho}{g^*}} \in (0,1), \tag{2.30}$$

where g^* is given by (2.27) and g_d^* by (2.28) and $1-\xi>\phi$. Note that the ladder condition restricts the knowledge spillover effect in order to ensure that $l_r \in (0,1)$ for the empirically plausible value $\theta>1$ (see Jones (2016)).

Further, (2.30) is independent of the tax rate because its positive effect (lower wages in the final goods sector makes employment in the R&D sector more attractive) and the negative effect (higher taxes in the final goods sector reduces the demand for intermediate goods (see (2.18)), thus reduces profits in the intermediate good sector and, hence, leads to lower wages in the R&D sector) via the intermediate goods cancel each other.

2.3.6 Comparative Statics

Before we proceed with deriving the social planner's problem, it is worth deriving some insights from comparative statics.

Proposition 3. *For $\theta>1$, the bgp growth rate (2.27) decreases with the increasing*

importance of data usage in the R&D sector, ξ, while the effect of the increasing importance of dis-utility stemming from private data usage or governmental surveillance, χ on the bgp is undetermined. Finally, if the importance of utility from governmental surveillance ϵ increases, households prefer lower bgp growth rates. In formal terms we have:

$$\frac{\partial \hat{g}}{\partial \xi} = \left\{ \frac{[(1-\theta)(1+\epsilon(1-\phi)) - (1-\kappa)(2-\phi)\chi]n\kappa\chi}{\xi[\theta - 1 + \chi(1-\kappa)] + (1-\phi)\chi\kappa]^2} \right\} < 0, \tag{2.31}$$

$$\frac{\partial \hat{g}}{\partial \chi} = \left\{ \frac{(\theta-1)[\kappa(1+\epsilon(1-\phi)) - \xi(1-\epsilon)(1-\kappa)]n\xi}{[\xi(\theta - 1 + \chi(1-\kappa)) - (1-\phi)\chi\kappa]^2} \right\} \lessgtr 0, \tag{2.32}$$

$$\frac{\partial \hat{g}}{\partial \epsilon} = \left\{ \frac{(1-\theta)n\xi}{\xi[\theta - 1 + \chi(1-\kappa)] - (1-\phi)\chi\kappa]} \right\} < 0. \tag{2.33}$$

Some comments regarding Proposition 3 are in order. First, from inspecting (2.31) we see that an increase of ξ that mirrors the importance of data usage in the innovation of new data algorithms decreases the bgp growth rate. This counter-intuitive result can be explained with a general equilibrium effect. An increase of ξ increases the usage of data at the cost of R&D employment in the production of new algorithms. To produce more data in the present, surveillance and/or private consumption in the present has to go up at the cost of future consumption. Alternatively, households might invest in new algorithms that increase their future consumption potential and also produce more future data as a by-product. Due to the tendency of consumption smoothing (note that $\theta > 1$), the first effect dominates the second, and, hence, households require a lower bgp growth rate if ξ goes up.[3] Next, (2.32) for $\theta > 1$ shows that if the importance of dis-utility from private data use and/or governmental surveillance increases, households accept lower bgp growth rates if the negative impact of governmental surveillance on utility is sufficiently small, i.e. if $1 - \kappa < \frac{(1-\epsilon)\xi}{1+\xi(1-\epsilon)+\epsilon(1-\phi)}$. This finding contrasts and complements recent studies such as Cong et al. (2021), which propose that consumers have to be compensated with higher bgp growth rates for the disutility from private data use. In our setting, where households also derive positive utility from surveillance, i.e. a higher subjective feeling of security, under some

[3] Cong et al. (2021) state on p. 6484 that an increase of ξ increases the gbp growth rate if $\theta > 1$. We believe that this is a typo as their bgp growth rate in fact decreases if ξ goes up, provided that the consumption smoothing motive is sufficiently large, i.e. $\theta > 1$.

conditions, in equilibrium smaller growth rates are required even if governmental surveillance as well as private data misuse is present. Finally, from (2.33) we observe that an increase of ϵ reduces the bgp growth rate as, unsurprisingly, households accept lower bgp growth rates if the importance of the utility-enhancing effect of governmental surveillance increases and the consumption smoothing motive is sufficiently pronounced.

2.4 The Social Planner's Problem

The equilibrium characterized in the decentralized economy is not socially optimal due to, (i) monopolistic competition, (ii) knowledge spillovers in the R&D production function as well as a (iii) reduction of R&D productivity due to government surveillance that is taken as given by the agents as reflected by (2.21) and (5).

In turn, a benevolent social planner maximizes the utility of the representative household subject to the rouse constraint. The latter requires that aggregate net output $Y(t) - \int_0^{N(t)} \hat{\psi} x_{i,t} di$ equals aggregate consumption and governmental surveillance expenditures, $G(t)$:

$$C(t) = c(t)L(t) = Y(t) - \int_0^{N(t)} \hat{\psi} x_{i,t} di - G(t), \tag{2.34}$$

where $G(t) = \frac{d_g(t)L(t)}{m(t)}$. Given $N(t)$, the social planner solves a static optimization problem that is at each point in time t she chooses the optimal level of intermediate goods given $N(t)$. In other words, she optimizes the following resource constraint:

$$C(t) + G(t) = L_Y(t)^{1-\alpha} \int_0^{N(t)} x_{i,t}^{\alpha} di - \int_0^{N(t)} \hat{\psi} x_{i,t} di \equiv Y_n(t). \tag{2.35}$$

From this optimization problem, we derive the optimal net output, $Y_n(t)$ as:

$$Y_n(t) = (1-\alpha) L_Y(t) N(t) \left(\frac{\alpha}{\hat{\psi}} \right)^{\frac{\alpha}{1-\alpha}}. \tag{2.36}$$

Unsurprisingly, given the same level of technology and labor input, compared to the decentralized economy, the output is larger. The difference is due to monopolistic competition in the intermediate goods sector in the decentralized economy. Given (2.36), the social planer solves the following problem:

$$\max_{\{c(t),d_c(t),G(t),l_y(t)\}} \int_0^\infty \exp[-(\rho-n)t]\left[\frac{[G(t)^\epsilon c(t)^{1-\epsilon}]^{1-\theta}}{1-\theta} - \iota[d_c(t)^\kappa G(t)^{1-\kappa}]^\chi\right] dt$$

subject to

$$c(t) \geq 0,$$

$$G(t) \geq 0,$$

$$d_c(t) \geq 0,$$

$$m(t) \geq 0,$$

$$\dot{N}(t) = \bar{\eta}\left[\frac{G(t)}{Y_n(t)}\right]^{-\omega} N(t)^\phi \left[(d_c(t)L(t))^\beta (G(t)m(t))^{1-\beta}\right]^\xi L_N(t)^{1-\xi}, \tag{2.37}$$

$$\frac{\dot{d}_c(t)}{d_c(t)} \leq \frac{\dot{c}(t)}{c(t)}, \frac{\dot{d}_g(t)}{d_g(t)} \leq \frac{\dot{G}(t)}{G(t)}$$

$$c(t) + \frac{G(t)}{L(t)} = (1-\alpha) l_y(t) N(t) \left(\frac{\alpha}{r(t)}\right)^{\frac{\alpha}{1-\alpha}}, \tag{2.38}$$

$$l_y(t) + l_n(t) = 1. \tag{2.39}$$

Equation (2.39) represents the labor market clearing condition, while equation (2.38) shows the simplified resource constraint. After solving the social planner's problem in Appendix 2.A.4, we can summarize our main findings with the following proposition:

Proposition 4. *In Appendix 2.A.4 it is shown that the bgp growth rates coincide with the bgp growth rates for the decentralized economy presented in Proposition 1. In other words, we have*

$$g^* = \hat{g} \leq \frac{n[\xi(1-(1-\kappa)\chi) - \chi\kappa] - \xi\rho}{\xi[1-\phi-\chi(1-\kappa)] - \kappa(1-\phi)\chi}. \tag{2.40}$$

One remark regarding the last (in)equality shown in Proposition 3 is in order. Although the social optimal and decentral bgp growth rates are identical, nevertheless we have to impose an upper limit on the social optimal growth rate in order to guarantee that the fraction of R&D workers does not exceed the value of one.

From an intuitive point of view, the result presented in Proposition 4 directly relates to

Jones (1995), who shows that a focus on growth rates per se is not sufficient to fully describe a country's economic performance.[4] For instance, an increase of governmental surveillance decreases the efficiency of R&D, thus leading to an immediate reduction of the growth of new ideas. However, more surveillance implies more governmental data that in turn can be used to generate new ideas that manifest themselves in new intermediate goods and an increase in the final output. Hence, in the medium run, growth rates return to the initial growth rates if the positive effect of additional data exactly offsets the negative effect of surveillance on creativity, as it is the case in our model.[5]

However, as we show below, we have *level effects*, i.e. the social optimal fraction of labor employed in the R&D sector is larger compared to the fraction of R&D workers employed in the decentralized economy. This, in turn, leads to a sub-optimal overuse of data in the decentralized economy. We show this in Proposition 5:

Proposition 5. *In Appendix 2.A.5, it is shown that the bgp fraction of R&D labor in the centrally planned economy is constant and given condition (2.29) holds, the social fraction of R&D employed labor exceeds the fraction of R&D labor in the decentralized economy. The bgp social fraction of R&D labor can be formally derived as*

$$l_n^*(t) = l_n^* = \frac{1}{1 + \frac{(1-\phi)(\xi-\chi\kappa)-\chi(1-\kappa)\xi}{(1-\xi)\xi} + \frac{\xi\rho-n[\xi-(\kappa-(1-\kappa)\xi)\chi]}{g(1-\xi)\xi}}. \tag{2.41}$$

Remark 3. In Appendix 2.A.5 it is also verified that given the inequality (2.40) holds, we have $l^* \in (0,1)$, which is automatically fulfilled for the empirically plausible case $\theta > 1$.

---

[4]The not surprising finding that the bgp growth rates between the decentralized and centralized economy coincide is also made by e.g. Cong et al. (2021) and Jones and Tonetti (2020), who both use a semi-endogenous growth setting with data usage.

[5]This can be seen directly inspecting equation (2.21) together with (5). As on the bgp, $G(t)$ grows with the same rate as $Y(t)$, the fraction $\frac{G(t)}{Y(t)}$ remains constant. Hence on the bgp, η in equation (5) does not grow. Hence, off the bgp, we expect transitional dynamics of $\eta(t)$, which dynamics is governed by ω. Hence, even if we modify (5) to $\eta(t) = \bar{\eta}\left(\frac{G(t)}{Y(t)}\right)^{-\omega} = \bar{\eta}\left(\frac{(d_g(t)L(t))^{-\omega_1}}{(m(t)Y(t))^{-\omega_2}}\right)$, with $\omega_1 \in [0,1)$ with $\omega_2 \in [0,1)$ and $\omega_1 \neq \omega_2$, on the bgp, the gap between $G(t)$ and $Y(t)$ is constant (and not zero as in our case for $\omega_1 = \omega_1 = \omega$). In other words, on the bgp, the finding that $\hat{g} = g^s$ remains valid even for the case that $\omega_1 \in [0,1)$, $\omega_2 \in [0,1)$ and $\omega_1 \neq \omega_2$.

2.5 Rent-Seeking Government

In the preceding sections, we have contrasted the decentralized with the central planning solution. We can interpret the social planner as a benevolent government. We found that bgp growth rates are the same, but the allocation of the labor force between the R&D sector and the final goods sector differs due to monopolistic competition in the decentralized economy's intermediate goods sector. This results in a sub-optimal overuse of data in the decentralized economy and hence excessive surveillance activities from a social's planner point of view. To obtain the socially optimal labor market allocation for the decentralized economy, a wage subsidy scheme with subsidy rate $s(t)$ has to be imposed in the intermediate goods sector, i.e.

$$p_N(t)(1-\xi)\eta N(t)^{\phi}(D(t))^{\xi}L_N(t)^{-\xi} = w_N(t)s(t). \tag{2.42}$$

A suited tax scheme brings down the sub-optimal use of data, and reduces surveillance activities by altering the R&D labor share towards an optimal level. In turn, as shown by Cong et al. (2021), a data tax addresses the miss-allocation in the labor market. This is also the case in our model. The same applies to the taxation of the final goods market with tax rate τ (see (2.11)). Hence, a tax rate greater than zero reduces welfare further. Therefore, the benevolent planner sets the tax rate to zero.

However, what happens with the tax rate if the government is self-interested, and how does this impact the economy? Allen (2011, p. 15) notes that "economic success is the result of secure property rights, low taxes, and minimal government. Arbitrary government is bad for growth because it leads to high taxes [...] and rent-seeking." In the context of our model, the tax rate $\tau(t)$ can be interpreted as a characteristic of an authoritarian state, where self-interested elites control the government and use tax revenues as an additional source for financing surveillance activities in order to consolidate their hold on power (see 2.11). For simplicity, we first assume that the elites are myopic and have a static objective function:

$$W(t) = \iota \ln[\tau Y(t)] + (1-\iota)\ln[c(t)], \tag{2.43}$$

where $\iota \in [0,1]$ reflects the weight that the elites place on surveillance with direct expenses

of the representative household in terms of consumption loss. We proceed by showing that on the bgp, consumption is a constant fraction of output:

$$\frac{c(t)}{y(t)} = [1-\tau]\underbrace{\left[1-\alpha^2-\frac{\xi(1-\alpha)(1-\beta)}{(1-\xi)}\left(\frac{l_n(t)}{l_y(t)}\right)\right]}_{>0}, \tag{2.44}$$

as $\frac{l_n(t)}{l_y(t)}$ is constant on the bgp. The latter results help to ensure that the bgp tax rate that is chosen by the myopic government is constant. Next, substitution of (2.44) and (2.19) in (2.43), and after having dropped all exogenous and pre-determined variables, we have

$$W(t) = \iota \ln[\tau] + (1-\iota)\ln[1-\tau] + \left(\frac{\alpha}{1-\alpha}\right)\ln[1-\tau]. \tag{2.45}$$

Finally, differentiating (2.45) with respect to the tax rate τ yields:

$$\tau = \iota(1-\alpha). \tag{2.46}$$

Thus, the tax rate chosen by the ruling elites is stationary and depends on two parameters: first, on the ruling elites' chosen degree ι of surveillance, and, second, on the intensity α of intermediate goods in final good production. (2.46) shows that the tax rate is increasing in the degree ι, i.e. the more important is surveillance for the ruling elites, the higher is the chosen tax rate. Moreover, the tax rate is decreasing in α, because a larger α increases the effect of the tax wedge on the production of intermediate goods as visualized in (2.18).

Now consider the case of forward-looking, dynamic optimization elites. These elites choose τ_t in order to maximize

$$W = \int_0^\infty \exp[-(\rho-n)t][\iota \ln[\tau Y(t)] + (1-\iota)\ln[c(t)]dt. \tag{2.47}$$

Again using (2.44) and (2.19) and noting the fact that the labor market outcome is not affected by the tax rate, dropping exogenous variables, we can rewrite (2.47) as

$$W = \frac{1}{\rho-n}\left[\iota \ln[\tau] + (1-\iota)\ln[1-\tau] + \left(\frac{\alpha}{1-\alpha}\right)\ln[1-\tau]\right]. \tag{2.48}$$

Hence, the tax rate that is chosen by the government under a dynamic rent-seeking regime is the same as that under a static rent-seeking regime and given by (2.46). We summarize the result with the following proposition.

Proposition 6. *Under some conditions, the tax rate chosen by a myopic government under static rent-seeking corresponds to the tax rate under dynamic rent-seeking.*

Hence, if the government is benevolent, the chosen tax rate τ is zero. While a tax rate greater than zero leaves the bgp growth rate unaffected, it nevertheless increases the R&D labor market distortion further that already exists due to data overuse.[6]

2.6 Calibration

Similar to Jones (1995), in our model the sub-optimal allocation of R&D labor in the decentralized economy is due to monopolistic competition in the production of intermediate goods. To compensate for the lower production and usage of intermediate goods, the final good producers employ more labor that in turn crowds out R&D labor. However, in our model the R&D labor market distortion is less severe than in Jones (1995), as data can be used as a direct substitute for R&D labor to produce new algorithms. Thus, like in Cong et al. (2021), the crowding-out of R&D labor is paralleled by a socially sub-optimal crowding-in of data.

To obtain a better understanding of the steady-stage labor market allocations, we perform a calibration exercise. In other words, we use calibrated values that are consistent with the relevant literature (see table (2.1) for an overview) and calculate the difference between labor that is allocated in the R&D sector between the social planner's problem and the decentralized economy by varying the parameter space $\{\phi, \kappa, \epsilon, \xi\}$[7].

The result of this exercise is presented in the first row of figure (2.3). In this exercise, we focus in particular on κ that weights the household's relative disutillity from private data misuse and from governmental surveillance.[8] We find that, (i), if household's derive

[6]The fraction of R&D labor in the decentralized economy decreases with decreasing α. In turn, τ is increasing with a decreasing α.

[7]Note that κ is calibrated in the way to insure a positive bgp growth rate (see Remark 2).

[8]For $\kappa = 0$ ($\kappa = 1$) household's members suffer only from governmental surveillance (private data misuse).

Table 2.1: Summary of parameterization for the baseline economy

Variable	Description	Value	Source
α	Intensity of intermediate goods in production	$\frac{1}{3}$	Standard
ϵ	Utility weight of governmental surveillance	0.2	Discretionary
θ	Relative risk aversion of consumption in utility	2.5	Standard
ξ	Contribution of data in innovation frontier	0.5	Cong et al. (2021)
χ	Disutility weight of data misuse or privacy violation	1.5	Cong et al. (2021)
ϕ	Degree of knowledge spillover in innovation frontier	0.85	Cong et al. (2021)
ρ	Subjective discount factor	0.03	Standard
n	Population growth rate	0.02	Standard
κ	Relative disutillity from private data misuse and from governmental surveillance	0.80	Discretionary

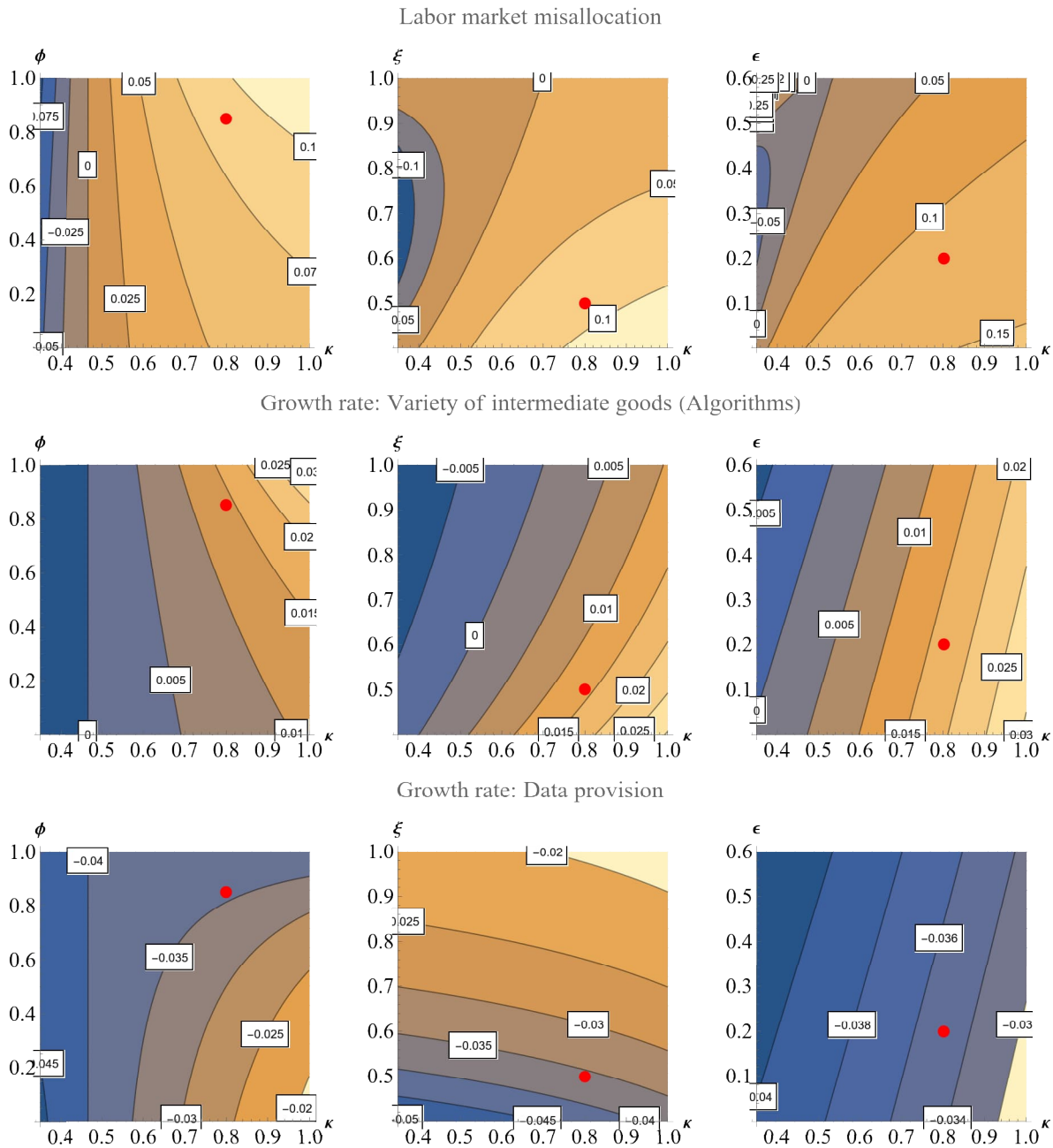
sufficient utility from governmental surveillance (high ϵ), the R&D misallocation of labor tends to zero, even if households put a large weight on governmental surveillance relative to private data misuse (small κ). The reason is that for a small κ and large ϵ, the social planner optimally allocates more people in the final goods sector, thereby reducing the R&D labor misallocation between the social planner's solution and the solution of the decentralized economy. On the other side, R&D labor misallocation increases with decreasing ϵ for constant κ. Moreover, (ii), we also observe that R&D misallocation is particularly pronounced if the degree of knowledge spillover ϕ is large and/or the importance of data for the development of new algorithms ξ is low. Importantly, we find that for a specific combination of utility and disutility from governmental surveillance, the decentralized labor market allocation coincides with the social optimal allocation. Hence, if households also derive utility from governmental surveillance, it is not necessarily the case that the decentralized data economy overuses data at the cost of an insufficiently small R&D sector. This finding complements the work of Cong et al. (2021).

To complement our exercise, the second row of figure (2.3) shows the equilibrium growth rates of the variety of intermediate goods which can be interpreted as algorithms. These figures confirm our findings summarized with Propositions 1 and 3. For empirically plausible values, we find that the growth rate of intermediate goods ranges between zero and 3 percent.

2.7 Conclusion

In this paper, we develop a political economy model to illustrate the trade-off between surveillance and data availability in an autocracy. As in traditional autocracies, surveillance

Figure 2.3: Decentralized economy vs. social planner's problem: Differences in R&D labor allocation



Notes: The first row shows the difference of R&D labor allocation between the central planner's problem and the decentralized economy. For the case of the first row, a positive (negative) value indicates a sub-optimal small (large) R&D sector for the decentral data economy. The printed numbers quantify the exact differences. The second (third) row computes the bgp growth rates of the variety of intermediate goodss (data provision) for different vales $\{\phi, \xi, \epsilon, \kappa\}$. For the second and third row: Yellow ocker shows larger bgp growth rates, while blue color reflects smaller bgp growth rates. The printed numbers reflect the bgp growth rates. The calibration ensures the convexity of the consumer's disutility function as well as the existence of the bgp. Finally, a red dote in each figure indicates the benchmark calibration based on Table 2.1.

and repression have a negative effect on innovation, as they render scientists and researchers less creative and productive. In modern informational autocracies, surveillance can however also be used to gather and bundle large amounts of data, which can then be used as an input – provided by the state – to accelerate innovation in data-intensive fields, such as artificial intelligence. Under certain conditions, this second effect might outweigh the first effect, rendering autocracies more productive than democracies when it comes to generating new blueprints and applied research, for example in the form of more precise and better performing algorithms.

Which effect ultimately prevails depends to a large extent on the future technological evolution of the field of artificial intelligence. If the technology remains roughly at the same level during the next couple of years, then informational autocracies such as China might well have an advantage in the new systemic competition of the 21st century. By leveraging large amounts of data – that might at least partially gathered through state surveillance – they will be better able to put applied solutions to the market than democratic states, where firms and research institutions are hampered by more restrictive data and privacy regulations.

If, however, cutting-edge scientist in some of the leading research institutions in the world come up with another breakthrough and paradigm shift in the field, the world's democracies might win the race once again, as – at least for now – most cutting-edge research institutions in the field of AI are still located in democracies, and in particular in the United States.

2.8 Disclosure statement

The authors declare that they have no relevant or material financial interests that relate to the research described in this paper. Moreover, the authors do not use generative AI and AI-assisted technologies

2.9 Bibliography

Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2019). Democracy Does Cause Growth. *Journal of Political Economy*, 127(1):47–100.

Acemoglu, D. and Robinson, J. A. (2012). *Why Nations Fail: The Origins of Power, Prosperity and Poverty*. Crown Publishers, New York, 1st ed. edition.

Aghion, P., Alesina, A., and Trebbi, F. (2007). Democracy, Technology and Growth. *NBER Working Paper*, No. 13180.

Agrawal, A., Gans, J., and Goldfarb, A. (2018). *Prediction Machines: The Simple Economics of Artificial Intelligence*. La Vergne: Ingram Publisher Services.

Allen, R. C. (2011). *Global Economic History: A Very Short Introduction*. Oxford University Press.

Beraja, M., Kao, A., Yang, D., and Yuchtman, N. (2023a). Exporting the Surveillance State via Trade in AI. *NBER Working Paper*.

Beraja, M., Kao, A., Yang, D. Y., and Yuchtman, N. (2023b). AI-tocracy. *The Quarterly Journal of Economics*, 138(3):1349–1402.

Beraja, M., Yang, D. Y., and Yuchtman, N. (2023c). Data-intensive Innovation and the State: Evidence from AI Firms in China. *The Review of Economic Studies*. rdac056.

Carayannis, E. and Campbell, D. (2014). Developed democracies versus emerging autocracies: arts, democracy, and innovation in Quadruple Helix innovation systems. *Journal of Innovation and Entrepreneurship*, 3(12).

Cockburn, I., Henderson, R., and Stern, S. (2018). The Impact of Artificial Intelligence on Innovation. *NBER Working Paper No. 24449*.

Cong, L. W., Wei, W., Xie, D., and Zhang, L. (2022). Endogenous growth under multiple uses of data. *Journal of Economic Dynamics and Control*, 141:104395. Markets and Economies with Information Frictions.

Cong, L. W., Xie, D., and Zhang, L. (2021). Knowledge accumulation, privacy, and growth in a data economy. *Management Science*, 67(10):6480–6492.

Dixit, A. and Stiglitz, J. (1977). Monopolistic Competition and Optimum Product Diversity. *American Economic Review*, 67(3):297–308.

Domingos, P. (2015). *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake our World*. London: Penguin Books.

Gao, Y., Zang, L., Roth, A., and Wang, P. (2017). Does democracy cause innovation? An empirical test of the popper hypothesis. *Research Policy*, 46(7):1272–1283.

Ghardallou, W. and Sridi, D. (2020). Democracy and Economic Growth: a Literature Review. *Journal of the Knowledge Economy*, 11:982–1002.

Gomez, A. and Canales, A. F., editors (2015). *Science Policies and Twentieth-Century Dictatorships: Spain, Italy and Argentina*. London: Routledge.

Graham, L. R. (1987). *Science, Philosophy, and Human Behavior in the Soviet Union*. New York: Columbia University Press.

Graham, L. R. (1993). *Science in Russia and the Soviet Union: A Short History*. New York: Cambridge University Press.

Guriev, S. and Treisman, D. (2022). *Spin Dictators: The Changing Face of Tyranny in the 21st Century*. Cambridge, MA: Princeton University Press.

Habich-Sobiegalla, S. and Kostka, G. (2022). Sharing is Caring: Willingness to Share Personal Data Through Contact Tracing Apps in China, Germany, And The US. *Information, Communication & Society*, pages 1–28.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep Residual Learning for Image Recognition. *CVPR*, pages 770–778.

Hey, T. (2009). *The Fourth Paradigm: Data-intensive Scientific Discovery*. Microsoft Research.

Hinton, G., Osindero, S., and Teh., Y. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7):1527–1554.

Hinton, G. E. and Salakhutdinov, R. R. (2006). Reducing the Dimensionality of Data with Neural Networks. *Science (New York, N.Y.)*, 313(5786):504–507.

Huang, H. and Xu, C. (1999). Institutions, Innovations, and Growth. *Economics of Transition and Institutional Change*, 26(3):335–362.

Jamali, K., Wandschneider, K., and Wunnava, P. (2007). The effect of political regimes and technology on economic growth. *Applied Economics*, 39(11):1425–1432.

Jones, C. (1995). R & D-Based Models of Economic Growth. *Journal of Political Economy*, 103(4):759–784.

Jones, C. I. (2016). Life and growth. *Journal of Political Economy*, 124(2):539–578.

Jones, C. I. and Tonetti, C. (2020). Nonrivalry and the Economics of Data. *American Economic Review*, 110(9):2819–2858.

Josephson, P. R. (2005). *Totalitarian Science and Technology*. Humanity Books.

Karpa, D., Klarl, T., and Rochlitz, M. (2022). *Artificial Intelligence, Surveillance, and Big Data*, pages 145–172. Springer International Publishing, Cham.

King, G., Pan, J., and Roberts, M. E. (2013). How Censorship in China Allows Government Criticism but Silences Collective Expression. *American Political Science Review*, 107(2):326–343.

King, G., Pan, J., and Roberts, M. E. (2014). Reverse-engineering Censorship in China: Randomized Experimentation and Participant Observation. *Science*, 345(6199):1251722.

Klomp, J. and de Haan, J. (2012). Political Regime and Human Capital: A Cross-Country Analysis. *Social Indicators Research*, 111:335–362.

Knutsen, C. H. (2013). Democracy, State Capacity, and Economic Growth. *World Development*, 43(1):1–18.

Knutsen, C. H. (2015). Why Democracies Outgrow Autocracies in the Long Run: Civil Liberties, Information Flows and Technological Change. *Kyklos*, 68(3):357–384.

Kostka, G. (2019). China's Social Credit Systems And Public Opinion: Explaining High Levels of Approval. *New Media & Society*, 21(7):1565–1593.

Krizhevsky, A., Sutskever, I., and Hinton, G. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25:1097–1105.

Lee, K.-F. (2018). *AI superpowers: China, Silicon Valley, and the new world order*. Houghton Mifflin Harcourt, Boston.

Libman, A. (2012). Democracy, size of bureaucracy, and economic growth: evidence from Russian regions. *Empirical Economics*, 43:1321–1352.

Murtin, F. and Wacziarg, R. (2014). The Democratic Transition. *Journal of Economic Growth*, 19(1):141-181.

Olson, M. (1982). *The Rise and Decline of Nations*. New Haven: Yale University Press.

Piatek, D., Szarzec, K., and Pilc, M. (2013). Economic Freedom, Democracy and Economic Growth: a Causal Investigation in Transition Countries. *Post-Communist Economies*, 25(3):267–288.

Pozuelo, J. R., Slipowitz, A., and Vuletin, G. (2016). Democracy Does not Cause Growth: the Importance of Endogeneity Arguments. *Inter-American Development Bank Working Paper*, page No. 694.

Qiang, X. (2019). The Road to Digital Unfreedom: President Xi's Surveillance State. *Journal of Democracy*, 30(1):53–67.

Roberts, M. E. (2018). *Censored: Distraction and Diversion Inside China's Great Firewall*. Princeton: Princeton University Press.

Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5):71–102.

Rosenberg, D. and Tarasenko, G. (2020). Innovation for despots? How dictators and democratic leaders differ in stifling innovation and misusing natural resources across 114 countries. *Energy Research & Social Science*, 68(101543).

Shadmehr, M. and Bernhardt, D. (2015). State Censorship. *American Economic Journal: Microeconomics*, 7(2):280–307.

Silve, F. and Plekhanov, A. (2018). Institutions, Innovation and Growth: Evidence from Industry Data. *Economics of Transition and Institutional Change*, 26(3):335–362.

Spitz, M. (2017). *Daten - das Öl des 21. Jahrhunderts? Nachhaltigkeit im digitalen Zeitalter*. Hamburg: Hoffman und Campe.

Stokes, R. G. (2000). *Constructing Socialism: Technology and Change in East Germany, 1945-1990*. Baltimore: John Hopkins University Press.

Strittmatter, K. (2020). *We Have Been Harmonized: Life in China's Surveillance State*. New York: Custom House.

Taffel, S. (2023). Data and Oil: Metaphor, Materiality and Metabolic Rifts. *New Media & Society*, 25(5):980–998.

Tang, R. and Tang, S. (2018). Democracy's Unique Advantage in Promoting Economic Growth: Quantitative Evidence for a New Institutional Theory. *Kyklos*, 71(4):642–666.

Tebaldi, E. and Elmslie, B. (2008). Institutions, Innovation and Economic Growth. *MPRA Working Paper*, No. 9683.

Truex, R. (2017). The Myth of Democratic Advantage. *Studies in Comparative International Development*, 52(3):261-277.

Veldkamp, L. L. (2005). Slow boom, sudden crash. *Journal of Economic Theory*, 124(2):230–257.

Wagner, C. S., Cai, X., and Mukherjee, S. (2020). China's scholarship shows atypical referencing patterns. *Scientometrics*, 124(3):2457–2468.

Wagner, C. S., Zhang, L., and Leydesdorff, L. (2022). A discussion of measuring the top-1% most-highly cited publications: quality and impact of Chinese papers. *Scientometrics*, 127(4):1825–1839.

Wang, Q.-J., Feng, G.-F., Wang, H.-J., and Chang, C.-P. (2021). The Impacts of Democracy on Innovation: Revisited Evidence. *Technovation*, 108:102333.

Zuazu, I. (2019). The Growth Effect of Democracy and Technology: An Industry Disaggregated Approach. *European Journal of Political Economy*, 56:115–131.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Public Affairs.

2.A Appendix

2.A.1 Proof of Proposition 1: BGP growth rate for the decentralized economy

In this Appendix, we derive the bgp growth rates (2.27) and (2.28). In our model, there is perfect labor mobility between final goods and the research sector. In equilibrium, mobility between sectors comes to a halt if $w_N(t) = w_Y(t)$ or equivalently,

$$p_N(t)(1-\xi)\eta N(t)^{\phi}(D(t))^{\xi}L_N(t)^{-\xi} \quad = \quad (1-\tau)(1-\alpha)\frac{Y(t)}{L_Y(t)}. \tag{49}$$

Next, we want to pin down the operating profits of intermediate goods producers. We start with,

$$p_N(t) = \pi(t) + \dot{p_N}(t) \tag{50}$$

which is the no-arbitrage condition that the market value of a patent $p_N(t)$ has to meet in equilibrium. In the absence of asset price bubbles (which we assume), condition (50) says that the market value of a patent equals the fundamental value of the patent, i.e. the present value of the expected future accounting profits from using the new invented algorithm in the intermediate goods sector. Hence, we have

$$p_N(t) = \int_t^{\infty} \exp[-(\Gamma(\tau))]\pi(t)d\tau, \tag{51}$$

with $\Gamma(\tau) = \int_t^{\tau} r(s)ds$ so that the discount rate is the market interest rate. With perfect foresight and the absence of uncertainty, the no-arbitrage condition can be handled as a differential equation for $p_N(t)$. The solution to the differential equation (50) is given by (51). In bgp, interest rates are constant, i.e. $r(t) = r$, while the operating profits have to grow with rate n (which becomes clear in the following analysis). Thus, (51) reduces to

$$p_N(t) = \frac{\pi(t)}{r-n}, \tag{52}$$

where operating profits are obtained as

$$\pi(t) = (1-\tau)(1-\alpha)\alpha\frac{Y(t)}{N(t)}. \tag{53}$$

With a constant interest rate r, $\frac{Y(t)}{N(t)}$ grows with a constant rate n (see (2.19)). Thus, in bgp, operating profits also grow with rate n (see equation (53)) as the price for blueprints does:

$$p_N(t) = \frac{(1-\tau)(1-\alpha)\alpha Y(t)}{(r-n)N(t)} \tag{54}$$

Thus, using (54) in (49) yields:

$$\alpha(1-\xi)\eta N(t)^{\phi-1}(D(t))^{\xi}L_N(t)^{-\xi} = \frac{r-n}{L_Y(t)}. \tag{55}$$

Now, we are prepared to derive the growth rate of $N(t)$ and $d_c(t)$. Writing (55) in growth rates gives (and assuming that r is constant on the the bgp):

$$(\phi+\omega-1)\frac{\dot{N}(t)}{N(t)} + \xi\beta\frac{\dot{d}_c(t)}{d_c(t)} + (\xi(1-\beta)-\omega)\frac{\dot{d}_g(t)}{d_g(t)} + \omega\frac{\dot{m}(t)}{m(t)} + n = 0. \tag{56}$$

The free-entry condition of private data (2.25) can be also written in growth rates, yielding the growth rate of the price for private data sets:

$$\frac{\dot{p}_{dc}(t)}{p_{dc}(t)} = (\xi\beta-1)\frac{\dot{d}_c(t)}{d_c(t)} + (\phi+\omega)\frac{\dot{N}(t)}{N(t)} + (\xi(1-\beta)-\omega)\frac{\dot{d}_g(t)}{d_g(t)} + \omega\frac{\dot{m}(t)}{m(t)} + n. \tag{57}$$

Similarly, the free-entry condition of governmental data (2.26) can be written in growth rates to obtain the growth rate of the price for governmental data sets from surveillance activities:

$$\frac{\dot{p}_{dg}(t)}{p_{dg}(t)} = (\xi(1-\beta)-1-\omega)\frac{\dot{d}_g(t)}{d_g(t)} + (\phi+\omega)\frac{\dot{N}(t)}{N(t)} + \xi\beta\frac{\dot{d}_c(t)}{d_c(t)} + \omega\frac{\dot{m}(t)}{m(t)} + n. \tag{58}$$

Combining (2.25) and (2.26), we can re-express this relationship in growth rates as

$$\frac{\dot{p}_{dc}(t)}{p_{dc}(t)} = \frac{\dot{p}_{dg}(t)}{p_{dg}(t)} + \frac{\dot{d}_g(t)}{d_g(t)} - \frac{\dot{d}_c(t)}{d_c(t)}. \tag{59}$$

As in the steady-state the data markets clear, we must have $p_{dc}(t) = p_{dg}(t)$. Hence, on

the bgp, governmental data must grow with the same rate as the prices of private data. Therefore, from (59) we find that on the bgp, the growth rates of private and governmental data sets are equal and grow with the constant rate $\hat{g}_d$, i.e.

$$\hat{g}_d = \frac{\dot{d}_g(t)}{d_g(t)} = \frac{\dot{d}_c(t)}{d_c(t)}. \tag{60}$$

Using (2.8) and (2.9) together with (57), we obtain

$$\frac{\dot{c}(t)}{c(t)} = \frac{1}{\tilde{\theta}} \left(\frac{\dot{p}(t)_{dc}}{p_{dc}(t)} - \chi \left[(\kappa - \frac{1}{\chi}) \frac{\dot{d}_c(t)}{d_c(t)} + (1-\kappa) \frac{\dot{G}(t)}{G(t)} \right] + (1-\theta)\epsilon \frac{\dot{G}(t)}{G(t)} \right) \tag{61}$$

Expressing $G(t)$ in growth rates, using this and (60) in (61), and, moreover, exploiting the fact that on the bgp, per capita consumption $c(t)$ and $N(t)$ grows with the same rate as $\hat{g}$, we arrive at

$$\hat{g}[1-(1-\theta)-\phi+(1-\kappa)\chi] = [\xi - \kappa\chi]\hat{g}_d + a_1 n \tag{62}$$

with $a_1 \equiv [1-\omega+(1-\theta)\epsilon-\chi(1-\kappa)]$. Note that in the absence of the government, we have that $a_1 = 1$.

To obtain (2.27) and (2.28) in Proposition 1, we make use of two equations, namely (56) and (62). We have two equations with two unknowns, $\hat{g}$ and $\hat{g}_d$. Inserting (62) in (56) delivers (2.27) and (2.28). Moreover, from (2.19), we find that output grows with rate $\hat{g}+n$. Together with the goods market clearing condition $Y(t) = C(t) + G(t)$, this implies that $C(t) = c(t)L(t)$ and $G(t)$ have to grow with rate $\hat{g}+n$, while $y(t)$, $c(t)$ and $N(t)$ each grow with the rate $\hat{g}$.

2.A.2 Proof of Remark 2: Existence of the bgp

To make sure that a bgp exists, two conditions have to be fulfilled. The first has to make sure that $\hat{g} > 0$ which is satisfied as long as the parameter restriction (2.29) is fulfilled. The second condition requires that the bgp growth rate also satisfies the transversality condition (2.10) under balanced growth. Along the bgp, (2.10) considerably simplifies. As along as

long as $r > \hat{g} + n$, along the bgp, (2.10) is satisfied. On the bgp, using (2.8) we find that $\theta\hat{g} + \rho = r - n$. This implies that as long as $\rho > \hat{g}(1-\theta)$ the bgp growth rate (2.27) derived with Proposition 1 fulfills the tvc. For the empirically plausible value $\theta > 1$ condition $\rho > \hat{g}(1-\theta)$ is automatically satisfied as $\rho > 0$. Hence, for $\theta > 1$ the bgp growth rate (2.27) is unique and exists as long as the parameter restriction (2.29) holds.

2.A.3 Proof of Proposition 2: BGP labor allocations for the decentralized economy

We start with the insight that the ratios employed in the final goods and research sectors are constant on the bgp. The proof is simple. Using the full employment condition, we have $L_Y(t) = L(t) - L_N(t)$ or expressed in ratios $l_Y(t) = 1 - l_n(t)$ with $L_Y(t) = l_y(t)L(t)$ and $L_N = l_n(t)L(t)$. If on the bgp, $l_y(t)$ and/or $l_n(t)$ grow with a constant rate (as other endogenous variables do), it might be that $l_y(t)$ become larger than one. On contrary, if one of the fractions grow with a negative rate on the bgp, the fractions might reach zero (or even become negative), which implies that all people work either in the final goods sector but do not innovate, or all people innovate but do not produce any final goods. These scenarios obviously contradicts with Proposition 1. Hence, to ensure that both sectors can produce on the bgp, wages have to equalize across sectors and this implies that the fractions $l_y(t)$ and $l_n(t)$ have to be constant on the bgp.

Next, we determine the constant fractions of the research and final goods sectors for the decentralized economy. Using the full employment condition $L_Y(t) = L(t) - L_N(t)$ in (49) together with (53) and (54) delivers on the bgp

$$\begin{aligned} L_Y(t)\alpha(1-\xi)g^* &= (r-n)L_N(t), \\ \alpha(1-\xi)g^*\left(\frac{L_Y(t)}{L(t)-L_Y(t)}\right) &= r-n. \end{aligned} \tag{63}$$

Moreover, inserting (63) in (2.8), on the bgp, we find

$$\begin{aligned} g^* &= \frac{1}{\tilde{\theta}}\left[\alpha(1-\xi)g^*\left(\frac{L_Y(t)}{L(t)-L_Y(t)}\right) - \rho + (1-\theta)\epsilon\left(g_d^* + n\right)\right] \\ \Rightarrow\quad l_n(t) &\equiv \frac{L_R(t)}{L(t)} = \hat{l}_n = \frac{\alpha(1-\xi)}{\left[\frac{g_d^*+n-g^*}{g^*}\right]\epsilon(\theta-1) + \theta + \alpha(1-\xi) + \frac{\rho}{g^*}}, \end{aligned} \tag{64}$$

where g^* is given by (2.27) and g_d^* by (2.28). Finally, $\hat{l}_y = 1 - \hat{l}_n$.

2.A.4 Proof of Proposition 3: Bgp growth rates for the socially planned economy

The corresponding current-value Hamiltonian for the social planner's problem is given by

$$\begin{aligned}\mathcal{H}(d_c(t), G(t), N(t), l_n(t)\lambda_1(t), \lambda_2(t)) &:= \left[\frac{[G(t)^{\epsilon} c(t)^{1-\epsilon}]^{1-\theta}}{1-\theta} - \iota[d_c(t)^{\kappa} G(t)^{1-\kappa}]^{\chi}\right] \\ &+ \lambda_1(t)\left[\left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\alpha}}(1-\alpha)N(t)l_y(t) - c(t) - \frac{G(t)}{L(t)}\right] \\ &+ \lambda_2(t)\left[\bar{\eta}\left[\frac{G(t)}{L(t)(1-l_n(t))\left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\alpha}}}\right]^{-\omega} N^{\omega+\phi}[d_c(t)^{\beta}(m(t)G(t))^{1-\beta}]^{\xi} l_n(t)^{1-\xi} L(t)\right]\end{aligned} \tag{65}$$

The derivation of the necessary first order conditions is straightforward and skipped here in order to safe space[9]. We proceed by showing with a brief sketch that the bgp growth rate of the decentralized economy corresponds to the bgp growth rate of the centralized economy. Writing the necessary first order condition for consumption $c(t)$ in growth rates yields

$$(1-\theta)\left[\epsilon\frac{\dot{G}(t)}{G(t)} + (1-\epsilon)\frac{\dot{c}(t)}{c(t)}\right] = \frac{\dot{\lambda}_1(t)}{\lambda_1(t)} + \frac{\dot{c}(t)}{c(t)} \tag{66}$$

Moreover, reformulating the necessary first order condition for $l_n(t)$ in growth rates, we arrive at

$$\frac{\dot{\lambda}_1(t)}{\lambda_1(t)} = \frac{\dot{\lambda}_2(t)}{\lambda_2(t)}. \tag{67}$$

Thus, on the bgp, the shadow prices grow with the same rate. Next, writing the necessary first order condition for $G(t)$ or $d_c(t)$ in growth rates, using (67) in (66), we arrive at a bgp growth rate g^* that is identical to the bgp growth rate $\hat{g}$ given in (2.27). Using the necessary first order conditions for $N(t)$ and $l_n(t)$, on the bgp, we obtain

$$g_d^* = \frac{1-\phi}{\xi} g^* - \frac{n}{\xi}$$

which is obviously identical to the first line given in (2.28). Thus, also the social bgp growth rate of data sets is identical to those derived for the decentralized economy.

[9]Details are of course available upon request from the authors.

2.A.5 Proof of Proposition 4: BGP labor allocations for the socially planned economy

We first assume that $l_n(t)$ and $l_y(t)$ are constant on the bgp (guess and verify). Solving the necessary first order condition of $l_n(t) = l_n$ for the fraction of shadow prices $\frac{\lambda_1(t)}{\lambda_2(t)}$, inserting this expression in the the necessary first order conditions for $N(t)$ yields on the bgp:

$$g^* l_1^{-1}[\phi l_n + (1 - l_n)(1 - \xi] = \frac{\dot{\lambda}_2(t)}{\lambda_2(t)} + \rho - n. \tag{68}$$

Next, using the necessary first order condition for $G(t)$ or $d_c(t)$, turning them in growth rates, and make use of the resulting expression in (68) in order to eliminate the growth rate of the shadow price. Finally, using first line given in (2.28), on the bgp we have

$$g^* \left[(1 - \xi)(1 - l_n) + l_n \chi(1 - \kappa)\right] = l_n \left[(\xi - \kappa\chi)\left(\frac{g^*(1 - \phi)}{\xi} - \frac{n}{\xi}\right) + \rho - n\chi(1 - \kappa)\right]. \tag{69}$$

Finally, inserting (2.27) in this expression, after solving for l_n yields expression (2.41) in the text. Hence, we have verified that l_n and l_y are indeed constant on the bgp.

3 Digital Governance Solutions

Essay title: Authoritarian surveillance and public support for digital governance solutions

Authors: David Karpa, Michael Rochlitz

Abstract

This study investigates factors influencing support for digital governance solutions and compares this support between autocracies and democracies. We conduct survey experiments in Russia, Germany, Turkey, the United States, and Estonia, and find that awareness of potential misuse of digital governance tools by the government reduces support. Importantly, while this effect has previously been documented for China, we find it irrespective of regime type for an autocracy, a hybrid regime and three democracies. Individuals relying on government-controlled information sources are more likely to endorse digital governance tools. Our study challenges prior findings by indicating that gaps in public service quality do not boost support. Instead, satisfaction with government services correlates with trust in the government's capacity to implement digital governance solutions.

Keywords: Digitization, big data, surveillance, survey experiment, digital authoritarianism, digital governance solutions, algorithmic governance

Publication: Under Review at Comparative Political Studies (since 09. October 2023). Preprint avaible: Karpa, David F. and Rochlitz, Michael, Authoritarian Surveillance and Public Support for Digital Governance Solutions (May 21, 2023). Available at SSRN: https://ssrn.com/abstract=4454627 or http://dx.doi.org/10.2139/ssrn.4454627

3.1 Introduction

Contemporary autocrats face a substantial dilemma when it comes to administering a modern state. On the one hand, it lies in the nature of authoritarian regimes to prevent political change and ensure political control (Svolik 2012). In order to achieve these objectives, dictators often weaken or dismantle existing institutions such as an independent press or judiciary, which could act as a counterweight to the regime (Brewer-Carias 2010; Aleksashenko 2018; Bugaric 2019; Tahiroglu 2020). On the other hand, authoritarian regimes must manage increasingly complex societies and economies in order to compete in a global system characterized by rapid economic and technological change. Managing a modern economy requires efficient institutions that can independently decide on a wide range of issues to prevent antisocial behavior and enforce legal rules.

How to solve this dilemma? One answer could be so-called *digital governance solutions* (DGS), i.e. systems that use digital technologies, artificial intelligence and big data to enforce existing legal rules and provide various services to the population, but also hold the potential for authoritarian surveillance and control. As Jee (2022) argues, such systems can be superimposed on existing institutions in a process she calls *institutional proliferation*. Autocrats thus do not need to reform or strengthen existing institutions but can add an additional layer of new institutions to the existing system. The best-known example of such a system to date are the various versions of a *social credit system* (SCS) that have been introduced in China in recent years (see e.g. Backer (2019) or Kostka (2019)), although other countries are introducing similar systems as well (Engvall and Flak 2022; Manoharan et al. 2022; Peters et al. 2022). During the next couple of years, the use of novel digital technologies for both public administration and authoritarian control could thus become a game changer in the way we think about modern authoritarian states.

Even strongly entrenched autocracies need the support of the population to make digital governance solutions work, however. While the autocrat can enforce adoption, at least to a certain extent, enforcing adoption can be costly. It might therefore be cost-efficient for the autocrat to instead rely on technologies that are useful and enjoy high public approval, for

example because they solve some acutely perceived deficiencies in public administration.[1] If citizens become aware that a digital governance solution is mainly used for repressive purposes, this could lead to citizens boycotting or subverting the system, rendering it less efficient. In addition, if repression becomes too visible, it could trigger public backlash, as happened for example when new lockdowns caused large-scale protests in China in November 2022, forcing the Chinese government to back down on its Covid policy.

Understanding the determinants of public support for digital governance solutions is thus crucial to evaluate their potential future role in autocratic states. For now, this question has only been studied for the Chinese context, where Kostka (2019) identify high rates of approval for SCSs, as citizens see them as a tool to improve the quality of life and fill institutional gaps, rather than as a mechanism of authoritarian control.[2] For example, SCSs are seen as convenient tools to establish trust between economic agents in China's banking sector and economy (Krause and Fischer 2020; Ding and Zhong 2021; Fang et al. 2024). Once Chinese citizens learn about the repressive potential of the new technology, however, their approval is significantly reduced (Xu et al. 2022).

To see if these results also hold in other authoritarian states, we investigate the question for an autocracy where digital governance solutions might play an equally important role in the future, the Russian Federation. We use a between-subjects computer-assisted telephone interviewing (CATI) experiment conducted in October 2022 in Russia (N = 2,462), to investigate how information about the repressive potential of a digital governance solution affects approval for the system. As a robustness-check and to better understand the effect of political regime type on public support, we complement our study with evidence from four additional online survey experiments (N = 1,000 each) in four countries that are in the process of introducing digital governance solutions and that feature different levels of political competitiveness, namely Estonia, Germany, the United States and Turkey.

We find that in the control condition, 70.1% of respondents in Russia approve of the introduction of a digital governance solution that increases bureaucratic efficiency and con-

[1]Indeed, as shown by Li and Kostka (2022), convincing citizens to actively engage with a DGS is not a trivial task, but might require some effort from the side of the authoritarian government.

[2]It seems that Kostka (2019) is for now one of the only papers that explicitly uses the concept of institutional gaps. We find the concept very useful, and build on it in our analysis.

tains some punitive legal capabilities, a number slightly lower than the 80% identified by Kostka (2019) for China. Once Russian citizens are specifically reminded that the system can be used to identify and prosecute political dissent, support drops by almost 25% to 45.9%. Looking at the mechanisms behind our findings, we show that citizens who generally approve of the government and consume information mainly from state-controlled media are more likely to be in favor of introducing a digital governance solution, linking our results to the literature on media effects in autocracies (Adena et al. 2015; Peisakhin and Rozenas 2018; Enikolopov et al. 2022). Other than in China (Kostka 2019), perceived institutional gaps do not seem to influence approval.

Our results for Estonia, Germany, the United States, and Turkey are mostly in line with what we find for Russia, allowing us to reject the hypothesis that political regime type has a significant effect on approval. Approval rates in the control condition are highest in Estonia (74.8%), followed by Turkey (66.7%), Germany (65.9%), and the US (45.9%) (see table 3.B). Once citizens are reminded that the system can be used to prosecute political dissent, support drops significantly in all of the countries (with the exception of the US, where the drop is also visible, but not statistically significant). In all four countries, satisfaction with public services significantly *increases* approval of a DGS, refuting the hypothesis that gaps in the quality of public services create a demand for digital governance solutions.

Our results have a number of important implications. First, as in the Chinese context, knowledge about the repressive potential of a digital governance solution can significantly reduce public support for the system. This is important, as even in autocracies public support remains essential for the proper functioning of a DGS.[3] Second, regime legitimacy matters. If citizens trust their government and receive information about the world mainly through state-approved sources, they are significantly more likely to support the introduction of a digital governance solution by the state. Third, contrary to our expectations, frustration with

[3]As politicians in democracies, authoritarian leaders *do* care about regime legitimacy and public approval (Gerschewski 2018), and are concerned about the possibility of backlash and public protests (Buckley et al. 2022), which might deter them from introducing a DGS. In addition, digital platforms such as a DGS need a sufficiently large amount of public support and participation to function properly, as shown for example by the literature on e-participation (Kneuer and Harnisch 2016; Lee and Kim 2018; Toots 2019). Finally, because for informational autocrats it is crucial to convince the public of their competence (Guriev and Treisman 2022), high citizen approval for a DGS might even be one of the objectives of the strategic autocratic planner.

the quality of public services does not increase support for a DGS. On the contrary, citizens who are satisfied with the state and the quality of public services are also more supportive of introducing a DGS. Finally, we do not find any conclusive evidence that regime type matters for the approval of digital governance solutions. Approval rates were highest in Estonia (a democracy), followed by Russia, Turkey, and Germany (an autocracy, a hybrid regime, and a democracy), with the US (another democracy) being somewhat of an outlier with much lower approval than in the other countries.

Our paper is organized as follows. Section 3.2 shows how our study fits into ongoing debates in the literature, and introduces our theory and hypotheses. Section 4.3 presents our methodology and data, and section 3.4 our results. Section 3.5 discusses the implications of our findings, and section 3.6 concludes.

3.2 Literature and Hypotheses

When are citizens in favor of sharing their data with the state? And what factors influence the willingness of citizens to support governance platforms with a data-sharing component? The literature on these questions is extensive, and our paper contributes to a number of ongoing debates.

One branch of the literature studies the trade-off between civil liberties and security. After the terrorist attacks of September 11, 2001, citizens in the United States were significantly more willing to give up certain civil liberties for greater personal security (Davis and Silver 2004; Lewis 2005). Respondents with lower trust in government institutions were less likely to give up their liberties, ceteris paribus (Davis and Silver 2004). Ziller and Helbling (2021) replicate the study for the European context, and expand the space of threats to also include pandemics. They find that public support for surveillance technologies that include data collection and might restrict civil liberties is generally high, but increases when surveillance is targeted at specific threats rather than being indiscriminate, and when a threat is salient. Concerns about privacy reduce support for sharing data with the state.

The Covid pandemic has led to a vast number of studies investigating this trade-off in the context of a global pandemic. Alsan et al. (2023) conduct a global survey with over 550,000 responses in 2020, and find that major crises, such as terrorist attacks and disease outbreaks,

can change preferences on the trade-off. However, the results are not uniform and reflect the heterogeneity of different countries and demographic groups. For example, disadvantaged citizens may be less willing to sacrifice their rights because they have fought harder to obtain them in the first place, or because they have comparatively fewer rights and privileges to lose than more advantaged social groups, a result also found in other studies (Davis and Silver 2004; Lewis 2005; Dragu 2011; Dietrich and Crabtree 2019). Others found cultural or economic vulnerabilities not to consistently predict concerns about internet surveillance in Arab countries (Martin et al. 2019).

Another development linked to the Covid pandemic is the emergence of contact tracing apps, and the trade-off between civil liberties and public health that they entail (Kitchin 2020; Huang et al. 2022). A range of empirical studies has investigated sources of public support for contact tracing apps, and found that lack of information and concerns over privacy lower support (Williams et al. 2021), while crisis perceptions only seem to play a limited role in explaining uptake (Habich-Sobiegalla and Kostka 2022). Habich-Sobiegalla and Kostka (2022) find that citizens in Germany, the US and China are willing to accept contact tracing apps despite concerns about privacy and government surveillance, as long as they perceive them to be efficient. They argue that this might explain higher approval rates in China, where usage of the app was mandatory, thus increasing its efficiency.

The debate about contact tracing apps fits into a larger literature on attitudes toward government surveillance and digital surveillance technologies. When are people concerned about surveillance by the state? If people trust the government (Trüdinger and Steckermeier 2017; Hillebrand 2021; Liu 2022; Kostka et al. 2023) or the police (Gurinskaya 2020), they seem to be less worried about state surveillance. Surveillance by the state also appears to be better tolerated than digital surveillance by corporations (Steinfeld 2017), with older people being more tolerant of state surveillance, and younger people more tolerant of corporate surveillance (Kalmus et al. 2022). Political views also seem to play a role, with for example US citizens holding libertarian political views being more critical of facial recognition software than those holding right-wing authoritarian views (Peng 2022). Once Peng (2022) informed respondents about potential demographic biases in the technology, however, support dropped among all respondents, irrespective of their political orientation.

In some societies, in particular China, citizens appear to be more tolerant of state surveillance than in others (Kostka 2019; Su et al. 2022; Kostka and Habich-Sobiegalla 2022). The high approval for state surveillance in these societies seems to be due to citizens perceiving digital governance platforms and digital surveillance not as a threat to their privacy and civil liberties, but simply as tools to enhance their convenience and security (Davis and Silver 2004; Kostka 2019; Kostka et al. 2021). Once infringement of privacy is explicitly mentioned, support for digital surveillance drops also in China (Kostka et al. 2023). Positive coverage on state media can also significantly increase public support, especially if citizens use state media as their main source of information (Xu et al. 2022). As the literature on informational autocracies has shown, information control and framing can thus become an important tool to foster public support for specific government policies (Guriev and Treisman 2020, 2022; Rochlitz et al. 2023). Although a specific theory trying to explain the difference in public acceptance of surveillance across different regime types and societies is still missing in the literature, Thompson et al. (2020) show that cultural factors such as societal collectivism can make it more likely that societies accept surveillance by the state. One factor that seems to increase support for digital surveillance across societies is fear, for example of terrorist attacks (Matthes et al. 2019; Hillebrand 2021; Yu and Wong 2023; Kaskeleviciute and Matthes 2022). Fear seems to remain an important predictor of approval even among people who do not trust their own government (Vasilopoulos et al. 2022).

The debate about the trade-off between civil liberties on the one hand and security and convenient governance solutions on the other exists both in democracies and autocracies (see for example Zuboff (2019) for a discussion of corporate surveillance in democracies). It is however particularly relevant in authoritarian and hybrid regimes, because of the new possibilities inherent in digital technologies for authoritarian control and repression (Feldstein 2019, 2021; Kendall-Taylor et al. 2020; Strittmatter 2020; Tirole 2021). Here it might well be possible that digital surveillance technologies in combination with big data and artificial intelligence, once in place, could become a powerful tool permitting contemporary dictators to firmly entrench their hold on power. In hybrid regimes and countries that swing between more and less authoritarian forms of government, such as Turkey or Hungary, but also the United States, a well-functioning surveillance infrastructure could play the role of

a technological "ratchet effect" – permitting a country to switch from softer to harder authoritarianism but preventing it from eventually switching back and democratizing again. As argued by Kostka (2019), citizens in autocracies or weak democracies could also be more tempted by the promises of digital governance solutions, as they might promise an easy fix to the deficiencies of corrupt state administrations.

It is therefore crucial to understand what factors play a role during the establishment of such technologies in authoritarian states and hybrid regimes, with one of the most important elements being public support for digital surveillance and governance platforms. However, for the time being, the literature on this question focuses almost exclusively on China (Kostka 2019; Kostka and Antoine 2020; Strittmatter 2020; Kostka and Habich-Sobiegalla 2022; Liu 2022; Yu and Wong 2023; Kostka et al. 2023), even though the phenomenon is relevant in many hybrid regimes and authoritarian states.[4] In our paper, we test these questions for a range of additional countries. Building on the above-cited literature, we derive four distinct hypotheses:

H1: Informing citizens of its repressive potential will reduce support for a DGS.

H2: Citizens whose main source of information are state-controlled media are more likely to approve of a DGS.

H3: In societies where public services suffer from perceived dysfunctionalities, public approval of a DGS will be higher.

H4: Citizens in democracies are less approving of the introduction of a DGS.

3.3 Data

To test our hypotheses, we conducted a computer-assisted telephone interviewing (CATI) survey experiment with 2,462 respondents in October 2022 in Russia. The survey was pre-registered,[5] and carried out by FOM, a leading Russian sociological and public opinion research organization.[6] Participants were contacted with the help of census data, and a

[4]One notable exception is Feldstein (2021), who studies digital repression in Thailand, the Philippines, and Ethiopia.

[5]ttps://www.socialscienceregistry.org/trials/9566

[6]https://fom.ru/

weighting-scheme was applied to make the sample nationally representative. We complemented our survey with four nationally representative online surveys conducted in September 2022 by the German company Bilendi in Estonia, Germany, the USA, and Turkey, with 1,000 respondents each.[7]

The question of whether public opinion polls can still be trusted in the increasingly repressive Russian context has been intensely debated during the last year, with various researchers having a more (Morris 2022, 2023) or less critical view (Pleines 2023; Reisinger et al. 2022; Rosenfeld 2022, 2023; Volkov 2023) of doing survey research in Russia. We believe that surveys are still a valuable tool to gauge public opinion in contemporary Russia. Unlike the Chinese authorities, the Russian government is not controlling what questions can be asked (Rosenfeld 2023). There are also still a large number of public opinion polls being conducted in Russia every week.[8] Finally, the fact that we find broadly comparable results across all of the five countries we study suggests that public opinion polls can still be used to understand the Russian context – especially with an experimental methodology that permits testing the reaction of the public to a set of randomized treatments.[9]

In all surveys, we use a two-factor between-subjects experiment. Participants were randomly assigned to either a control group or three different treatment conditions. The treatments were exposure to various forms of information about the possible uses of DGS, either in spoken language (CATI) or in written text (online survey). Slight technical differences between the Russian survey and the other surveys are discussed in Appendix 3.A. Beyond technical differences we added two questions to the online survey that are not in the Russian version of the questionnaire, a question on political orientation (*liberal* vs. *conservative*, 10-point slider) and a question on different media sources of information (*Television/News/Radio/Social Media/Smartphone/Apps/Newspaper/Personal conversations*, multiple-choice). The latter is a substitute for the state vs online media question in the Russian version of the questionnaire. We undertook minimal changes trying to maxi-

[7]https://www.bilendi.de/

[8]https://www.themoscowtimes.com/2022/12/06/what-secret-russian-state-polling-tells-us-about-support-for-the-war-a79596

[9]In conducting the same survey in 5 different countries, our paper thus also makes a valuable contribution to the debate about the validity of survey research in authoritarian political contexts.

mize comparability and only changing questions were necessary to ensure validity with the given context. Appendix 3.A.1 provides an English translation of the full text for the Russian survey, and Appendix 3.A.2 the English-language version of the full text for the online surveys.

Respondents were presented with a scenario where a digital platform uses data collected from citizens to make interaction with the government more efficient. The *institutional gap* (Kostka 2019) addressed in our scenario is thus inefficient or cumbersome government bureaucracy, a scenario that respondents in all five countries – despite some between-country heterogeneity – can relate to. In addition to rendering governance more efficient, the digital platform is also able to connect data gathered from citizens with data gathered from other sources, such as CCTV cameras. This data can then be used to hold citizens who "violate law and order" accountable.

Our treatment conditions differ with respect to the amount of detail that is used to describe what is meant by "violating law and order", with figure A2 in the Appendix presenting our control and treatment scenarios. In the no-exposure control condition, we do not explicitly state what *"prosecuting those who violate law and order"* means, with respect to both *prosecution* and *violation of law and order*. In the treatment conditions, different violations of what might be considered *law and order* are specified. The first treatment uses a petty crime – breaking a bench on a playground – as an example of a breach of law and order. In the second treatment, we use "participating in an unauthorized protest" as a proxy for behavior that is illegal but can also be understood as an expression of political dissent.[10] In the third treatment, we follow Xu (2022) and combine both treatments. Across all treatments, the punishment we mention as a consequence of illegal behavior is being excluded from public transport for a period of time.[11]

[10] We specifically decided to use this treatment, as it is particularly salient in the context of our main experiment. In Russia, "participating in an unauthorized protest" ("uchastvovat v nesankcionirovannom mitinge") has become a synonym of expressing political dissent, as all protest actions against the government have in recent years been forbidden by the authorities, and are therefore "unauthorized". For comparability reasons, we keep the same treatment in the other four countries, where, although the costs of participating in "unauthorized protests" is lower, unauthorized protests are also salient (for example in the case of climate activists in Germany, student protests in the US, or anti-government protests in Turkey).

[11] This comes close to some of the standard sanctions in the Chinese SCS. Similarly, being banned from driving was introduced as a punishment in Russia in early 2023 for not complying with electronic summons

To ensure that participants are not overwhelmed with information, we introduce the description of the experimental scenario in two steps. We first introduce the concept of a "unified state digital archive" that will store data about all citizens and has the objective to "increase the efficiency of the government and simplify interaction with citizens" – our digital governance solution. The next question then introduces the possibility that the DGS can also be used for policing and exposes respondents to the four treatment conditions. Questions are being kept as short and easily worded as possible, and answer options for our outcome variables are 4-point Likert scales (*Would definitely approve/Would rather approve/Would rather not approve/Would definitely disapprove*). If respondents do not want to answer a specific question, they have the option to select *don't know/no answer*. In most other cases, we used binary answer options to further reduce complexity and enhance comprehensibility.

Table 3.B presents summary statistics for all our variables. Our sample was 55.5% female and 50.7 years old (SD=16.27), on average. We asked participants on a 1 - 4 scale whether they were satisfied with their income (M=2.92, SD=0.82), and about their employment status. From the latter we build a dummy taking the value of 1 if employed, and 0 otherwise (M=0.59, SD=0.49). Participants reported having received education on a scale ranging from 1 - 6 (M=4.45, SD=1.56), living in either urban or suburban areas (M=0.42, SD=0.49), and living specifically in Moscow (M=0.07, SD=0.26). We asked participants about their sources of news consumption, where 42.7% (SD=0.50) reported consuming mainly news from traditional media sources (TV, radio and newspapers), and 37.6% (SD=0.49) mainly from online media, such as news sites, blogs or social media. This distinction is important, as in October 2022, all traditional media sources in Russia were tightly controlled by the state. Online media, on the other hand, remain less tightly controlled, even though state control is increasing here as well. Regarding trust in public institutions, when asked a binary question 78.9% (SD=0.41) reported trusting public services, 77.5% (SD=0.41) had interacted with online government services at least once in the past year, and 83.4% (SD=0.37) thought the country was developing in a good direction, which we use as a proxy for government

to the military over the DGS Gosuslugi (https://www.washingtonpost.com/world/2023/04/11/russia-conscription-military-mobilization-war/).

approval.[12] Using a 1 - 4 scale, we asked participants about the frequency of previous experiences with online digital services (M=2.59, SD=0.99).

3.4 Results

3.4.1 Survey Experiment Russia

What factors determine public support for a digital governance solution in an autocracy? We start by looking at the descriptive statistics. Figure 3.1 illustrates the effect of the four treatment conditions on support for a DGS that includes a policing component, and the last three columns of Table 3.B in the Appendix provide summary statistics. Our results show that when the policing component is described in general terms ("find and prosecute those who violate law and order"), 70.1% of respondents support the introduction of a DGS. When respondents are provided with the specific example that "a person who broke a bench on a playground could be banned from public transport" (our *crime* treatment), support slightly drops to 67.3%. However, when participants were treated with the *repression* treatment – reminding them that the digital governance solution can be used to identify and prosecute those who participated in "an unauthorized protest" – support decreases by almost 25% to 45.9%. Here it is important to emphasize again that participating in an unauthorized protest, "uchastvovat v nesankcionirovannom mitinge", has become a synonym of political dissent in contemporary Russia, where all anti-government protests are unauthorized. When both treatments are combined, support drops less substantially, to 59.6%, potentially because adding the crime treatment might somewhat lessen respondents' concerns that the policing component of the DGS is mainly aimed at political repression.

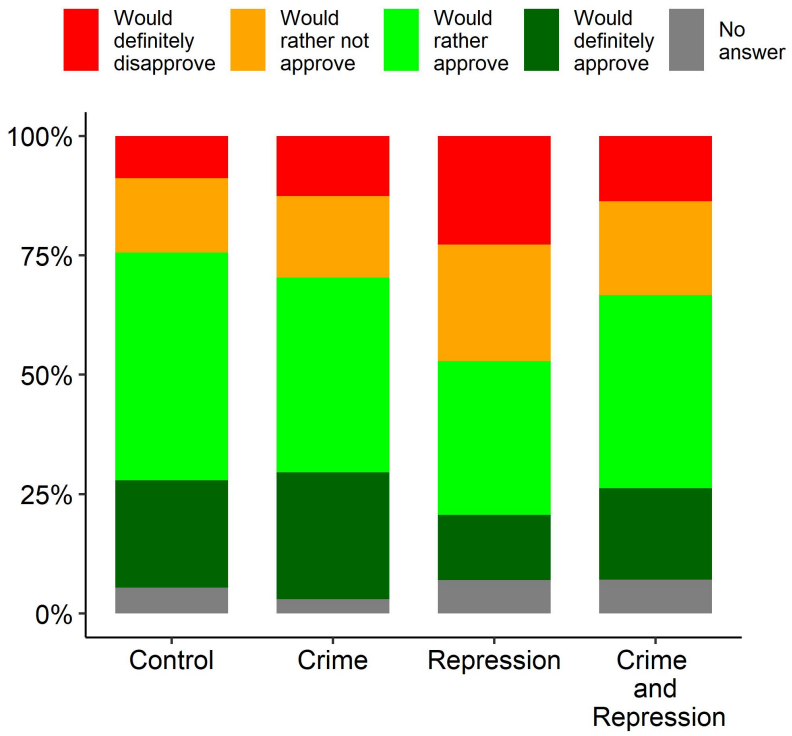
In the next step, we use the following linear equation to conduct a multivariate analysis:

$$y_i = \beta_0 + \beta_1 Crime_i + \beta_2 Repression_i + \beta_3 Crime\ and\ Repression_i + X_i + \epsilon_i, \tag{3.1}$$

where y denotes support for a digital governance solution by respondent i, *Crime, Repression*, and *Crime and Repression* are dummy variables that take a value of 1 if a respondent was

[12] Although 83.4% might seem high, this result is in line with the results of other recent surveys, including studies that use list experiments to test if Russian public opinion polls about government approval reflect the true opinion of the population, see e.g. Frye et al. (2017, 2023).

Figure 3.1: Approval for digital governance solutions in Russia



Notes: See section 3.B in the Appendix for the underlying calculation.

in the respective treatment group and 0 otherwise, X_i is a vector of socio-demographic controls, and ϵ_i is the error term. Figure 3.2 presents our results, and Table A6 in the Appendix presents the corresponding regression analysis.[13]

Our main hypothesis (H1) is that approval of a DGS remains high as long as citizens are not aware of its repressive potential, but that citizens – even in a relatively repressive autocracy – will *reduce* their support for a DGS if treated with information that the new technology can be used to suppress political protest. Conversely, our null hypothesis would be that citizens in repressive autocracies are already aware of constant state surveillance, and have therefore internalized the notion of surveillance by the state to the extent that they no longer update their attitudes when reminded about it once again.

Our empirical results confirm H1 and permit us to reject the null hypothesis. While the

[13] In order to add robustness to our study, we further estimate a multinomial model which we report in Figure A4. In this model we do not find evidence for systematic non-reporting between treatment conditions, which is a blind-spot for our linear model.

crime treatment only weakly reduces approval of the DGS, the repression treatment leads to a significant reduction in public support. Compared to the control group, the repression treatment reduces support by 0.24 SDs[14], which, given a mean of 2.7, is a reduction by 8.8%, with the effect being statistically significant at the 1 percent level. As before, the combined treatment leads to a somewhat smaller reduction in approval, by 4.3%, potentially because adding the crime treatment might reduce concerns that the DGS is aimed exclusively at political repression. The comparatively strong and negative effect of the repression treatment on approval shows that even in the repressive environment of October 2022, Russian citizens remain genuinely concerned about the surveillance capabilities that a big data governance solution might imply, once they are informed about these capabilities.

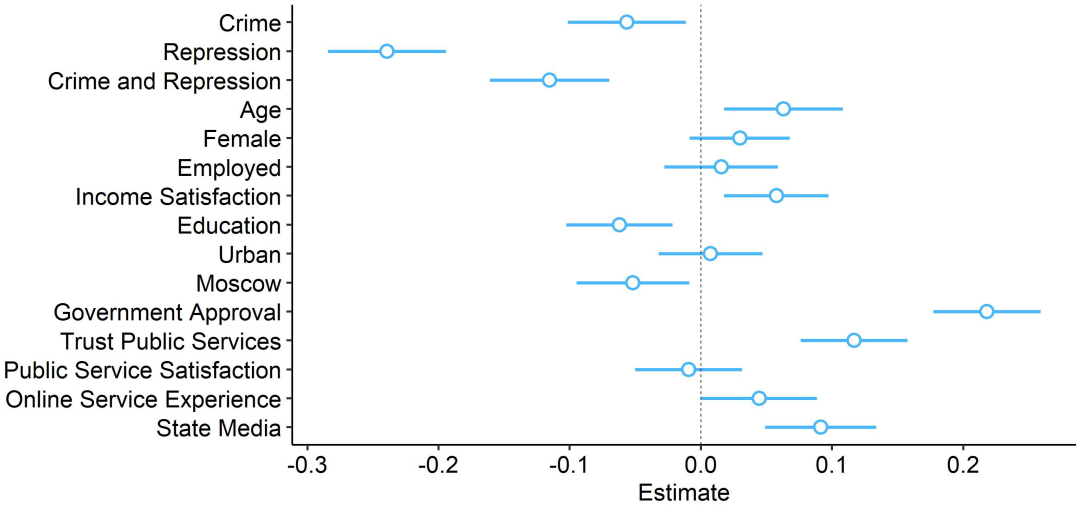
In addition to the effect of information on support for a DGS, our empirical analysis reveals a number of further interesting findings, that can help us understand the mechanisms behind our results. First, government approval appears to be a strong predictor of support for a DGS.[15] Citizens who approve of the general direction the country is taking are also significantly more likely to support the introduction of a DGS.

We believe that one important mechanism through which government approval affects preferences for a DGS are channels of information – our hypothesis H2. Citizens who inform themselves mainly through government-controlled sources of information such as state-controlled TV are also significantly more likely to support the introduction of a DGS. Figure 3.3 splits our sample into respondents who mainly receive information through state-controlled media, and respondents whose main source of information are online media. We see that consumers of online-media such as news sites, blogs or social media – which offer a more pluralistic choice of views – are significantly more sceptical about the introduction of a DGS than consumers of state-controlled media, irrespective of the treatment. We estimate interaction effects between the treatment conditions and state media, finding no heteroge-

[14]This effect size can be considered to be small to medium. However, because the control condition also implicitly entails the potential for misuse, the treatment effects are potentially underestimated, which should be taken into consideration.

[15]To measure government approval, we use a standard question that has been used for many years in social science surveys to measure general support for the policies of the government in Russia, see Question 10 in Appendix 3.A.1.

Figure 3.2: Treatment effects Russia



Notes: Standardized OLS estimates with 95% confidence intervals. See Table A6 in the Appendix for the underlying regression results (specification (3) displayed here).

neous effects of state media, which is in line with the previous assessment of Figure 3.3 (see Table A7 in the Appendix).

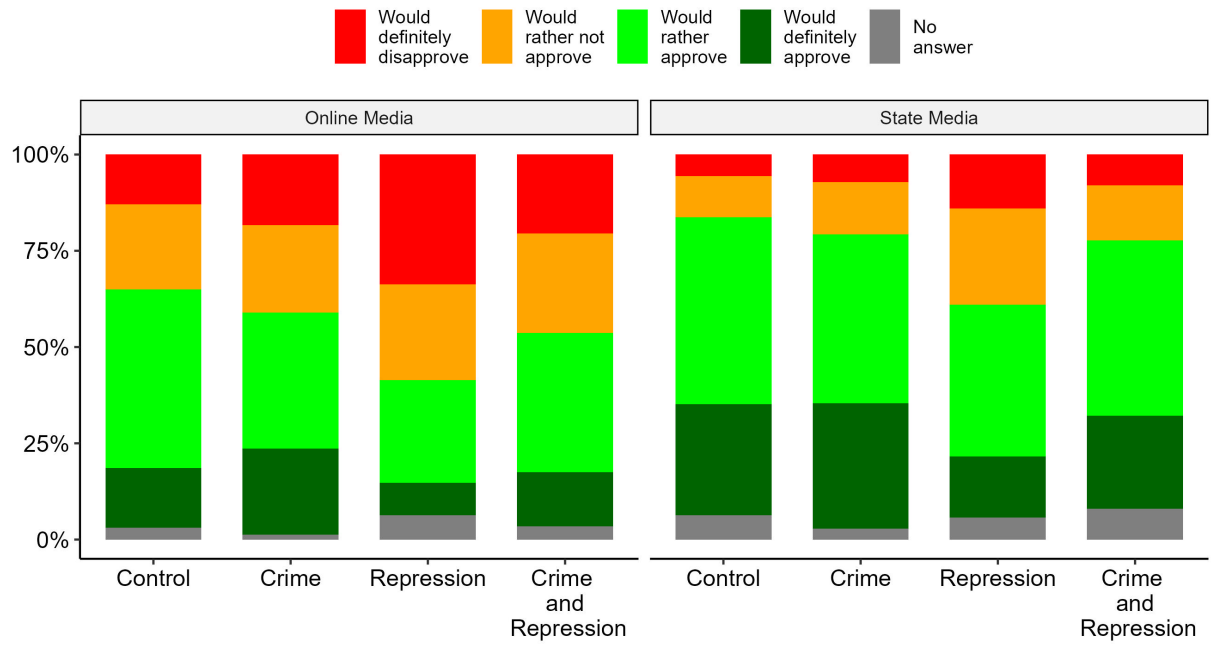
Third, general trust in the authorities providing public services also positively predicts support for a DGS.[16] However, other than with the four online surveys discussed in section 3.4.2, a recent positive or negative experience with government service provision has no effect on support. We thus cannot say that existing institutional gaps (i.e. deficiencies with public service provision) might have an influence on support for a digital platform that could address some of these deficiencies (our hypothesis H3). [17]

Finally, citizens who are older and have higher incomes and previous exposure to online services are also slightly more supportive of a digital governance solution, while employment status, gender, and living in a city or a rural location have no clear effect. Conversely, citizens with higher levels of education are less in favor of a DGS, similar to those living in

[16]When using a different coding for this question, as specified in Table A6 under "Coding", the coefficient is still positive but smaller and no longer significant. We thus advise the reader to handle this result with caution.

[17]In addition, we estimate interaction effects for these variables in the specifications (2) to (5) which we report in Table A7. We do not find statistically significant interaction effects between our treatment conditions and trust in and satisfaction with public services.

Figure 3.3: Effects of different media sources on approval



Notes: The sample was split between different primary sources of information for all treatments

Moscow.

3.4.2 Online Survey Experiments

To complement the experiment we carried out in Russia, we conducted four additional survey experiments in Estonia, Germany, the US and Turkey. The countries were selected to provide us with a range of different regime types, to test if approval of a DGS varies between countries with different political institutions.[18] Estonia and Germany are both consolidated democracies, with the Internet being slightly less free in Germany.[19] The United States, clearly a focal point for the study of contemporary institutional development, scores slightly lower on most institutional indicators. Turkey's political institutions have deteriorated significantly over the last decade, which is reflected in its low democracy scores across the indicators. Similarly, Russia has the lowest institutional indicators in our sample. Both

[18]See Tables A4 and A5 in the Appendix for the range of different institutional indicators used in our approach.

[19]Estonia scores 93 out of 100 on the Freedom on the Net index, while Germany scores 77 out of 100.

countries have become increasingly authoritarian over the last two decades. It is important to note that the difference between Russia and Turkey is small, while the gap to the other countries is quite substantial.

Figure 3.4: Approval for digital governance solutions

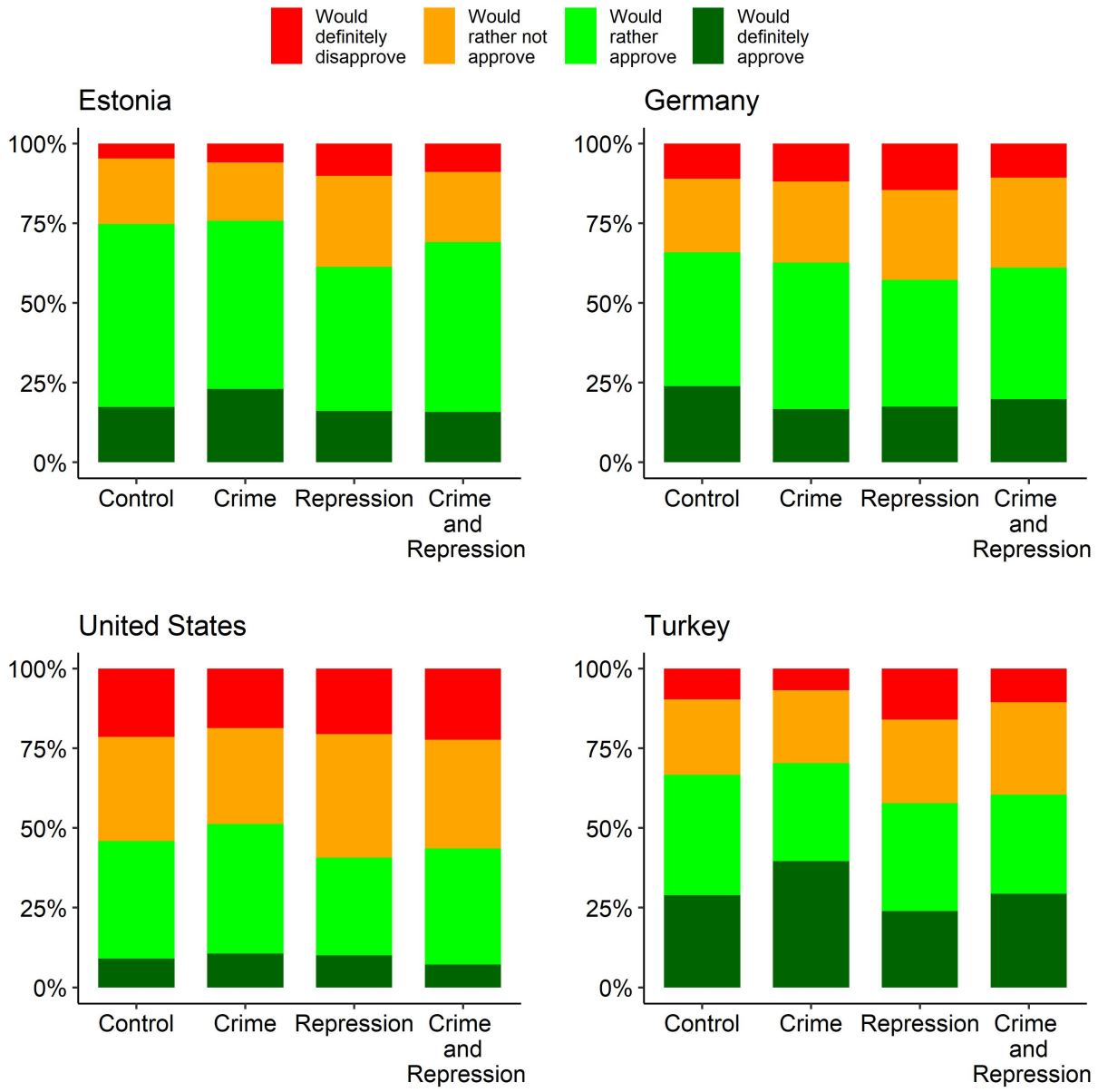


Figure 3.4 presents the descriptive statistics of our treatment effects, while Table 3.B in the Appendix provides the underlying summary statistics. We find that approval rates for a DGS vary across countries and treatments. In general, average approval rates in the control

condition are highest in Estonia (74.8%), followed by Turkey (66.7%), Germany (65.9%), and the US (45.9%)

First, the fact that the approval rate in the control condition for Russia (70.1%) is somewhere in the middle of what we find for the other countries suggests that the Russian data are meaningful and that survey participants in Russia did not self-censor (Robinson and Tannenberg 2019) or falsify their preferences (Eck et al. 2021), at least not more than in the other countries.

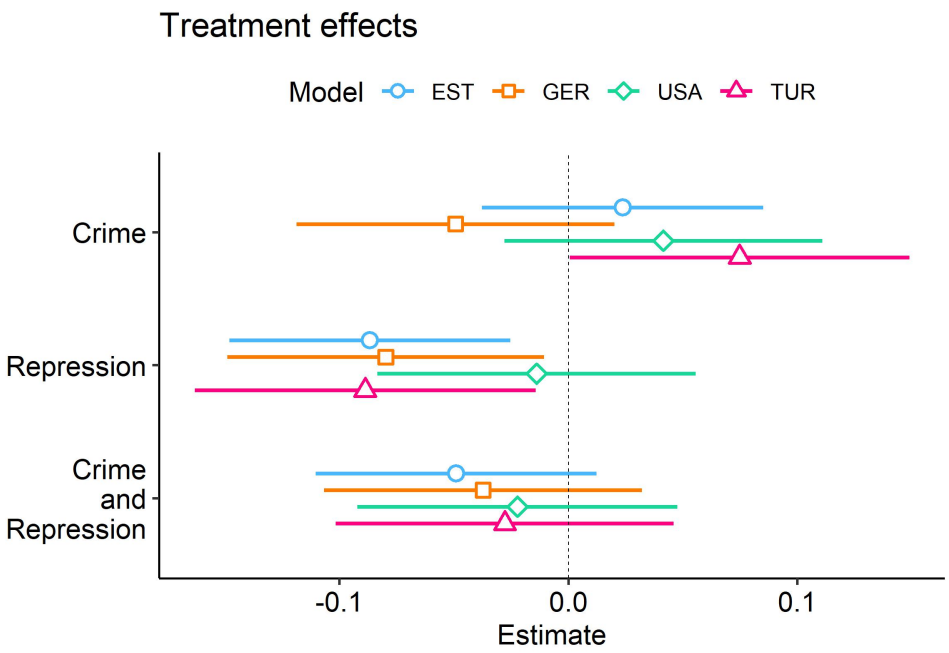
Second, and more interestingly, we do *not* find a linear relationship between political freedom and approval of a new institution that might potentially limit it, allowing us to reject our hypothesis H4. Our data suggest at least two possible explanations. Either there is non-linearity in the correlation, for example a U-shaped dependence between institutional freedom and support for digital government solutions. Alternatively, individual country characteristics, such as culture, economic development or history, might play a more critical role than indices designed only to measure political institutions. For example, it might be the case that in the United States, a country whose constitution, founding myth and economic system put particular emphasis on the importance of freedom, institutions with the potential to limit freedom are met with more scepticism than in other countries.[20] The US might therefore be an outlier with respect to absolute approval rates for a DGS.

While we find differences across countries with respect to overall approval rates, when comparing the variation between treatments, results are relatively similar across countries. For all four countries, the *repression* treatment caused a visible dent in approval rates, similar to what we found for Russia in section 3.4.1. The effects for the *crime* and the combined treatment are less pronounced.

For a more precise analysis, we again use equation 3.1 to conduct a multivariate analysis. Figure 3.5 presents the results of our treatments, Figure A3 in the Appendix the full set of results, and Tables A8, A9, A10 and A11 in the Appendix the underlying regression analyses. Additionally, we estimated a logistic regression model that serves as a robustness check (see Figure A5).

[20]See for example Kennedy (1999), McPherson (2003), Rana (2014) or Schmidli (2022) for a discussion of

Figure 3.5: Treatment effects in Estonia, Germany, the United States, and Turkey



Notes: Standardized OLS estimates with 95% confidence intervals. See appendix 3.C for the underlying regression results.

In Estonia, Germany, and Turkey, our main finding from the first experiment could be replicated, i.e., the coefficient for the *repression* treatment has a negative effect and is significant at least at the 95%-level, providing additional evidence in support of hypothesis H1. Results for the US are also negative, but not statistically significant, with a smaller coefficient than for the other three countries. To better understand our results for the US, we looked at the effect of our treatment on different age groups, and found evidence for significant age-related polarization (see Figure A1 in the Appendix, as well as Table A13). It appears that citizens in the US who are 60 years and older are actually *more* in favor of a DGS if the technology features a repression component, explaining why our *repression* treatment remains insignificant for the country as a whole. This effect, however, can only be found for the US, and not for any other country in our study. In Turkey, there are even statistically significant additional *negative* treatment effects for citizens above 60 years.[21]

the importance of the concept of freedom in recent US history.

[21] In addition to age, we also test for heterogeneous treatment effects for political orientation and government approval. We find that the repression and combined treatments reduce approval for a DGS among US citizens with liberal political views (Table A14). We find no heterogeneous treatment effects with respect

Finally, other than in the Russian experiment where we find a negative effect, the *crime* and the combined treatment had no statistically significant effects in Turkey, Estonia, Germany and the US. This might be explained by the very low trust rates Russians have for the police (see e.g. Semukhina and Reynolds (2014)), while trust in the police is substantially higher in the other four countries in our sample.[22]

Our results suggest that the effects we find for Russia seem to be robust across countries with different regime types and different levels of political competitiveness (see Figure 3.6). While in general approval of a DGS is high, once respondents are primed that the technology can be used for purposes of political control, the drop in approval rates is substantial. Interestingly, this drop only occurs in the case of the political repression treatment, and not when respondents are treated with the possibility that the technology could be used for policing petty crime. If anything, the results from countries other than Russia suggest that the policing functionality of a DGS *increases* approval, although not significantly.

Figure A3 in the Appendix presents some additional interesting results. First, trust in the quality of public services, as well as a positive experience with government services during the previous year are positive predictors of approval for a DGS. These results refute our "institutional gaps" hypothesis (H3). People do not seem to be more likely to approve (or wish for) a DGS because they are experiencing difficulties with existing public services. Rather, they approve of introducing a DGS if they already have a positive view of public services – probably because they expect the authorities to implement it in a responsible and efficient way.

Second, respondents who are satisfied with their personal income situation and who approve of "the direction into which the country is developing" (our "government approval" indicator) are also more likely to approve of a DGS. These results suggest that being satisfied with your personal situation and the overall affairs of the country also makes people more open to the introduction of new governance technology. Interestingly, though, education has

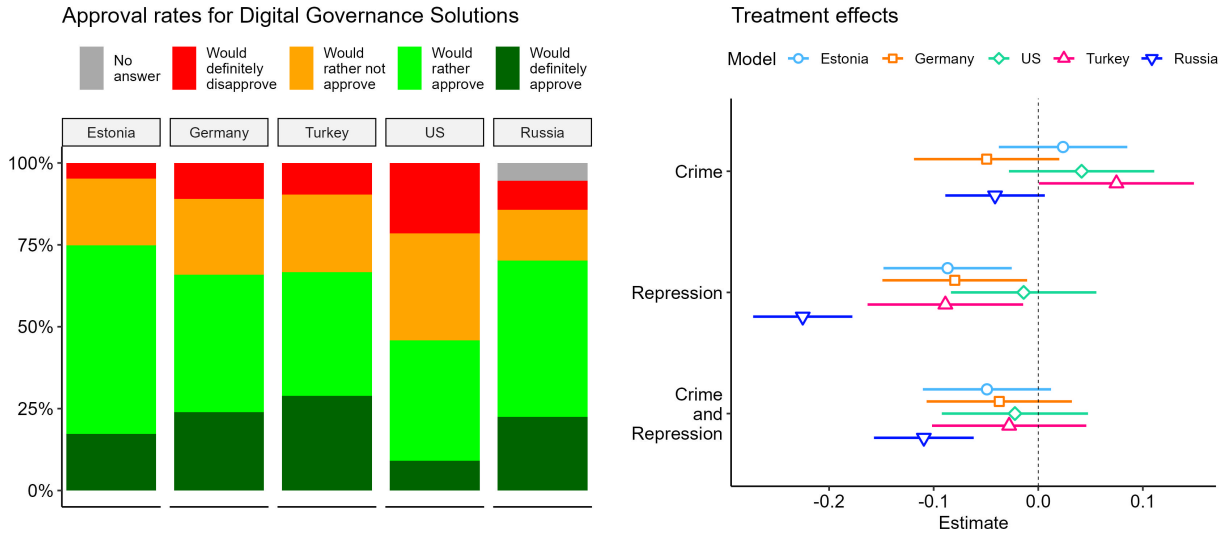
to government approval (Table A15).

[22] Recent survey data show a big variation in trust in the police across countries, with 86% of respondents trusting the police in Estonia, 49% in Germany, 48% in the US, 39% in Turkey, but only 20% in Russia in 2022. See Goldie (2023) for data on Germany, the US, Turkey and Russia, and Kaitseministeerium (2023) for Estonia.

the opposite effect. More educated respondents, on average, are more sceptical towards the introduction of a DGS, with the effect being significant for Estonia, Germany, and the US. These findings are robust for specifications including the full sample, see Table A12.

Finally, while we also find a positive effect of TV news consumption on DGS approval for Estonia, Germany, and the US (as in the case of Russia), here the results of the experiments cannot be directly compared with the results we obtain for Russia. While in Russia receiving news about the state of the world mainly from television has become a synonym for being a recipient of state propaganda, the content of TV news in the other countries in our sample is much more heterogeneous, and does not allow a similar conclusion.

Figure 3.6: Approval for digital governance solutions and treatment effects summary



Notes: Left panel gives approval rates for Digital Governance Solutions as measured in the control condition. In Russia, due to the nature of CATI, the data includes "don't know/no answer", while no such option was available in the online surveys in the other four countries. Right panel shows treatment effects as estimated in the models in the Section 3.C.

3.5 Discussion of the Experimental Findings

Perhaps the most important finding of our study is that *information* – about the potential costs of a DGS – *matters*. In other words, public support for a DGS seems to depend on citizens not being aware of its repressive potential, or, more generally, of the costs imposed by the new technology. Once they are being informed about the political risks of a new

institution, approval drops significantly, irrespective of the context and country they live in.

When having to decide about the advantages and disadvantages of adopting a new institution, a rational citizen will infer information from priors that stem from her or his *informational* and *institutional* background. Our results show that in this situation, sources of information can play an important role. An extensive literature has shown that the media can be instrumental in influencing human behavior, both in democracies (DellaVigna and Kaplan 2007; Gerber et al. 2009; Barone et al. 2015; Durante et al. 2019) and in autocracies (Enikolopov et al. 2011, 2022; Adena et al. 2015; Peisakhin and Rozenas 2018). Autocracies such as Russia are particular, however, in that the state has agenda-setting power over the media, and can suppress alternative sources of information. How the government frames certain issues can then have an important effect on citizens' perceptions and behavior (Kazun 2016; Pan et al. 2022). Xu et al. (2022) illustrate this point by showing how Chinese citizens who receive information about the Chinese social credit system through state media subsequently show higher levels of support for the system. In our paper, we document a similar effect for Russia, with Russian citizens who receive information mainly through state-controlled media being significantly more likely to support the introduction of a DGS.[23]

Another informational input influencing citizens' cost-benefit analysis is the institutional environment. When citizens perceive the government and existing government services as trustworthy, they might infer that new institutions can also be trusted. This is indeed what we find. Russian citizens who approve of the direction the country is taking – our proxy for government approval – are almost three standard deviations more supportive of a DGS. We find a similar, albeit smaller effect for trust in public services. Both effects can also be found for Estonia, Germany, the United States and Turkey, although here the effects are not always statistically significant.

In this, our paper relates to an extensive literature investigating the role of institutions in

[23]Indeed, the subject of the digitization of the Russian state has been extensively and positively covered on Russian state media, in particular since Mikhail Mishustin took over as prime minister in January 2020, who made the digitization of Russia's state administration one of his priorities (see for example https://tass.com/society/1571181 and https://foreignpolicy.com/2020/01/20/russia-incoming-prime-minister-techno-authoritarianism/).

determining acceptance rates for new technologies and tolerance for government surveillance. Similar to our study, a number of experimental papers have found that trust in the government or its institutions plays a key role in the acceptance of facial recognition technology (Kostka et al. 2021, 2023) and correlates positively with support for government surveillance (Trüdinger and Steckermeier 2017; Su et al. 2022) and sacrificing civil liberties for security (Davis and Silver 2004; Alsan et al. 2023). It also increases support for social credit systems (Kostka 2019) and leads to a higher willingness for data-sharing in the context of Covid-19 contact tracing apps (Kostka and Habich-Sobiegalla 2022; Huang et al. 2022). Looking at the issue from the opposite direction, Kostka and Antoine (2020) show that digital governance solutions such as China's social credit systems work better when confidence in the government is high. Ziller and Helbling (2021) reverse this causality and examine the effect of surveillance on trust in governments. Some studies argue that strong emotions such as fear can moderate the effect of trust in the government, so that even people who normally would not trust their government approve of restricting civil liberties when exposed to risks such as political unrest (Yu and Wong 2023) or Covid-19 (Vasilopoulos et al. 2022).

In sum, it seems that even in authoritarian contexts such as Russia or China, trust in government institutions is crucial to make digital governance solutions work. This is why contemporary dictatorships often attach a lot of importance to creating public legitimacy for the regime, often through control over channels of information and the media (Guriev and Treisman 2020, 2022). The media can thus become an important driver of trust in institutions, while also helping to build support for specific policies favoured by the state – as we document in our paper.

3.6 Conclusion

In our study, we conducted a survey experiment in five different countries to understand the determinants of public approval for digital governance solutions. We find that information about the potential abuse of a DGS by a government for purposes of political repression can significantly reduce public support for the new technology. Crucially, this effect is consistent across different institutional settings that feature different levels of institutional safeguards

against government abuse of power.[24]

Why does this matter? Big data governance technologies as the one described in our paper have the potential to become a game changer in the way we think about government surveillance and political control, in particular – but not only – in autocracies. China is a pioneer in this respect. During the last couple of years, China has tested a number of social credit systems in different regions of the country (Kostka 2019; Kostka and Antoine 2020; Strittmatter 2020; Li and Kostka 2022; Liu 2022). The Covid pandemic (Knight and Creemers 2021; Chen et al. 2022) and recent breakthroughs in artificial intelligence, big data and facial recognition have given the technology an additional boost. China is now using the data gathered via its surveillance systems as an input subsidy for Chinese firms, to promote its domestic industry and further improve its surveillance capabilities in what has become a positive feedback loop (Beraja et al. 2023c,b). This has made China an undisputed leader in the market for digital governance, big data and facial recognition technologies (Feldstein 2023), which are now being offered as integrated packages, for example via so called "smart city" or "safe city" solutions (Yang and Xu 2018; Große-Bley and Kostka 2021). Beraja et al. (2023a) find that China is now actively exporting these technologies, in particular to other authoritarian countries. If an authoritarian government was recently challenged by domestic political protests, the likelihood that it will import a DGS with a policing component from China is particularly high (Beraja et al. 2023a). Often, surveillance technologies are offered in bundles with infrastructure and other technologies, as part of global strategies such as the "Belt and Road Initiative". As with the democratizing effects of trade with democratic countries (Tabellini and Magistretti 2022), trade with China can thus foster authoritarian consolidation, via the export of digital governance solutions (Beraja et al. 2023b; Feldstein 2023).

One country that is particularly at risk is Russia. Russia has been learning from China with respect to mechanisms of authoritarian governance for some time (Libman and Rochlitz 2019), has recently been digitizing its service sector and economy (Østbø 2021), while also

[24]This finding begs the question of whether citizens perceive existing institutional safeguards as *insufficient*, or whether institutional safeguards do not play a significant role in moderating this effect, opening up interesting new avenues for further research.

becoming increasingly authoritarian. While this paper has been written, Russia started using facial recognition technology from China to identify draft dodgers in the Moscow metro,[25] and its novel DGS "gosuslugi" to deliver conscription orders to Russian men who were supposed to join the ranks of the Russian army in Ukraine.[26] Other countries have also started implementing big data governance technologies from China, for example Kenya, Laos, Mongolia, Uganda, Saudi Arabia and Uzbekistan (Feldstein 2023). As argued by Feldstein (2023), once put in place, technologies combining digital governance and big data with capabilities of surveillance and control might make it much more difficult than today for civic accountability and democratization movements to challenge autocratic governments, potentially introducing a new wave of autocratization (Snyder 2018; Lührmann and Lindberg 2019).

Understanding the factors that determine the often surprisingly high rates of public approval for digital governance solutions is therefore crucial. As our study shows, it seems that once citizens become aware that such technologies could play the role of a "Trojan horse" for introducing methods of authoritarian control, they are much more circumspect about adopting the new technology. This is important, not only in autocracies such as Russia, but also – and probably even more so – in hybrid regimes and democracies that have proven vulnerable to populist leaders with authoritarian tendencies, such as for example Turkey or the United States.

3.7 Bibliography

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., and Zhuravskaya, E. (2015). Radio and the Rise of the Nazis in Prewar Germany. *The Quarterly Journal of Economics*, 130(4):1885–1939.

Aleksashenko, S. (2018). *Putin's Counterrevolution*. Washington, D.C.: Brookings Institution Press.

Alsan, M., Braghieri, L., Eichmeyer, S., Kim, M. J., Stantcheva, S., and Yang, D. Y. (2023). Civil liberties in times of crisis. *American Economic Journal: Applied Economics*, 15(4):389–421.

[25] https://www.hrw.org/news/2022/10/26/russia-uses-facial-recognition-hunt- down-draft-evaders

[26] https://carnegieendowment.org/politika/89553

Backer, L. C. (2019). China's Social Credit System: Data-Driven Governance for a 'New Era'. *Current History*, 118(809):209–214.

Barone, G., D'Acunto, F., and Narciso, G. (2015). Telecracy: Testing for Channels of Persuasion. *American Economic Journal: Economic Policy*, 7(2):30–60.

Beraja, M., Kao, A., Yang, D., and Yuchtman, N. (2023a). Exporting the Surveillance State via Trade in AI. *NBER Working Paper*, 31676.

Beraja, M., Kao, A., Yang, D. Y., and Yuchtman, N. (2023b). AI-tocracy. *The Quarterly Journal of Economics*, 138(3):1349–1402.

Beraja, M., Yang, D. Y., and Yuchtman, N. (2023c). Data-intensive Innovation and the State: Evidence from AI Firms in China. *The Review of Economic Studies*, 90(4):1701–1723.

Brewer-Carias, A. (2010). *Dismantling Democracy in Venezuela: The Chávez Authoritarian Experiment*. Cambridge: Cambridge University Press.

Buckley, N., Reuter, J., Rochlitz, M., and Anton, A. (2022). Staying Out of Trouble: Criminal Cases Against Russian Mayors. *Comparative Political Studies*, 55(9):1539–1568.

Bugaric, B. (2019). Central Europe's descent into autocracy: A constitutional analysis of authoritarian populism. *International Journal of Constitutional Law*, 17(2):597–616.

Chen, W., Huang, G., and Hu, A. (2022). Red, Yellow, Green or Golden: the Post-Pandemic Future of China's Health Code Apps. *Information, Communication & Society*, 25(5):618–633.

Davis, D. and Silver, B. (2004). Civil Liberties vs. Security: Public Opinion in the Context of the Terrorist Attacks on America. *American Journal of Political Science*, 48(1):28–46.

DellaVigna, S. and Kaplan, E. (2007). The Fox News Effect: Media Bias and Voting. *The Quarterly Journal of Economics*, 122(3):1187–1234.

Dietrich, N. and Crabtree, C. (2019). Domestic Demand for Human Rights: Free Speech and the Freedom-Security Trade-Off. *International Studies Quarterly*, 63(2):346–353.

Ding, X. and Zhong, D. (2021). Rethinking China's Social Credit System: A Long Road to Establishing Trust in Chinese Society. *Journal of Contemporary China*, 30(130):630–644.

Dragu, T. (2011). Is there a trade-off between security and liberty? executive bias, privacy protections, and terrorism prevention. *The American Political Science Review*, 105(1):64–78.

Durante, R., Pinotti, P., and Tesei, A. (2019). The Political Legacy of Entertainment TV. *American Economic Review*, 109(7):2497—-2530.

Eck, K., Hatz, S., Crabtree, C., and Tago, A. (2021). Evade and Deceive? Citizen Responses to Surveillance. *The Journal of Politics*, 83(4):1545–1558.

Engvall, T. and Flak, L. (2022). Digital Governance as a Scientific Concept. In Charalabidis, Y., Flak, L., and Pereira, G., editors, *Scientific Foundations of Digital Governance and Transformation: Concepts, Approaches and Challenges*, chapter 2, pages 25–50. SpringerLink.

Enikolopov, R., Petrova, M., and Zhuravskaya, E. (2011). Media and Political Persuasion: Evidence from Russia. *American Economic Review*, 101(7):3253–85.

Enikolopov, R., Rochlitz, M., Schoors, K., and Zakharov, N. (2022). The Effect of Independent Online Media in an Autocracy. *SSRN Working Paper No. 4346225.*

Fang, H., Chen, L., Xiong, J., and Zhu, Y. (2024). Social Trust and Entrepreneurship: Insights from China's Social Credit System. *Finance Research Letters*, 62(Part A):105118.

Feldstein, S. (2019). The Road to Digital Unfreedom: How Artificial Intelligence Is Reshaping Repression. *Journal of Democracy*, 30(1):40–52.

Feldstein, S. (2021). *The Rise of Digital Repression: How Technology Is Reshaping Power, Politics, and Resistance*. Oxford: Oxford University Press.

Feldstein, S. (2023). Surveillance in the Illiberal State. In Sajó, A., Uitz, R., and Holmes, S., editors, *Routledge Handbook of Illiberalism*, chapter 24, pages 351–364. Routledge.

Frye, T., Gehlbach, S., Marquardt, K., and Reuter, O. J. (2017). Is Putin's Popularity Real? *Post-Soviet Affairs*, 33(1):1–15.

Frye, T., Gehlbach, S., Marquardt, K., and Reuter, O. J. (2023). Is Putin's Popularity (Still) Real? A Cautionary Note on Using List Experiments to Measure Popularity in Authoritarian Regimes. *Post-Soviet Affairs*, 39(3):213–222.

Gerber, A. S., Karlan, D., and Bergan, D. (2009). Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions. *American Economic Journal: Applied Economics*, 1(2):35–52.

Gerschewski, J. (2018). Legitimacy in Autocracies: Oxymoron or Essential Feature? *Perspectives on Politics*, 16(3):652–665.

Goldie, J. (2023). Interactive: Trust in the police around the world. *360info*.

Große-Bley, J. and Kostka, G. (2021). Big Data Dreams and Reality in Shenzhen: An Investigation of Smart City Implementation in China. *Big Data & Society*, 8(2):1–14.

Guriev, S. and Treisman, D. (2020). A Theory of Informational Autocracy. *Journal of Public Economics*, 186:104158.

Guriev, S. and Treisman, D. (2022). *Spin Dictators: The Changing Face of Tyranny in the 21st Century*. Cambridge, MA: Princeton University Press.

Gurinskaya, A. (2020). Predicting citizens' Support For Surveillance Cameras. Does Police Legitimacy Matter? *International Journal of Comparative and Applied Criminal Justice*, 44(1-2):63–83.

Habich-Sobiegalla, S. and Kostka, G. (2022). Sharing is Caring: Willingness to Share Personal Data Through Contact Tracing Apps in China, Germany, And The US. *Information, Communication & Society*, pages 1–28.

Hillebrand, K. (2021). The Role of Fear and Trust When Disclosing Personal Data to Promote Public Health in a Pandemic Crisis. In Ahlemann, F., Schütte, R., and Stieglitz, S., editors, *Innovation Through Information Systems: Volumne I: A Collection of Latest Research on Domain Isues*, chapter 18, pages 247–262. Springer.

Huang, G., Hu, A., and Chen, W. (2022). Privacy at Risk? Understanding the Perceived Privacy Protection of Health Code Apps in China. *Big Data & Society*, 9(2).

Jee, H. (2022). *Credit for Compliance: How Institutional Layering Ensures Compliance in China*. Phd-thesis, Stanford University.

Kaitseministeerium (2023). Public Opinion on National Defense. *Kaitseministeerium*.

Kalmus, V., Bolin, G., and Figueiras, R. (2022). Who is Afraid of Dataveillance? Attitudes Toward Online Surveillance in a Cross-cultural and Generational Perspective. *New Media & Society*.

Kaskeleviciute, R. and Matthes, J. (2022). A Vicious Cycle? Threat of Terror, Perceived Media Bias, and Support for Surveillance Policies. *Mass Communication and Society*.

Kazun, A. (2016). Framing Sanctions in the Russian Media: The Rally Effect and Putin's Enduring Popularity. *Demokratizatsiya: The Journal of Post-Soviet Democratization*, 24(3):327–350.

Kendall-Taylor, A., Frantz, E., and Wright, J. (2020). The Digital Dictators: How Technology Strengthens Autocracy. *Foreign Affairs*, 99(103-115).

Kennedy, D. (1999). *Freedom from Fear: The American People in Depression and War, 1929-1945*. Oxford: Oxford University Press.

Kitchin, R. (2020). Civil Liberties or Public Health, or Civil Liberties and Public Health? Using Surveillance Technologies to Tackle the Spread of COVID-19. *Space and Polity*, 24(3):362–381.

Kneuer, M. and Harnisch, S. (2016). Diffusion of e-government and e-participation in Democracies and Autocracies. *Global Policy*, 7(4):548–556.

Knight, A. and Creemers, R. (2021). Going Viral: The Social Credit System and COVID-19. *Working Paper, Leiden University*.

Kostka, G. (2019). China's Social Credit Systems And Public Opinion: Explaining High Levels of Approval. *New Media & Society*, 21(7):1565–1593.

Kostka, G. and Antoine, L. (2020). Fostering Model Citizenship: Behavioral Responses to China's Emerging Social Credit Systems. *Policy & Internet*, 12(3):256–289.

Kostka, G. and Habich-Sobiegalla, S. (2022). In Times of Crisis: Public Perceptions Toward COVID-19 Contact Tracing Apps in China, Germany, and the United States. *New Media & Society*, 0(0).

Kostka, G., Steinacker, L., and Meckel, M. (2021). Between Security and Convenience: Facial Recognition Technology in the eyes of citizens in China, Germany, the United Kingdom, and the United States. *Public Understanding of Science*, 30(6):671–690.

Kostka, G., Steinacker, L., and Meckel, M. (2023). Under Big Brother's Watchful Eye: Cross-country Attitudes Toward Facial Recognition Technology. *Government Information Quarterly*, 40(1):101761.

Krause, T. and Fischer, D. (2020). An Economic Approach to China's Social Credit System. In Everling, O., editor, *Social Credit Rating: Reputation und Vertrauen beurteilen*, chapter 22, pages 437–454. SpringerGabler.

Lee, J. and Kim, S. (2018). Citizens' E-participation on Agenda Setting in Local Governance: Do Individual Social Capital and E-participation Management Matter? *GPublic Management Review*, 20(6):873–895.

Lewis, C. W. (2005). The Clash Between Security and Liberty in the U.S. Response to Terror. *Public Administration Review*, 65(1):18–30.

Li, H. and Kostka, G. (2022). Accepting but not Engaging with it: Digital Participation in Local Government-run Social Credit Systems in China. *Policy & Internet*, 14(4):845–874.

Libman, A. and Rochlitz, M. (2019). *Federalism in China and Russia: Story of Success and Story of Failure?* Cheltenham: Edward Elgar.

Liu, C. (2022). Who Supports Expanding Surveillance? Exploring Public Opinion of Chinese Social Credit Systems. *International Sociology*, 37(3):391–412.

Lührmann, A. and Lindberg, S. (2019). A Third Wave of Autocratization is Here: What is new about it? *Global Policy*, 7(4):548–556.

Manoharan, A., Melitski, J., and Holzer, M. (2022). Digital Governance: An Assessment of Performance and Best Practices. *Public Organization Review*.

Martin, J. D., Naqvi, S. S., and Schoenbach, K. (2019). Attribute substitution and stereotypes about the online arab public sphere: Predictors of concerns about internet surveillance in five arab countries. *New Media & Society*, 21(5):1085–1104.

Matthes, J., Schmuck, D., and von Sikorski, C. (2019). Terror, Terror Everywhere? How Terrorism News Shape Support for Anti-Muslim Policies as a Function of Perceived Threat Severity and Controllability. *Political Psychology*, 40(5):935–951.

McPherson, J. (2003). *Battle Cry of Freedom: The Civil War Era*. Oxford: Oxford University Press.

Morris, J. (2022). Political Ethnography and Russian Studies in a Time of Conflict. *Post-Soviet Affairs*, 39(1-2):92–100.

Morris, J. (2023). Public Opinion Still Does Not Exist; War in Ukraine and Dictatorship in Russia Can Help Us Acknowledge That. *Russian Analytical Digest*, 292:7–8.

Pan, J., Shao, Z., and Xu, Y. (2022). How government-controlled media shifts policy attitudes through framing. *Political Science Research and Methods*, 10(2):317–332.

Peisakhin, L. and Rozenas, A. (2018). Electoral Effects of Biased Media: Russian Television in Ukraine. *American Journal of Political Science*, 62(3):535–550.

Peng, Y. (2022). The Role of Ideological Dimensions in Shaping Acceptance of Facial Recognition Technology and Reactions to Algorithm Bias. *Public Understanding of Science*, 32(2):535–550.

Peters, B., Pierre, J., Sørensen, E., and Torfing, J. (2022). The Promise, Perils and Pitfalls of Digital Governance. In Peters, B., Pierre, J., Sørensen, E., and Torfing, J., editors, *A Research Agenda for Governance*, chapter 8, pages 133–151. Edward Elgar.

Pleines, H. (2023). Levada Polls on the Discuss Data Online Platform: Accessing and Discussing Russian Surveys of Public Opinion. *Russian Analytical Digest*, 292:9–10.

Rana, A. (2014). *The Two Faces of American Freedom*. Cambridge, MA: Harvard University Press.

Reisinger, W., Zaloznaya, M., and Woo, B.-D. (2022). Fear of Punishment as a Driver of Survey Misreporting and Item Non-response in Russia and its Neighbors. *Post-Soviet Affairs*, 39(1-2):49–59.

Robinson, D. and Tannenberg, M. (2019). Self-censorship of Regime Support in Authoritarian States: Evidence From List Experiments in China. *Research & Politics*, 6(3):2053168019856449.

Rochlitz, M., Schoors, K., and Zakharov, N. (2023). Can Authoritarian Propaganda Compete with the Opposition on Social Media? Experimental Evidence from Russia. *SSRN Working Paper No. 4433145*.

Rosenfeld, B. (2022). Survey Research in Russia: in the Shadow of War. *Post-Soviet Affairs*, 39(1-2):38–48.

Rosenfeld, B. (2023). Curious What Russians Think about the War? Ask Yourself This Before You Read the Polls. *Russian Analytical Digest*, 292:4–6.

Schmidli, W. M. (2022). *Freedom on the Offensive: Human Rights, Democracy Promotion, and Us Interventionism in the Late Cold War*. Ithaca: Cornell University Press.

Semukhina, O. and Reynolds, K. M. (2014). Russian citizens' perceptions of corruption and trust of the police. *Policing and Society*, 24(2):158–188.

Snyder, T. (2018). *The Road to Unfreedom: Russia, Europe, America*. Tim Duggan Books, New York, first paperback edition edition.

Steinfeld, N. (2017). Track me, Track me not: Support and Consent to State and Private Sector Surveillance. *Telematics and Informatics*, 34(8):1663–1672.

Strittmatter, K. (2020). *We Have Been Harmonized: Life in China's Surveillance State*. New York: Custom House.

Su, Z., Xu, X., and Cao, X. (2022). What Explains Popular Support for Government Monitoring in China? *Journal of Information Technology & Politics*, 19(4):377–392.

Svolik, M. (2012). *The Politics of Authoritarian Rule*. Cambridge: Cambridge University Press.

Tabellini, M. and Magistretti, G. (2022). Economic Integration and the Transmission of Democracy. *NBER Working Paper No. 30055*.

Tahiroglu, M. (2020). How Turkey's Leaders Dismantled the Rule of Law. *The Fletcher Forum of World Affairs*, 44(1):67–96.

Thompson, N., McGill, T., Bunn, A., and Rukshan, A. (2020). Cultural factors and the role of privacy concerns in acceptance of government surveillance. *Journal of the Association for Information Science and Technology*, 71(9):1129–1142.

Tirole, J. (2021). Digital Dystopia. *American Economic Review*, 111(6):2007–2048.

Toots, M. (2019). Why E-participation Systems Fail: The Case of Estonia's Osale.ee. *Government Information Quarterly*, 36(3):546–559.

Trüdinger, E.-M. and Steckermeier, L. C. (2017). Trusting and Controlling? Political trust, Information and Acceptance of Surveillance Policies: The Case of Germany. *Government Information Quarterly*, 34(3):421–433.

Vasilopoulos, P., McAvay, H., Brouard, S., and Foucault, M. (2022). Emotions, governmental trust, and support for the restriction of civil liberties during the covid-19 pandemic. *European Journal of Political Research*.

Volkov, D. (2023). Are Meaningful Public Opinion Polls Possible in Today's Russia? *Russian Analytical Digest*, 292:2–3.

Williams, S., Armitage, C., Tampe, T., and Dienes, K. (2021). Public Attitudes Towards COVID-19 Contact Tracing Apps: A UK-based Focus Group Study. *Health Expectations*, 24(2):377–385.

Xu, X. (2022). The Unintrusive Nature of Digital Surveillance and Its Social Consequences. *Working Paper*.

Xu, X., Kostka, G., and Cao, X. (2022). Information Control and Public Support for Social Credit Systems in China. *Journal of Politics*, 84(2).

Yang, F. and Xu, J. (2018). Privacy Concerns in China's Smart City Campaign: The Deficit of China's Cybersecurity Law. *Asia & the Pacific Policy Studies*, 5(3):533–543.

Yu, Z. and Wong, S. H.-w. (2023). Social Media, Fear, and Support for State Surveillance: The Case of China's Social Credit System. *China Information*, 37(1):51–74.

Ziller, C. and Helbling, M. (2021). Public Support for State Surveillance. *European Journal of Political Research*, 60(4):994–1006.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Public Affairs.

Østbø, J. (2021). Hybrid Surveillance Capitalism: Sber's Model for Russia's Modernization. *Post-Soviet Affairs*, 37(5):435–452.

3.A Full Questionnaire

As evident below, the questionnaires differ slightly in design for two reasons. First, some questions are only meaningful in the Russian context or are very carefully formulated in order to not be too sensitive for the increasingly repressive context. Second, some questions are only meaningful in other countries and are thus not asked in Russia.

The punishment we used in Russia is very closely designed to resemble Chinese social credit systems. Not being able to use public transportation applies to millions of Chinese citizens who have been blacklisted due to bad credit, mostly for minor offenses like not paying bills.[27] Social credit systems work with data from courts or financial transaction applications. What is new about social credit systems is not the surveillance capabilities – but the capabilities for executive power, i.e., enforcing laws, regulations, and social norms.

3.A.1 CATI Questionnaire in Russia

1) Gender

(Codify without asking the question.)

male/ female

2) Age

[27] see: https://www.scmp.com/economy/china-economy/article/2186606/chinas-social-credit-system-shows-its-teeth-banning-millions

Please tell me, how old are you?

years

3) Education

What is your level of education?

incomplete secondary or lower/ general secondary (school)/ primary vocational (vocational school, lyceum, etc.)/ specialized secondary school (colleges, colleges, technical school, medical school, etc.)/ incomplete higher education (training in a university without a diploma)/ higher education (specialist diploma, bachelor's degree, master's degree, etc.)/ difficult to answer, refusal to answer

4) Employment

What is your occupation?

working/ studying and working/ studying and not working/ retired and working/ retired and not working/ do not work, but I am looking for a job/ do not work and do not look for a job/ other/ find it difficult to answer

5) Income satisfaction

How would you rate your current financial situation?

very good/ good/ average/ poor/ very bad/ difficult to answer

6) Sources of Information

Tell me, please, from what sources do you prefer to learn news: from traditional mass media (TV, radio, newspapers) or from the Internet (news sites, blogs, social networks, messengers)? (If necessary, read positions 1 and 2 again. One answer.)

from traditional media/ from the Internet/ difficult to answer, other answers

7) Residency

Do you live in a city or a village?

city, township/ village/ no answer

Question 8 is asked if item 1 ("town, city-type settlement") or item 3 ("Don't answer") is selected in question 7.

8) Location

What is the name of the town in which you live?

List of cities

Question 9 is asked if you have chosen item 2 ("village") in question 7.

9) District

What district do you live in?

List of districts (villages)

10) Government Approval

In your opinion, in the last year, has Russia been moving in the right or rather the wrong direction?

Rather in the right direction/ Rather the wrong direction/ Difficult to answer, other answer

11) Trust in Public Services

Speaking in general, do you trust or do not trust the authorities providing public services to the population (e.g. services for receiving benefits, certificates, documents)? I will read out the answer options.

Definitely trust/ rather trust/ rather do not trust/ definitely do not trust/ Difficult to answer, other answer

12) Public Service Satisfaction

In the last year, did you or did you not apply to the authorities to receive public services? (Pause after the first part of the question. If the respondent says that he/she did not, mark item 1 and do not read out the second part of the question). If you did, in the last such application were you rather satisfied or rather dissatisfied with the quality of the service?

Not at all/ Rather satisfied/ Rather dissatisfied/ Difficult to answer

13) DGS

Now the idea of creating a unified state digital archive, which would store data about all Russians, is being discussed. It is assumed that the creation of such an archive will increase the efficiency of government bodies and simplify interaction with them for citizens. How do you feel about the idea of creating a single state digital archive - rather positive or rather negative?

Rather positive/ Rather negative/ Difficult to answer

Only one of the questions 14a - 14d will be asked at random.

14a) Comparing information from the digital archive and data from video surveillance cameras will make it possible to find and prosecute those who violate law and order. Would you

personally approve or disapprove of the introduction of such a system in our country? I will read out the answer choices.

would definitely approve/ would rather approve/ would rather not approve/ would definitely disapprove/ find it difficult to answer

14b) Comparing information from the digital archive and data from video surveillance cameras will make it possible to find and prosecute those who violate law and order. For example, a person who broke a bench on a playground could be banned from public transport for a while. Would you personally approve or disapprove of the introduction of such a system in our country? I will read out the answer choices.

would definitely approve/ would rather approve/ would rather not approve/ would definitely disapprove/ find it difficult to answer

14c) Comparing information from the digital archive and data from video surveillance cameras will make it possible to find and prosecute those who violate law and order. For example, a person who participated in an unauthorized protest could be banned from public transport for a while. Would you personally approve or disapprove of the introduction of such a system in our country? I will read out the answer choices.

would definitely approve/ would rather approve/ would rather not approve/ would definitely disapprove/ find it difficult to answer

14d) Comparing information from the digital archive and data from video surveillance cameras will make it possible to find and prosecute those who violate law and order. For example, a person who participated in an unauthorized protest or broke a bench on a playground could be banned from public transport for a while. Would you personally approve or disapprove of the introduction of such a system in our country? I will read out the answer choices.

would definitely approve/ would rather approve/ would rather not approve/ would definitely disapprove/ find it difficult to answer

15) data security

A single digital archive is supposed to store citizens' personal information (employment, tax information, health information, etc.). Would you rather or rather not worry about the security of your personal data in such an archive?

Would rather worry about it/ Would rather not worry/ Difficult to answer

16) Onlince Service Experience

One last question. How often do you use the "Gosuslugi" portal or other online government services? I will read out the answer choices.

never use it/ very rarely use it/ sometimes use it/ often use it/ find it difficult to answer

3.A.2 Online Questionnaire

0) Introduction (Button)

Thank you for your participation! This is a political science survey overseen by David Karpa as the primary researcher from the University of Bremen. Your participation makes a valuable contribution to academic research. It is therefore important that you read all questions closely and answer them truthfully. If you prefer not to respond to a question, please choose the alternative "prefer not to answer". Completing the questionnaire should take approximately 10 minutes. Participation in this survey is voluntary and unauthorized persons will not be given access to your responses. If at any time you wish to quit the survey you may do so by leaving this site or simply closing your browser window. If you wish to participate, please click "Next" below. If you do not wish to participate, please leave this site or close your browser window.

1) Age (Dropdown)

What is your year of birth?

2) Gender (Selection)

How do you describe yourself? *Male/ Female/ Prefer not to answer*

3) Political views (Two-sided slider, 11 scale) (Different wording for different countries to represent the respective poles, this is for the US.)

Here is a scale on which the political views that people might hold are arranged from extremely liberal (left) to extremely conservative (right). Where would you place yourself on this scale? *Liberal/Conservative/Prefer not to answer*

4) Employment (Selection)

What best describes your employment status over the last three months? *Working full-time/ Working part-time/ Unemployed and looking for work/ A homemaker or a stay-at-home parent/ Student/ Retired/ Other/ Prefer not to answer*

5) Income (Selection)

What was your total household income before taxes in Euros in the past 12 months? *Less than $30.000 per year (2.500 $/month)/ $30.000 - $59.999 per year (2.500-5.00 $/month)/ $60.000 - $119.999 per year (5.000-10.000 $/month)/ $120.000 - $239.999 per year (10.000-20.000 $/month)/ More than $240.000 per year (20.000 $/month)/ Prefer not to answer*

6) Income satisfaction (Selection)

How satisfied were you with your household income in the past 12 months? *Definitely satisfied/ Rather satisfied/ Rather dissatisfied/ Definitely dissatisfied/ Prefer not to answer*

7) Ethnicity (Selection) (Different ethnicities for different countries, this is for Germany. US: White, Hispanic and Latino, Black or African American, Asian, Other, Don't want to answer; Turkey: Turkish, Kurdish, Arabic, Greek, Other, Don't want to answer; Estonia: Estonian, Russian, Ukrainian, Belarusians, Other, Don't want to answer)

Which of the following describes you the best? *German/ Turkish/ Russian/ Polish/ Other/ Prefer not to answer*

8) Education (Selection)

What is the highest level of education you have completed? *Some or completed Primary Education/ Some or completed Secondary Education/ Vocational or Similar/ Some University but no degree/ University Bachelor's degree/ Graduate or professional degree (MA, MS, MBA, PhD, JD, MD, DDS etc.)*

9) Residency (Selection)

Which of these best describes the general area where you live? *Urban/ Suburban/ Rural/ Remote*

10) Sources of Information (Multiple Choice)

Which of the following media are your primary sources of information (multiple answers possible)? *Television/News/ Radio/ Social Media/Smartphone/Apps/ Newspaper/ Personal conversations*

11) quality (Selection) (screenout if this does not match with the first question, or if participants are under 18)

How old are you? *Under 18/ Between 18 and 25/ Between 26 and 35/ Between 36 and 45/ Between 46 and 55/ Between 56 and 65/ Older than 65*

12) Country development (Selection)

In your opinion, in the last year, has country been moving more in the right direction or more in the wrong direction? (put in the respective country for country) *Rather in the right direction/ Rather in the wrong direction*

13) trust authorities (Selection)

Generally speaking, do you trust the authorities that provide public services to the population? *Definitely trust/ Rather trust/ Rather do not trust/ Definitely do not trust*

14) government satisfaction (Selection)

During the last year, were you rather satisfied or rather dissatisfied with the quality of government services (for example, getting benefits, certificates, documents)? *Rather satisfied/ Rather dissatisfied*

15) single archive (Selection)

The idea of creating a unified state digital archive, which would contain data about all citizens, is currently being discussed. The creation of such an archive would increase the efficiency of government bodies and simplify interaction with them for citizens. How do you feel about the idea of creating a single digital archive - rather positive or rather negative? *Rather positive/ Rather negative*

Only one of the questions 16a - 16d will be asked at random.

16a) control (Selection)

The comparison of information from the digital archive and data from video surveillance cameras will make it possible to find and prosecute those who violate law and order. Would you personally approve or disapprove of the introduction of such a system in country? *Would definitely approve/ Would rather approve/ Would rather not approve/ Would definitely disapprove*

16b) crime (Selection)

A comparison of information from the digital archive and video surveillance camera data would make it possible to find and prosecute those who violate the law and order. For example, a person who broke a bench on a playground could be held accountable. Would you personally approve or disapprove of the introduction of such a system in country? *Would definitely approve/ Would rather approve/ Would rather not approve/ Would definitely dis-*

approve

16c) repression (Selection)

A comparison of information from the digital archive and data from video surveillance cameras would make it possible to find and prosecute those who violate the law and order. For example, a person who participated in an unauthorized political protest could be held accountable. Would you personally approve or disapprove of the introduction of such a system in country? *Would definitely approve/ Would rather approve/ Would rather not approve/ Would definitely disapprove*

16d) crime and repression (Selection)

A comparison of information from the digital archive and data from video surveillance cameras would make it possible to find and prosecute those who violate the law and order. For example, a person who participated in an unauthorized political protest or broke a bench on a playground could be held accountable. Would you personally approve or disapprove of the introduction of such a system in country? *Would definitely approve/ Would rather approve/ Would rather not approve/ Would definitely disapprove*

17) data security (Selection)

A single digital archive is supposed to store citizens' personal information (employment, tax information, health information, etc.). Would you be concerned about the safety of your data? *Would definitely be concerned/ Would rather be concerned/ Would rather not be concerned/ Would definitely be concerned*

18) online services (Selection)

How often do you use online government services? *Never/ Very rarely/ Sometimes/ Often*

3.B Summary Statistics

Table A1: Summary Statistics Russia

Variable	N	Mean	SD
Control	589	2.886	0.8742
Crime	585	2.838	0.9718
Repression	572	2.395	1.012
Crime and Repression	577	2.7	0.9585
Approval	2323	2.707	0.9732
Age	2462	50.72	16.27
Female	2462	0.5548	0.4971
Employed	2462	0.589	0.4921
Income Satisfaction	2462	2.922	0.8179
Education	2462	4.455	1.565
Urban	2462	0.4188	0.4935
Government Approval	2462	0.7201	0.449
Trust in Public Services	2314	0.7887	0.4083
Public Service Satisfaction	2462	0.3765	0.4846
Onlince Service Experience	2448	2.587	0.9878
State Media	2462	0.4261	0.4946
Online Media	2462	0.3761	0.4845
Moscow	2462	0.07067	0.2563

Table A2: Summary Statistics

ctry	Estonia			Germany			Turkey			US		
Variable	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Control	254	2.9	0.74	255	2.8	0.93	249	2.9	0.95	242	2.3	0.92
Crime	252	2.9	0.81	252	2.7	0.89	250	3	0.95	246	2.4	0.91
Repression	248	2.7	0.86	241	2.6	0.94	263	2.7	1	248	2.3	0.91
Crime and Repression	246	2.8	0.83	252	2.7	0.91	238	2.8	0.98	264	2.3	0.89
Age	1000	43	16	1000	50	17	1000	38	12	1000	48	17
Female	1000	0.54	0.5	1000	0.51	0.5	1000	0.51	0.5	1000	0.5	0.5
Liberal	1000	0.49	0.5	1000	0.46	0.5	1000	0.62	0.48	1000	0.56	0.5
Conservative	1000	0.28	0.45	1000	0.3	0.46	1000	0.25	0.43	1000	0.25	0.43
Education	1000	3.6	1.6	1000	3.8	1.4	1000	4.3	1.2	1000	4.2	1.5
Urban	1000	0.58	0.49	1000	0.44	0.5	1000	0.92	0.27	1000	0.21	0.4
Government Approval	1000	0.57	0.49	1000	0.3	0.46	1000	0.29	0.46	1000	0.3	0.46
Trust in Public Services	1000	0.76	0.43	1000	0.62	0.49	1000	0.32	0.47	1000	0.47	0.5
Public Service Satisfaction	1000	0.73	0.45	1000	0.53	0.5	1000	0.32	0.47	1000	0.47	0.5
Onlince Service Experience	1000	0.78	0.41	1000	0.35	0.48	1000	0.78	0.41	1000	0.36	0.48
Minority	1000	0.098	0.3	1000	0.029	0.17	1000	0.068	0.25	1000	0.21	0.41
TV	1000	0.7	0.46	1000	0.81	0.39	1000	0.76	0.43	1000	0.75	0.43
Radio	1000	0.38	0.49	1000	0.53	0.5	1000	0.25	0.43	1000	0.28	0.45
New Media	1000	0.73	0.44	1000	0.52	0.5	1000	0.9	0.31	1000	0.54	0.5
Newspaper	1000	0.34	0.47	1000	0.4	0.49	1000	0.36	0.48	1000	0.26	0.44
Conversations	1000	0.26	0.44	1000	0.33	0.47	1000	0.28	0.45	1000	0.27	0.45

Table A3: Summary Statistics Approval

ctry	Estonia			Germany			Turkey			US			Russia		
Variable	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
General Approval (binary)	1000	0.7	0.46	1000	0.62	0.49	1000	0.64	0.48	1000	0.45	0.5	2462	0.61	0.49
General Approval (1 - 4 scale)	1000	2.8	0.82	1000	2.7	0.92	1000	2.8	0.98	1000	2.3	0.91	2323	2.7	0.97
Control Approval (binary)	254	0.75	0.44	255	0.66	0.48	249	0.67	0.47	242	0.46	0.5	623	0.7	0.46
Crime Approval (binary)	252	0.76	0.43	252	0.63	0.48	250	0.7	0.46	246	0.51	0.5	603	0.67	0.47
Repression Approval (binary)	248	0.61	0.49	241	0.57	0.5	263	0.58	0.49	248	0.41	0.49	615	0.46	0.5
Crime and Repression Approval (binary)	246	0.69	0.46	252	0.61	0.49	238	0.61	0.49	264	0.44	0.5	621	0.6	0.49

Table A4: V-dem Indicator

Countrys[a]	Electoral Democracy	Liberal Democracy	Participatory Democracy	Deliberative Democracy	Egalitarian Democracy
Estonia	0.893	0.847	0.619	0.754	0.798
Germany	0.857	0.807	0.615	0.816	0.792
United States	0.819	0.741	0.582	0.711	0.582
Turkey	0.276	0.118	0.164	0.084	0.209
Russia	0.209	0.071	0.142	0.073	0.15

[a]2022 Edition; Index ranging from 0 to 1; https://v-dem.net/

Table A5: Additional Institutional Indicators

Index	**Global Freedom**[a]	**Internet Freedom**[b]	**Polity5**[c]	**Democracy Index**[d]
Russia	19 Not Free	23 Not Free	4 Open Anocracy	2.28 Authoritarian Regime
Estonia	94 Free	93 Free	9 Democracy	7.96 Flawed Democracy
Germany	94 Free	77 Free	10 Full Democracy	8.80 Full Democracy
United States	83 Free	76 Free	8 Democracy	7.85 Flawed Democracy
Turkey	32 Not Fee	32 Not Free	-4 Closed Anocracy	4.35 Hybrid Regime

[a]2022 Edition; Index ranging from 0 to 100; https://freedomhouse.org/countries/freedom-net/scores

[b]2022 Edition; Index ranging from 0 to 100; https://freedomhouse.org/countries/freedom-net/scores

[c]2018 Edition; Index ranging from -10 to 10; https://www.systemicpeace.org/inscrdata.html

[d]2022 Edition; Index ranging from 0 to 10; https://www.eiu.com/n/campaigns/democracy-index-2022/

3.C Regression Tables

Table A6: Russia

	Approval For Digital Governance Solution				
	(1)	(2)	(3)	(4)	(5)
Crime	−0.041* (0.024)	−0.043* (0.024)	−0.056** (0.023)	−0.039 (0.024)	−0.091** (0.038)
Repression	−0.225*** (0.024)	−0.229*** (0.024)	−0.239*** (0.023)	−0.226*** (0.024)	−0.303*** (0.038)
Crime and Repression	−0.109*** (0.024)	−0.107*** (0.024)	−0.115*** (0.023)	−0.108*** (0.024)	−0.139*** (0.040)
Age		0.100*** (0.022)	0.063*** (0.023)	0.094*** (0.022)	0.084** (0.041)
Female		0.041** (0.020)	0.030 (0.019)	0.037* (0.020)	0.011 (0.032)
Employed		0.012 (0.022)	0.015 (0.022)	0.007 (0.023)	0.020 (0.037)
Income Satisfaction		0.128*** (0.020)	0.058*** (0.020)	0.125*** (0.020)	0.057* (0.031)
Education		−0.075*** (0.020)	−0.062*** (0.021)	−0.073*** (0.021)	−0.028 (0.037)
Urban		−0.001 (0.021)	0.007 (0.020)	−0.016 (0.021)	−0.054 (0.034)
Moscow		−0.064*** (0.023)	−0.052** (0.022)	−0.061*** (0.023)	−0.042 (0.034)
Government Approval			0.218*** (0.021)		0.283*** (0.035)
Trust in Public Services			0.117*** (0.021)		0.045 (0.041)
Public Service Satisfaction			−0.009 (0.021)		0.007 (0.038)
Onlince Service Experience			0.044** (0.022)		0.039 (0.039)
State Media			0.091*** (0.022)		0.110*** (0.038)
Constant	2.724*** (0.020)	2.724*** (0.020)	2.730*** (0.019)	2.745*** (0.020)	2.810*** (0.035)
Coding		Don't know = No	Don't know = No	Don't know = NA	Don't know = NA
N	2,323	2,323	2,199	2,196	745
R^2	0.041	0.078	0.180	0.080	0.261
Adjusted R^2	0.040	0.074	0.174	0.076	0.246
Residual Std. Error	0.959 (df = 2319)	0.941 (df = 2312)	0.880 (df = 2183)	0.934 (df = 2185)	0.854 (df = 729)

*p < .1; **p < .05; ***p < .01
Columns (2) and (3) use a different coding of variables as the columns (4) and (5). In (2) and (3), "Don't know" and "Find it difficult to answer" replies have been coded as "No", whereas in (4) and (5) these replies have been coded as missings in the data. Accordingly, N is smaller in (4) and (5)

Table A7: Russia (interactions with treatments)

	Approval For Digital Governance Solution				
	(1)	(2)	(3)	(4)	(5)
Crime	−0.040* (0.024)	−0.052** (0.023)	−0.054** (0.024)	−0.040* (0.024)	−0.042* (0.024)
Repression	−0.228*** (0.024)	−0.233*** (0.023)	−0.235*** (0.024)	−0.224*** (0.024)	−0.225*** (0.024)
Crime and Repression	−0.107*** (0.024)	−0.120*** (0.023)	−0.112*** (0.024)	−0.107*** (0.024)	−0.109*** (0.024)
State Media	0.178*** (0.020)				
State Media:Crime	0.004 (0.024)				
State Media:Repression	0.004 (0.024)				
State Media:Crime and Repression	0.020 (0.024)				
Government Approval		0.287*** (0.019)			
Government Approval:Crime		0.013 (0.023)			
Government Approval:Repression		0.022 (0.023)			
Government Approval:Crime and Repression		0.003 (0.023)			
Trust in Public Services			0.220*** (0.019)		
Trust in Public Services:Crime			−0.013 (0.024)		
Trust in Public Services:Repression			−0.018 (0.024)		
Trust in Public Services:Crime and Repression			−0.010 (0.024)		
Public Service Satisfaction				0.017 (0.020)	
Public Service Satisfaction:Crime				−0.023 (0.024)	
Public Service Satisfaction:Repression				−0.022 (0.024)	
Public Service Satisfaction:Crime and Repression				−0.008 (0.025)	
Onlince Service Experience					0.019 (0.020)
Onlince Service Experience:Crime					−0.063** (0.025)
Onlince Service Experience:Repression					−0.030 (0.025)
Onlince Service Experience:Crime and Repression					−0.022 (0.024)
Constant	2.725*** (0.020)	2.723*** (0.019)	2.743*** (0.020)	2.725*** (0.020)	2.724*** (0.020)
N	2,323	2,323	2,204	2,323	2,316
R^2	0.074	0.127	0.096	0.042	0.044
Adjusted R^2	0.071	0.124	0.093	0.039	0.041
Residual Std. Error	0.943 (df = 2315)	0.915 (df = 2315)	0.922 (df = 2196)	0.959 (df = 2315)	0.958 (df = 2308)

*p < .1; **p < .05; ***p < .01

Table A8: Estonia

	Approval For Digital Governance Solution	
	(1)	(2)
Crime	0.024 (0.031)	0.034 (0.029)
Repression	−0.087*** (0.031)	−0.082*** (0.029)
Crime and Repression	−0.049 (0.031)	−0.044 (0.029)
Age		0.030 (0.028)
Female		0.065** (0.025)
Employed		−0.005 (0.024)
Income Satisfaction		0.046* (0.026)
Education		−0.055** (0.025)
Urban		−0.003 (0.024)
Government Approval		0.198*** (0.028)
Trust in Public Services		0.048 (0.030)
Public Service Satisfaction		0.103*** (0.028)
Onlince Service Experience		−0.047* (0.024)
Conservative		−0.015 (0.024)
Minority		0.016 (0.025)
TV		0.087*** (0.026)
Radio		0.018 (0.026)
New Media		0.007 (0.025)
Newspaper		−0.104*** (0.027)
Conversations		−0.030 (0.025)
Constant	2.810*** (0.026)	2.810*** (0.024)
N	1,000	1,000
R^2	0.015	0.184
Adjusted R^2	0.012	0.167
Residual Std. Error	0.810 (df = 996)	0.744 (df = 979)
F Statistic	4.986*** (df = 3; 996)	11.001*** (df = 20; 979)

*p < .1; **p < .05; ***p < .01

Table A9: Turkey

	Approval For Digital Governance Solution	
	(1)	(2)
Crime	0.075** (0.038)	0.085** (0.034)
Repression	−0.089** (0.038)	−0.083** (0.034)
Crime and Repression	−0.028 (0.038)	−0.045 (0.034)
Age		−0.084*** (0.029)
Female		0.014 (0.028)
Employed		0.068** (0.030)
Income Satisfaction		0.149*** (0.032)
Education		−0.032 (0.029)
Urban		−0.021 (0.028)
Government Approval		0.063 (0.041)
Trust in Public Services		0.083** (0.042)
Public Service Satisfaction		0.204*** (0.042)
Onlince Service Experience		0.052* (0.028)
Conservative		0.046 (0.029)
Minority		−0.052* (0.028)
TV		0.014 (0.030)
Radio		−0.0001 (0.032)
New Media		0.044 (0.029)
Newspaper		−0.027 (0.033)
Conversations		0.024 (0.029)
Constant	2.834*** (0.031)	2.834*** (0.027)
N	1,000	1,000
R^2	0.019	0.234
Adjusted R^2	0.016	0.218
Residual Std. Error	0.974 (df = 996)	0.868 (df = 979)
F Statistic	6.509*** (df = 3; 996)	14.942*** (df = 20; 979)

*p < .1; **p < .05; ***p < .01

Table A10: Germany

	Approval For Digital Governance Solution	
	(1)	(2)
Crime	−0.049 (0.035)	−0.048 (0.033)
Repression	−0.080** (0.035)	−0.095*** (0.033)
Crime and Repression	−0.037 (0.035)	−0.046 (0.034)
Age		0.055 (0.034)
Female		0.023 (0.029)
Employed		−0.031 (0.030)
Income Satisfaction		0.090*** (0.029)
Education		−0.082*** (0.028)
Urban		−0.002 (0.028)
Government Approval		0.068** (0.031)
Trust in Public Services		0.127*** (0.034)
Public Service Satisfaction		0.080** (0.033)
Onlince Service Experience		0.078*** (0.028)
Conservative		0.035 (0.028)
Minority		0.065** (0.028)
TV		0.082*** (0.029)
Radio		0.092*** (0.029)
New Media		0.054* (0.031)
Newspaper		0.021 (0.031)
Conversations		−0.007 (0.029)
Constant	2.693*** (0.029)	2.693*** (0.027)
N	1,000	1,000
R^2	0.005	0.141
Adjusted R^2	0.002	0.124
Residual Std. Error	0.917 (df = 996)	0.860 (df = 979)
F Statistic	1.755 (df = 3; 996)	8.048*** (df = 20; 979)

*p < .1; **p < .05; ***p < .01

Table A11: USA

	Approval For Digital Governance Solution	
	(1)	(2)
Crime	0.041 (0.035)	0.045 (0.033)
Repression	−0.014 (0.035)	−0.020 (0.033)
Crime and Repression	−0.022 (0.036)	−0.005 (0.033)
Age		−0.053 (0.034)
Female		0.079*** (0.028)
Employed		0.014 (0.029)
Income Satisfaction		0.083*** (0.028)
Education		−0.069** (0.028)
Urban		0.056** (0.028)
Government Approval		0.035 (0.032)
Trust in Public Services		0.175*** (0.034)
Public Service Satisfaction		0.140*** (0.034)
Onlince Service Experience		0.029 (0.027)
Conservative		0.029 (0.027)
Minority		0.076*** (0.028)
TV		0.101*** (0.028)
Radio		0.001 (0.027)
New Media		0.003 (0.031)
Newspaper		−0.016 (0.028)
Conversations		−0.030 (0.027)
Constant	2.337*** (0.029)	2.337*** (0.026)
N	1,000	1,000
R^2	0.004	0.177
Adjusted R^2	0.001	0.160
Residual Std. Error	0.907 (df = 996)	0.832 (df = 979)
F Statistic	1.298 (df = 3; 996)	10.535*** (df = 20; 979)

*p < .1; **p < .05; ***p < .01

Table A12: Treatment effects full sample

	Approval For Digital Governance Solution		
	(1)	(2)	(3)
Crime	0.006 (0.014)	−0.0003 (0.014)	0.001 (0.013)
Repression	−0.121*** (0.014)	−0.126*** (0.014)	−0.128*** (0.013)
Crime and Repression	−0.052*** (0.014)	−0.054*** (0.014)	−0.054*** (0.013)
US	−0.133*** (0.013)	−0.105*** (0.013)	−0.021 (0.019)
Turkey	0.047*** (0.013)	0.123*** (0.014)	0.213*** (0.019)
Germany	0.037*** (0.013)	0.069*** (0.013)	0.070*** (0.018)
Estonia	−0.005 (0.013)	0.018 (0.013)	0.083*** (0.019)
(Reference=Russia)			
Age		0.046*** (0.013)	0.039*** (0.012)
Female		0.041*** (0.012)	0.040*** (0.011)
Employed		0.009 (0.012)	0.006 (0.012)
Income Satisfaction		0.189*** (0.012)	0.104*** (0.012)
Education		−0.066*** (0.012)	−0.080*** (0.012)
Urban		0.006 (0.013)	−0.003 (0.012)
Government Approval			0.179*** (0.014)
Trust in Public Services			0.147*** (0.014)
Public Service Satisfaction			0.048*** (0.013)
Onlince Service Experience			0.023 (0.021)
Constant	2.683*** (0.012)	2.682*** (0.011)	2.683*** (0.011)
N	6,323	6,323	6,199
R^2	0.044	0.083	0.166
Adjusted R^2	0.043	0.081	0.164
Residual Std. Error	0.925 (df = 6315)	0.906 (df = 6309)	0.862 (df = 6181)
F Statistic	41.375*** (df = 7; 6315)	44.075*** (df = 13; 6309)	72.365*** (df = 17; 6181)

*p < .1; **p < .05; ***p < .01

Table A13: Interactions effects age and treatments

	Approval For Digital Governance Solution				
	(1)	(2)	(3)	(4)	(5)
Crime	0.026 (0.031)	−0.049 (0.036)	0.079** (0.038)	0.042 (0.035)	−0.020 (0.024)
Repression	−0.088*** (0.031)	−0.085** (0.035)	−0.096** (0.038)	−0.016 (0.035)	−0.213*** (0.024)
Crime and Repression	−0.047 (0.031)	−0.039 (0.035)	−0.027 (0.037)	−0.022 (0.036)	−0.080*** (0.024)
Age	−0.411*** (0.152)	0.096 (0.176)	0.065 (0.185)	−0.341* (0.174)	0.179 (0.116)
Age square	0.424*** (0.152)	−0.006 (0.177)	−0.202 (0.185)	0.301* (0.174)	−0.063 (0.117)
Age square:Crime	−0.021 (0.032)	−0.018 (0.036)	0.001 (0.038)	−0.007 (0.035)	0.019 (0.024)
Age square:Repression	0.037 (0.031)	−0.011 (0.036)	−0.068* (0.038)	0.072** (0.035)	−0.003 (0.024)
Age square:Crime and Repression	−0.004 (0.031)	−0.006 (0.037)	−0.032 (0.037)	0.021 (0.036)	−0.018 (0.024)
Constant	2.811*** (0.026)	2.693*** (0.029)	2.831*** (0.031)	2.337*** (0.029)	2.707*** (0.020)
Country	Estonia	Germany	Turkey	United States	Russia
N	1,000	1,000	1,000	1,000	2,323
R^2	0.026	0.015	0.043	0.015	0.054
Adjusted R^2	0.018	0.007	0.035	0.007	0.051
Residual Std. Error	0.808 (df = 991)	0.915 (df = 991)	0.964 (df = 991)	0.904 (df = 991)	0.948 (df = 2314)

*p < .1; **p < .05; ***p < .01

Table A14: Interaction effects political views and treatments

	Approval For Digital Governance Solution			
	(1)	(2)	(3)	(4)
Crime	0.024 (0.031)	−0.046 (0.036)	0.076** (0.038)	0.044 (0.035)
Repression	−0.087*** (0.031)	−0.076** (0.035)	−0.091** (0.038)	−0.013 (0.035)
Crime and Repression	−0.049 (0.031)	−0.033 (0.036)	−0.024 (0.037)	−0.020 (0.036)
Liberal	0.008 (0.033)	−0.037 (0.036)	0.034 (0.045)	−0.063* (0.037)
Conservative	−0.038 (0.033)	0.008 (0.036)	0.157*** (0.045)	−0.050 (0.038)
Liberal:Crime	−0.020 (0.040)	0.046 (0.043)	0.078 (0.054)	−0.075 (0.046)
Liberal:Repression	−0.054 (0.039)	0.047 (0.045)	0.020 (0.054)	−0.112** (0.047)
Liberal:Crime and Repression	−0.008 (0.040)	0.034 (0.044)	0.052 (0.054)	−0.134*** (0.046)
Conservative:Crime	−0.016 (0.040)	−0.011 (0.044)	0.057 (0.055)	−0.049 (0.047)
Conservative:Repression	−0.008 (0.040)	−0.011 (0.044)	0.056 (0.053)	−0.039 (0.047)
Conservative:Crime and Repression	0.023 (0.040)	0.003 (0.044)	0.083 (0.055)	−0.035 (0.047)
Constant	2.810*** (0.026)	2.692*** (0.029)	2.831*** (0.031)	2.336*** (0.029)
Country	Estonia	Germany	Turkey	United States
N	1,000	1,000	1,000	1,000
R^2	0.021	0.011	0.043	0.019
Adjusted R^2	0.010	−0.0005	0.032	0.008
Residual Std. Error (df = 988)	0.811	0.919	0.966	0.904
F Statistic (df = 11; 988)	1.959**	0.955	3.998***	1.774*

*p < .1; **p < .05; ***p < .01

Table A15: Interactions effects government approval and treatments

	Approval For Digital Governance Solution				
	(1)	(2)	(3)	(4)	(5)
Crime	0.028 (0.030)	−0.053 (0.035)	0.081** (0.036)	0.044 (0.035)	−0.052** (0.023)
Repression	−0.092*** (0.030)	−0.082** (0.035)	−0.085** (0.036)	−0.015 (0.035)	−0.233*** (0.023)
Crime and Repression	−0.051* (0.030)	−0.040 (0.035)	−0.038 (0.035)	−0.009 (0.035)	−0.120*** (0.023)
Government Approval	0.263*** (0.024)	0.154*** (0.029)	0.324*** (0.029)	0.206*** (0.028)	0.287*** (0.019)
Government Approval: Crime	−0.040 (0.030)	−0.060* (0.035)	−0.039 (0.036)	0.009 (0.034)	0.013 (0.023)
Government Approval: Repression	0.024 (0.030)	−0.039 (0.035)	0.023 (0.036)	0.020 (0.034)	0.022 (0.023)
Government Approval: Crime and Repression	0.011 (0.030)	−0.051 (0.035)	0.009 (0.035)	0.013 (0.035)	0.003 (0.023)
Constant	2.808*** (0.024)	2.694*** (0.029)	2.833*** (0.029)	2.337*** (0.028)	2.723*** (0.019)
Country	Estonia	Germany	Turkey	United States	Russia
N	1,000	1,000	1,000	1,000	2,323
R^2	0.123	0.037	0.132	0.055	0.127
Adjusted R^2	0.117	0.030	0.126	0.049	0.124
Residual Std. Error	0.766 (df = 992)	0.905 (df = 992)	0.918 (df = 992)	0.886 (df = 992)	0.915 (df = 2315)

*p < .1; **p < .05; ***p < .01

3.D Additional Figures

Figure A1: Results USA with different age groups

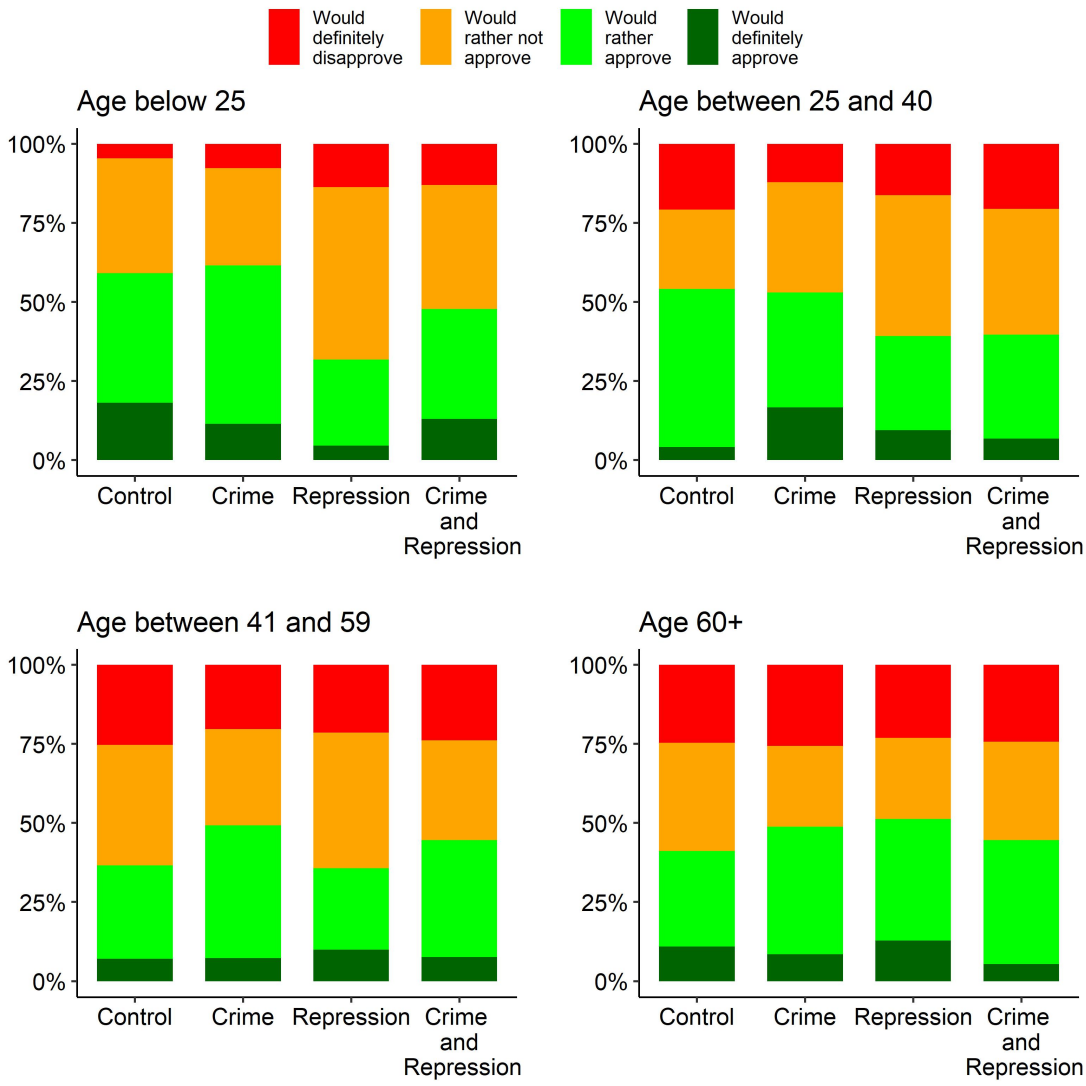
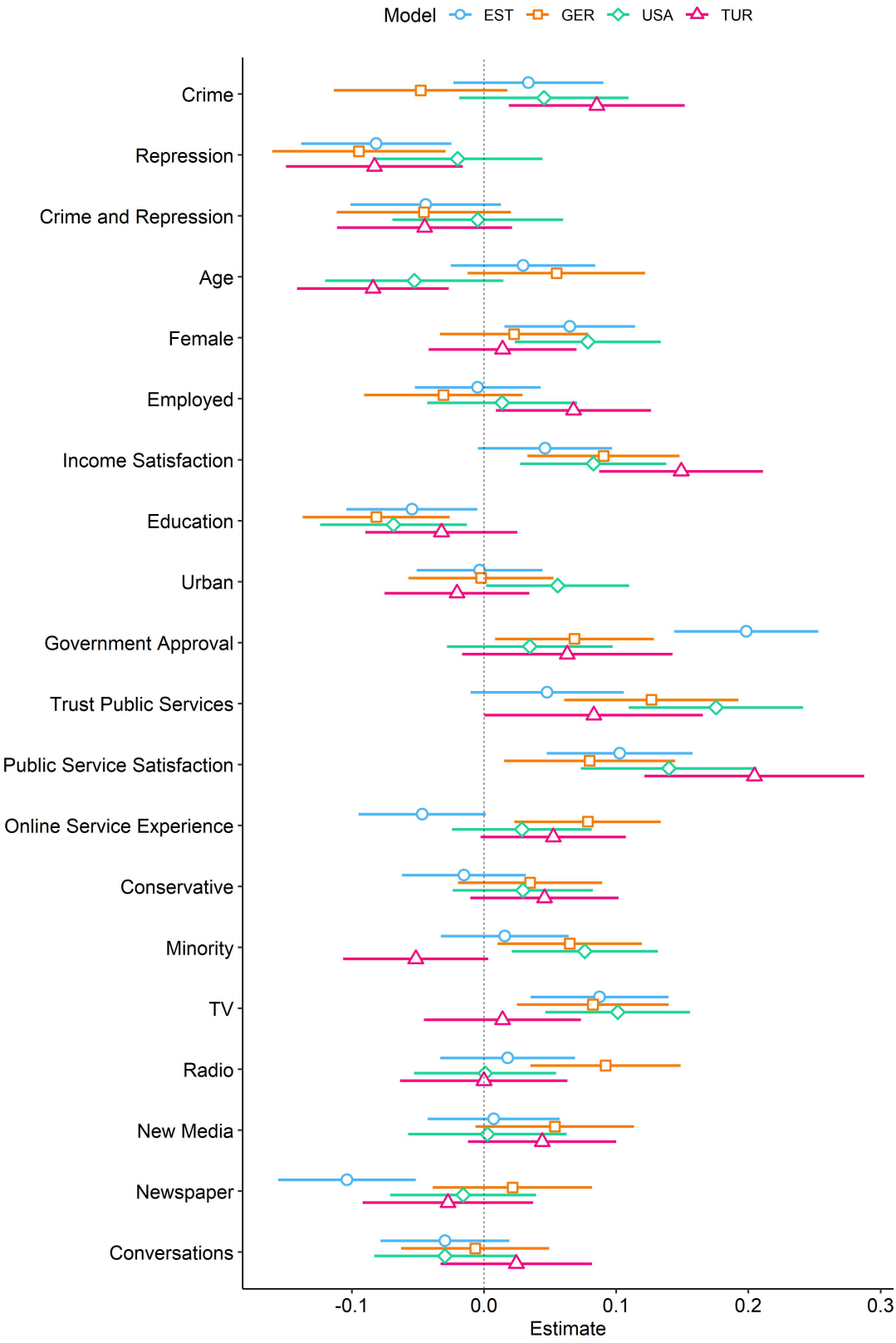


Figure A2: Treatment design

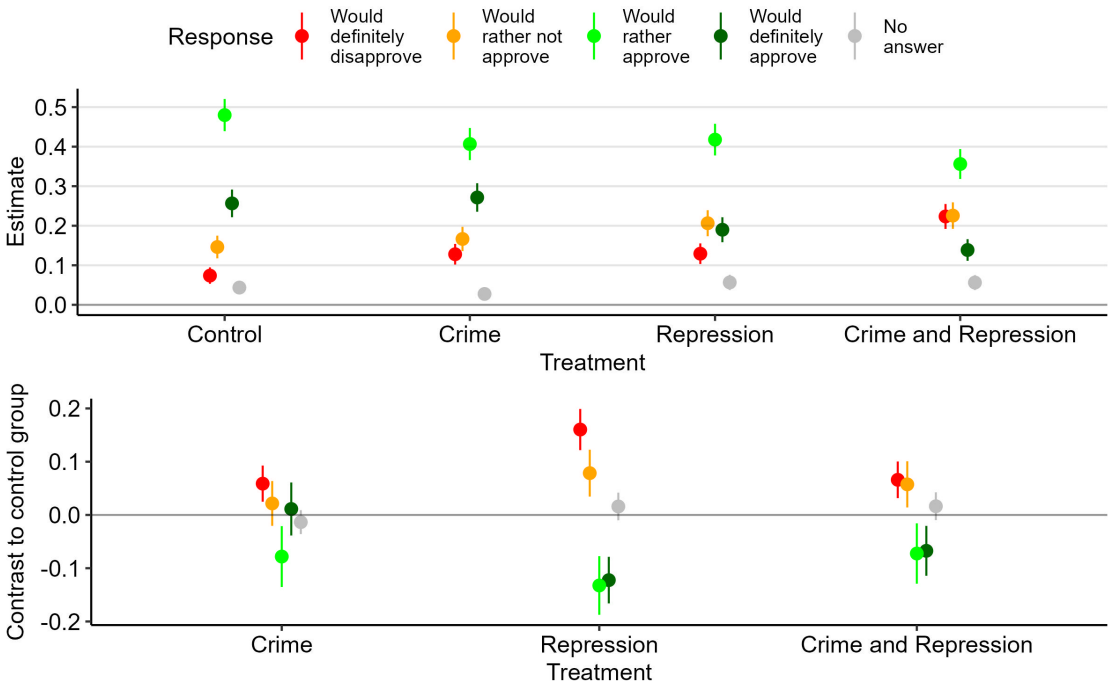
Political repression Information	Policing Crimes Information: No	Policing Crimes Information: Yes
No	Comparing information from the digital archive and data from video surveillance cameras will make it possible to find and prosecute those who violate law and order. Would you personally approve or disapprove of the introduction of such a system in our country? I will read out the answer choices.	...those who violate law and order. **For example, a person who broke a bench on a playground could be banned from public transport for a while.** Would you personally approve or...
Yes	...those who violate law and order. **For example, a person who participated in an unauthorized protest could be banned from public transport for a while.** Would you personally approve or...	...those who violate law and order. **For example, a person who participated in an unauthorized protest or broke a bench on a playground could be banned from public transport for a while.** Would you personally approve or...

Figure A3: Treatment effects in Estonia, Germany, United States, and Turkey



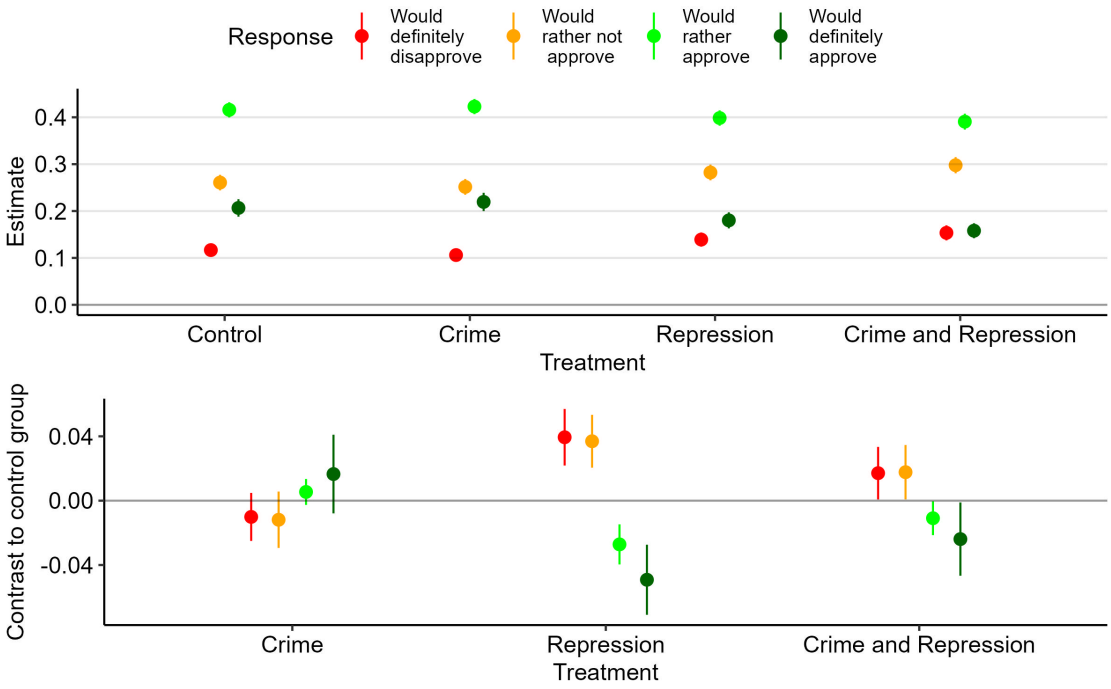
Notes: Standardized OLS estimates with 95% confidence intervals. See appendix 3.C for the underlying regression results.

Figure A4: Treatment effects in Russia



Notes: Multinomial regression coefficients with 95% confidence intervals. Full set of control variables used as in the main analysis in section 3.4.

Figure A5: Treatment effects in Estonia, Germany, United States, and Turkey



Notes: Ordered logistic regression coefficients with 95% confidence intervals. Full set of control variables used as in the main analysis in section 3.4, in addition to country fixed effects.

4 Propaganda and Sanction Evasion

Essay title: Language Affinity, Russian Media and Circumvention of Trade Sanctions: Evidence from Kazakhstan

Authors: Andrey Tkachenko, David Karpa, Meruyert Tatkeyeva, Michael Rochlitz, Galiya Sagyndykova

Abstract

Does language affinity facilitate the export of propaganda by authoritarian regimes? Through which channels can propaganda be exported? And what are its effects? We study these questions with the help of six survey waves conducted in Kazakhstan between May 2021 and September 2023, as well as an original survey conducted in October 2023. We find that for surveys conducted *after* February 2022, Russian-speaking respondents were more likely to justify Russia's war in Ukraine, to negate Russian responsibility and to blame Western countries for the war, even after controlling for ethnicity, while before February 2022 no measurable difference between Russian-speaking and other respondents can be found. We document that these shifts of opinion correspond to a significant increase in the consumption of Russian internet and social media content among Russian-speakers after February 2022, suggesting that this might be the channel through which Russian propaganda was effective. Finally, we show how consumers of Russian propaganda are more likely to support the war and condone the circumvention of sanctions against Russia, and document that war-relevant exports to Russia increased disproportionately in those Kazakh regions where citizens were more likely to consume Russian propaganda and condone the war.

Keywords: Authoritarian propaganda, language, sanctions circumvention, Kazakhstan, Russia-Ukraine war

JEL Classification: D84, P16, P52

"Russkiy mir" – the Russian world, a previously obscure historical term for a Slavic civilisation based on shared ethnicity, religion and heritage. The Putin regime has revived, promulgated and debased this idea into an obscurantist anti-Western mixture of Orthodox dogma, nationalism, conspiracy theory and security-state Stalinism.

(The Economist)

4.1 Introduction

Even after their demise, empires can keep an influence over their former colonies through language. English – the language of the former British Empire – has become a global language that has developed into an important channel through which Western concepts and values are distributed across the world (Phillipson 2008; Crystal 2012). Similarly, through "la Francophonie", France has managed to promote its economic, political and cultural interests in its former colonies, many years after the end of the French empire (Neathery-Castro and Rousseau 2005; Vigouroux 2013). Often, the mechanism through which values and concepts are communicated are not even direct government communications or propaganda, but citizens of former colonies self-selecting into consuming specific content because the language they speak in their everyday lives is the one of the former colonizing nation.[1]

In this paper, we explore this channel and its relevance in the Russia-Ukraine war for Kazakhstan – one of Russia's former colonies in Central Asia. Kazakhstan is particular in that more than 30 years after its independence and despite the state language being Kazakh, the first language of a large percentage of the population remains Russian. Many of these Russian-speaking citizens of Kazakhstan are torn between loyalty to their home country, and a close cultural affinity to what Russian state media is increasingly describing as the "Russkiy mir", or "Russian world". Through their language, they are also exposed to Russian propaganda, much more than their fellow citizens whose first language is not Russian.

We investigate whether this language affinity makes the Russian-speaking population in Kazakhstan more vulnerable to Russian propaganda than the part of the population that mainly speaks Kazakh in their daily lives. By controlling for ethnicity and other demographic

[1] A good case in point are Radio France Internationale or the BBC World Service, which for many years continued to be crucial sources of information in former French and British colonies in sub-Saharan Africa, long after these countries had become independent (Vaillant 2017; Leyris 2022).

characteristics, we focus specifically on language as a channel for propaganda transmission and study its political and economic consequences. Specifically, we investigate the role of Russian propaganda in political beliefs, such as the justification of the Russian full-scale invasion of Ukraine and the willingness to help Russia circumvent sanctions, as well as its role in undertaking actions to circumvent sanctions imposed on Russia.

To answer these questions, we use data from six waves of a telephone survey biannually conducted by the Central Asia Barometer (CAB) in Kazakhstan.[2] The first two waves were carried out before the Russian full-scale invasion of Ukraine in the Spring and Autumn of 2021; the other waves were carried out after the invasion in the Spring and Autumn of 2022 and 2023. We complement this data with an original online survey conducted in November 2023, containing several list experiments. In addition, we use Yandex search data to document patterns of consumption for Russian language news around the time of Russia's full-scale invasion of Ukraine. The set of surveys and list experiments allow us to identify the effect of language and media consumption on political beliefs. Finally, we use transaction-level export data to investigate whether language patterns and shifts in public opinion translate into an increase in war-related exports from Kazakhstan to Russia, after the start of the Russian invasion.

We investigate *potential channels* that could have an effect on public opinion. Looking at the various survey waves of the CAB survey, we find that consumption of news from the internet and Russian social media is consistently higher among the Russian-speaking population in our sample, while consumption of Russian TV news has declined among this group since the start of the war, and the use of Telegram has increased. As various pro-Russian Telegram channels have been particularly prominent in covering the war (see e.g. Oleinik 2024), we believe Telegram in particular to be a potential channel influencing public opinion among Kazakhstan's Russian-speaking population. In addition, we also document a spike in searches for Russian-language information channels in Kazakhstan from February 2022 onwards. This spike is more pronounced in regions with a larger share of Russian speakers.

We hypothesize that the consumption of Russian media positively affects support for the

[2]https://ca-barometer.org/en/cab-database

war, and indeed, find that 4 months after the start of Russia's full-scale invasion of Ukraine, Russian-speaking respondents in Kazakhstan were 18% more likely to justify Russia's war in Ukraine than non-Russian-speaking respondents. They were also 8.8% less likely to attribute responsibility for the war to Russia, and 12% more likely to attribute responsibility to the US, the EU and NATO. Although these results fluctuate somewhat over time, they remain broadly stable until the end of 2023. As questions about the war were only asked from June 2022 onwards, we gauge the evolution of opinions over time by using a question about general attitudes towards Russia, the US and the Eurasian Economic Union, which was asked in all survey waves. Here we see that opinions towards Russia only diverged *after* February 2022, but not before. Before the invasion, Russian-speaking respondents in Kazakhstan even had a significantly more positive view of the United States than non-Russian-speakers.

In a next step, we then use the survey experiments from our original survey to pinpoint the effect of Russian media consumption on political beliefs. Both in the direct question and when using list-experiments, we find that consumers of Russian-language media are more likely to condone Russia's actions in the war, and to support the circumvention of economic sanctions against Russia. Crucially, we are able to hold the effects for ethnicity and other demographics constant, thus being able to exclude other alternative mechanisms in play.

Finally, using detailed regional trade data we identify a spike in sanctions-related exports from Kazakhstan to Russia, after February 2022. This spike is particularly pronounced in Kazakh regions with larger shares of Russian speakers, as well as larger shares of people trusting Russian media and condoning Russia's actions in the war.

Taken together, our interpretation of these findings is the following. When the start of the full-scale Russian invasion of Ukraine triggered interest for the event in Kazakhstan, those citizens in Kazakhstan mainly using Russian in their daily lives were more likely than other Kazakh citizens to end up on the Russian internet and social media platforms, in their search for information. Watching the propaganda-content on these platforms led to a divergence in opinions about the war and about Russia's role in global politics in Kazakhstan, a divergence that so did not exist before February 2022. This divergence was then also reflected in a higher propensity to condone the circumvention of economic sanctions against Russia, potentially convincing a certain number of Kazakh entrepreneurs to participate in making war-related

deliveries to Russia. Importantly, we do not argue that Russian propaganda was the main or the only reason responsible for Kazakhstan's prominent role in sanctions circumvention during the early months of the war (Borozna and Kochtcheeva 2024). But the empirical evidence presented in this paper lets us believe that Russian propaganda played a meaningful role in shaping attitudes and behaviour.

Our paper builds on a vast literature that has investigated the effects of the media and political propaganda on political attitudes and voting. The effect of traditional media such as TV and newspapers has been extensively documented, both for democracies (DellaVigna and Kaplan 2007; Gerber et al. 2009; Durante and Knight 2012) and autocracies (Yanagizawa-Drott 2014; Adena et al. 2015; Peisakhin and Rozenas 2018; Mattingly and Yao 2022; Pan et al. 2022). Notably, a literature focusing on Russia has shown that on traditional platforms such as TV news, media that are independent from the government are able to oppose the government effectively and influence voting behavior, even in authoritarian states (Enikolopov et al. 2011, 2022).

More recently, social media has started to play an important role in influencing public opinions (Bond et al. 2012). Less accessible to government control, social media also lack the reputation mechanisms that ensure minimum content quality and are thus prone to the spread of misinformation and polarization (Zhuravskaya et al. 2020). While social media has been used by the opposition to organize protests (Enikolopov et al. 2020) or mobilize voters (Enikolopov et al. 2022), and has been hailed as a "liberation technology" (Morozov 2011), authoritarian states also increasingly spread their messages through social media. Authoritarian governments employ different methods ranging from "flooding" popular platforms and hashtags with positive or distracting misinformation to shaping public opinion by providing different – and sometimes objectively wrong – narratives (Roberts 2018).

Beyond influencing the domestic population, research has shown that authoritarian propaganda is also increasingly spreading beyond borders, as a tool of foreign influence. China, for example, has been extensively criticized for spreading misinformation related to Covid-19 and other topics.[3] Scholars have argued that for Russia in particular, influencing popular

[3]https://www.grid.news/story/global/2022/05/18/how-china-uses-global-media-to-spread-its-views-and-misinformation/

opinion and elections abroad has become an important political objective (Snyder 2018). Indeed, there is empirical evidence for Russia being the most active autocracy in this respect (Martin et al. 2019), with the Brexit referendum and the 2016 presidential elections in the US as two prominent examples (Martin et al. 2019; Eady et al. 2023).

As part of this strategy, Russian diasporas and Russian-speaking communities abroad have become an important tool in the Kremlin's foreign policy strategy. For example, since February 2022 several pro-Putin events that were then joined by far-right extremists were organized in Germany by the Russian diaspora.[4] A similar pattern can be observed in the US, where the Russian-speaking community was active in spreading a pro-Kremlin agenda.[5] Russian foreign influence is thus not limited to its neighboring countries, but can be described as a worldwide phenomenon. Sometimes, the results of exposure to Russian propaganda might not even be intentional, as in the case of Latvia, where the discrediting of Western vaccines on Russian TV led to a significant drop in vaccination rates among the ethnic Russian population (Larreguy and Martinez 2024).

However, with the exception of Larreguy and Martinez (2024), empirical research on the specific effects, influence and channels of Russian propaganda on Russian-speaking communities abroad remains rare. It is this research gap that we try to fill with this paper, by investigating one context where the Russian-speaking community plays a particular important role – the former Soviet republic of Kazakhstan in Central Asia.

Theoretically, the first part of our paper is based on Guiso and Makarin (2020), in that we hypothesize that Russian speakers think of information conveyed in Russian as *trustworthy*. According to this theory, genetic, cultural, or religious affinity enhances trust (Guiso and Makarin 2020). In our paper, we denote language as – perhaps the most important – trait of cultural proximity, and try to separate it from other connections, such as for example personal ties, with a family member working in Russia and transmitting remittances. We thus hypothesise that Russian-speakers are particularly susceptible to information that is conveyed in the Russian language, while they are also more likely to seek information from

[4]https://theins.ru/politika/258094 See also Sablina (2023), who shows the role social media played in mobilizing the Russian-language community in Germany.

[5]https://theins.ru/politika/256770

Russian-language sources in the case of high-profile events.

The second part of our paper builds on Korovkin and Makarin (2023), who find that Ukrainian firms from districts with fewer ethnic Russians experienced a deeper decline in trade with Russia, after the Russian annexation of Crimea in 2014. Along a similar vein, we document that firms in regions with a higher share of Russian speakers are more likely to circumvent the sanctions and export war-relevant goods to Russia, after the imposition of trade sanctions following Russia's full-scale invasion of Ukraine.

The remainder of our paper is organized as follows. Section 4.2 introduces the background and context of our study, and section 4.3 presents our data and provides summary statistics. Section 4.4 documents how Russian speakers increased their consumption of Russian media after February 2022, while simultaneously their views of Russia improved and their views of the West deteriorated, as compared to the rest of the population in Kazakhstan. Section 4.5 uses our original survey to show how Russian media consumption is indeed related to being more likely to condone Russia's actions in the war and to support the circumvention of economic sanctions. Section 4.6 connects these findings to our trade data, and documents how Kazakh regions with a higher share of Russian speakers are significantly more likely to export war-related goods to Russia after February 2022, but not before. Section 4.7 concludes.

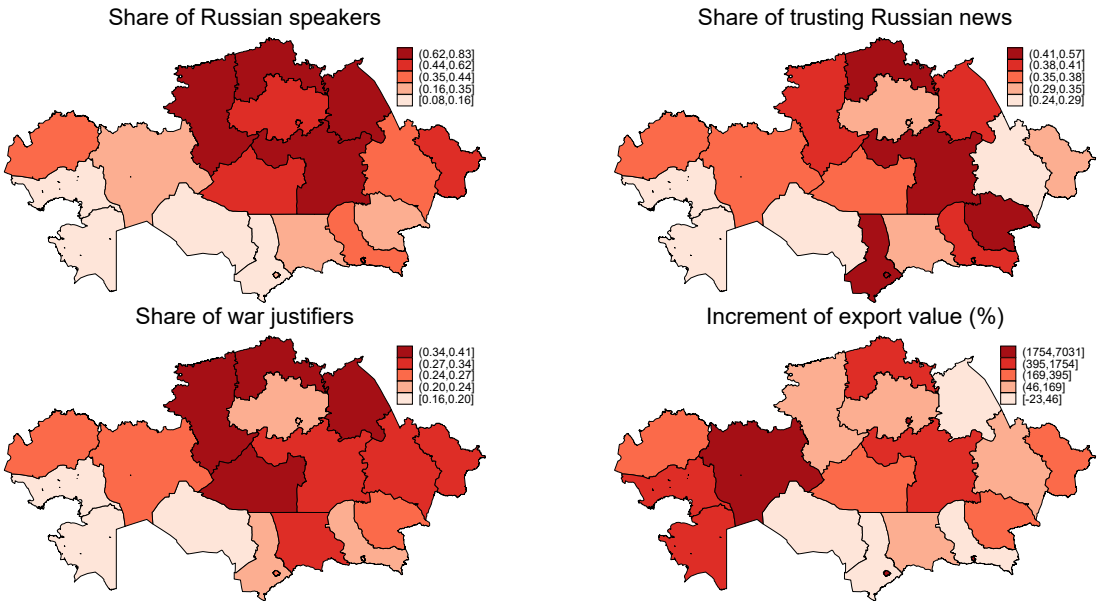
4.2 Background and Context

One of the successor states of the Golden Horde, the Kazakh Khanate was gradually conquered and absorbed by the Russian Empire from the late 18th to the mid-19th century. Following the Russian Revolution, the territory became part of the Soviet Union. After several territorial reorganizations, Kazakhstan's modern borders were established in 1936 with the formation of the Kazakh Soviet Socialist Republic. Kazakhstan declared its independence from the Soviet Union on December 16, 1991, and since then has been an independent state.

As a result of two centuries of Russian colonization, a large Russian minority continues to live in Kazakhstan. Although Kazakh is the official state language, Russian also has the status of an official language, and a large share of the population continues to speak Russian

in their daily lives. Map 4.1 illustrates the distribution of Russian speakers across Kazakh regions, showing how Russian speakers are concentrated in the north of the country, as well as in the two main cities Astana and Almaty. As we will show throughout the paper, this regional variation in language speaking capabilities can be linked to differences in media consumption (section 4.4), political opinions (section 4.5), and economic behaviour in the form of trading sanctioned products to Russia (section 4.6).

Figure 4.1: Regional heterogeneity by language, Russian media trust, war justification and sanctions circumvention



Importantly for our study, while the share of ethnic Russians in 2024 was 14.9% and that of ethnic Kazakhs 71%, 84.8% of the Kazakh population are able to read and write in Russian, and between 40% and 50% indicate Russian as the language they speak at home.[6]

Since the start of Russia's war in Ukraine in 2014, the Kazakh government has been careful in maintaining a low profile with respect to the conflict, avoiding too extensive coverage of the conflict in the official media (Lehtisaari et al. 2018) and employing "strategic silence" Dadabaev and Sonoda (2023), in order not to antagonize neither Russia nor the West, even

[6]https://stat.gov.kz/en/ See Table 4.1 in Section 4.3 for the share of Kazakh citizens that speak Russian at home.

though at times concerns and criticism about the war are voiced.[7] As a result, almost all of the Kazakh- and Russian-language media published within Kazakhstan take either a neutral or a moderately critical stance towards Russia when covering the conflict.[8] However, at the same time Russian-language media from Russia is widely available in Kazakhstan. Especially when searching for information on the internet, the probability that Russian speakers will end up on a site from Russia featuring pro-Russian content is therefore relatively high, while Kazakh speakers are more likely to find content that is more neutral or critical with respect to Russia's invasion of Ukraine.

While the Kazakhstani government has been careful to maintain a neutral stance towards the conflict in its official statements, economically, Kazakhstan has taken a pro-Russian stance. Exports from Kazakhstan to Russia have increased significantly since February 2022. However, as illustrated by Figure 4.2, this is *only* the case only for electronic goods, machines, and weaponry, i.e. products on which sanctions were imposed by Western nations after Russia's full-scale invasion of Ukraine in February 2022, while exports of other goods have remained remarkably stable.

In our paper, we investigate if Kazakhstan's role in circumventing the sanctions is related to Russia's continuing influence on the Russian-speaking community in Kazakhstan. In other words, are Russian speakers more likely to condone the war and approve of the circumvention of the sanctions after encountering Russian propaganda? And if there is such an effect, does it translate into increased exports of sanctioned products to Russia?

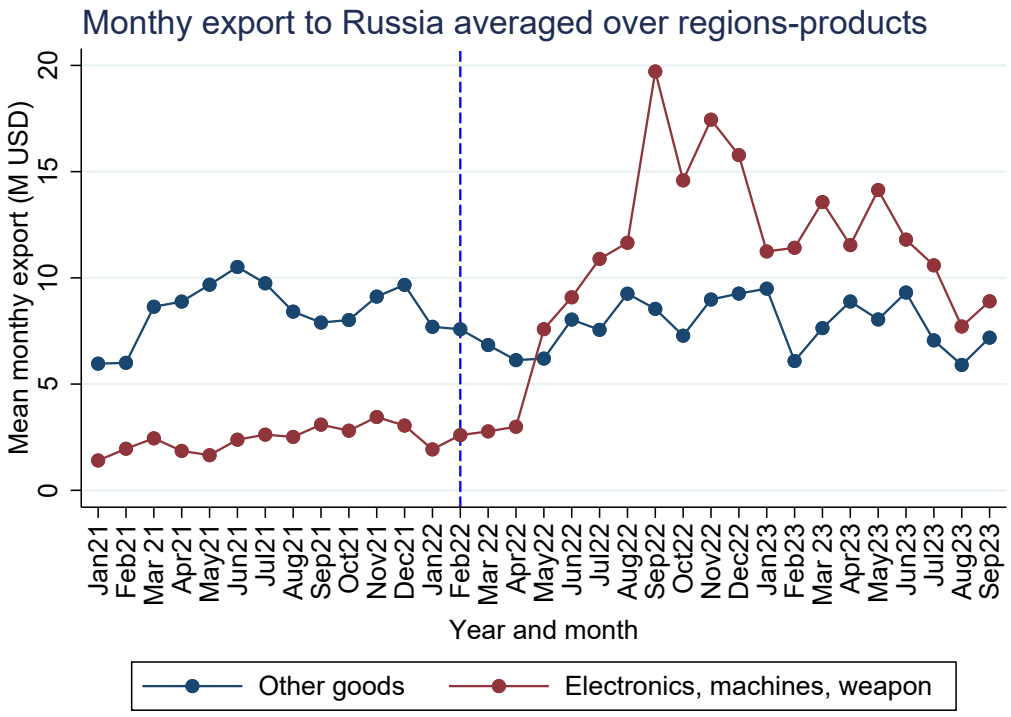
4.3 Data

To answer these questions, we use four distinct sources of data. First, six waves of a telephone survey conducted by the Central Asia Barometer of the population of Kazakhstan for the time period between 2021 and 2023 to test for a potential divergence in political views and media consumption between Russian speakers and other Kazakh citizens after the start of

[7] https://ecfr.eu/publication/steppe-change-how-russias-war-on-ukraine-is-reshaping-kazakhstan/

[8] This is also a result of Kazakh media being largely state-controlled. According to Freedom House, "the dominant media outlets in Kazakhstan are either in state hands or owned by government-friendly businessmen", https://freedomhouse.org/country/kazakhstan/freedom-world/2023.

Figure 4.2: Export from Kazakhstan to Russia



the full-scale invasion. Second, we justify this divergence by analyzing Yandex online search for conflict information and media content. Third, we conducted a large-scale online survey with list experiments in October 2023, permitting us to reveal attitudes toward the war and sanctions circumvention and study their discrepancy by media consumption. Fourth, transaction-level trade data of Kazakhstan within the Eurasian Economic Union (EAEU) for 2021 to 2023, permitting us to investigate the export of sanctioned goods from Kazakhstan to Russia before and after February 2022, for Kazakh regions with varying shares of Russian-speakers.

4.3.1 Central Asia Barometer

Since June 2017, the Central Asia Barometer (CAB) has been conducting regular and detailed public opinion surveys in the four Central Asian republics of Kazakhstan, Kyrgyzstan, Uzbekistan and Tajikistan.[9] For our study, we use waves 9 to 14 of CAB data for Kazakhstan

[9]https://www.ca-barometer.org/en

carried out in the Spring and Autumn of 2021 – 2023. In addition to detailed demographics, all six survey waves include a wealth of information on questions of public interest and individual attitudes, including attitudes towards Russia, the United States, China, and the Eurasian Economic Union. The waves starting from June 2022 also include several questions related to Russia's full-scale invasion of Ukraine, of which we use in particular the following two: (i) *"In your view, who is mainly responsible for the situation in Ukraine?"* (pre-coded answers: *"Russia", "Ukraine", "The United States", "NATO, Europe/the EU"*), (ii) *"To what extent do you think Russia's special military operation in Ukraine is justified or unjustified?"* (pre-coded answers: *"Completely justified", "Somewhat justified", "Somewhat unjustified", "Completely unjustified", "Refused to answer", "Don't know"*).

To measure attitudes towards Russia, the United States, China, and Iran, the survey asked the following question: *"Thinking about other countries, please tell me if you have a very favorable, somewhat favorable, somewhat unfavorable, or very unfavorable opinion of (Russia/the United States/China/Iran)."* Attitudes towards the Eurasian Economic Union were gauged through the following question: *"As you may know, since 2015, our country has been a member of the Eurasian Economic Union with Russia, Belarus, and Armenia. Do you strongly agree, somewhat agree, somewhat disagree, or strongly disagree that joining the Eurasian Economic Union has benefited our country's national economy?"*

All six survey waves also include an extensive block of questions related to: (i) the *main* source of news most often used to get information about what is going on outside of Kazakhstan (including relatives and friends, national and Russian traditional media, and the internet), (ii) the messaging apps most often used (WhatsApp, Telegram, etc.), (iii) the social media platform most often used (Facebook, Odnoklassniki, VKontakte, Instagram, Twitter, TikTok, etc.), (iv) the language that is spoken at home and the language of the interview, (v) and a full set of standard demographics, including ethnicity. Finally, the wave of September 2023 also includes the following question: *"Would you say that you strongly trust, somewhat trust, somewhat distrust, or strongly distrust news and information from (Russia/the United States/China/Turkey/Iran/)?"*

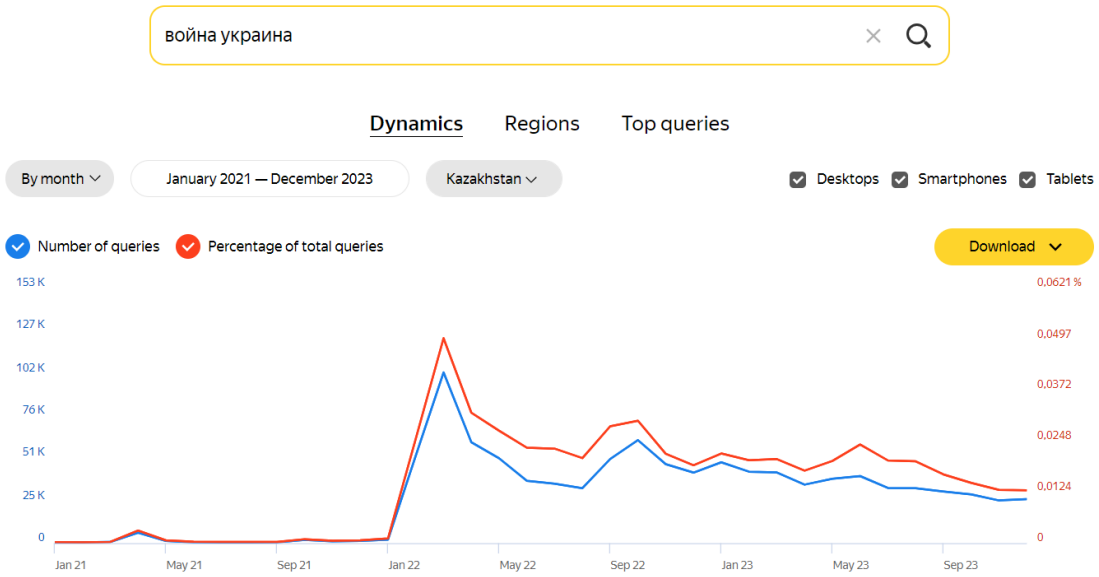
Section 4.4.1 provides descriptive statistics of the survey data, section 4.4.3 investigates how language proficiency and opinions about the war are related, section 4.4.4 looks at the

mechanisms behind these findings, and section 4.4.5 looks at dynamics over time.

4.3.2 Yandex Search Data

We use Yandex search data to see how information-seeking behavior has evolved before and during the invasion of Ukraine.[10] The data is taken from yandex.kz and available on a monthly-regional level. Figure 4.3 shows a substantial spike in searches of the phrase "war Ukraine" in February 2022, with a smaller spike around the Russian mobilization in September 2022 and a stabilisation after that.

Figure 4.3: Search for "War Ukraine" on Yandex



Figures A5, A6 and A7 of Appendix A document that the interest in internet news also had a spike in February 2022. In addition, we see that while the search for information from pro-Kremlin news aggregators Yandex Dzen and Yandex News (Figures A5 and A6 of Appendix A) is significant and stable after February 2022, the increase in searches for Euronews, the most popular pro-European news channel in Kazakhstan, is only a fraction (Figure A7 of Appendix A).

[10] Yandex has a market share of roughly 20% in Kazakhstan, whereas Google carries out a vast majority of searches; however, Google search data is not accessible in a systematic way. The market share of Yandex has been oscillating between roughly 10 and 24% for the last three years https://gs.statcounter.com/search-engine-market-share/all/kazakhstan

4.3.3 Survey List Experiment

The third dataset we use is the result of an online survey list experiment conducted in November 2023 in Kazakhstan with 5,025 respondents. The survey was designed to reveal unfalsified preferences of Kazakhstani people about their attitude to the war and readiness to help Russia circumvent sanctions. The survey was pre-registered[11] and carried out by NAC Analytica, a leading Kazakh sociological and public opinion research organization.[12] Participants were recruited through advertisements in social media (river sampling), and a weighting-scheme was applied to make the sample nationally representative.

After having answered questions about their socio-demographic background and their opinions on the lists (see section 4.C.2), participants were asked three questions in random order, two of which are politically sensitive, and one that is not sensitive and acts as a placebo in order to confirm the list experiment worked as intended. The sensitive questions concerned geopolitics (*In your opinion, is helping Russia avoid Western sanctions generally justified or not justified?* and *In your opinion, is Russia's Special Military Operation/ invasion of Ukraine generally justified or not justified?*). The framing *Special Military Operation* and *invasion of Ukraine* was assigned at random, in order to balance invoked framing effects. Arguably, the way one describes Russia's invasion of Ukraine gives away their view on this war and thus invokes demand effects and social desirability bias. A neutral stance between the two mutually exclusive narratives of an illegitimate invasion or a 'Special Military Operation' is hard to find. We controlled for the effect of the framing in our regression analysis. While there are differences in the proportions, the dynamics of self-censorship develop analogously across treatments. The third question acted as a placebo, in order to control for design effects (*In your opinion, is working more than 50 hours per week generally justified or not justified?*). Answer options for these questions were *Justified*, *Not justified*, and *Prefer not to answer*. We coded *Justified* as 1, and *Not justified* and *Prefer not to answer* as 0.

Quality controls included attention checks (two questions on respondents age had to match), speeding filters (minimum of 200 seconds), allowing only two completes per IP

[11] https://aspredicted.org/YPM_LSH

[12] https://nacanalytica.com/en/

address, and allowing phone numbers to participate only once (payment was carried out by phone number). Out of 28,201 participants, 5,025 completed the survey, passed quality checks, were unique respondents, and were compensated 700 Tenge (approx. 1.50 USD). A large majority of the participants that left the survey before finishing did so in the very first pages of the survey.

4.3.4 Trade data

Finally, we use the international trade data between Kazakhstan and Russia to study the economic consequences of the Kazakhstan population's exposure to Russian media. The data was provided by the Statistical Bureau of Kazakhstan (www.stat.gov.kz). It contains all transactions between EAEU countries and Kazakhstan for January 2021 — September 2023, including the region of Kazakhstan's firms, the trading product's classification code called TNVED, and the trading partner's country. We consider only export activity from Kazakhstan to Russia, reaching more than 810K transactions, and classify all the products into five categories: (i) electronics, machines and weapons, (ii) food, (iii) cloth, (iv) metals, and (v) others.[13] Most of the products from the first class are under Western sanctions, forbidding them from exporting to Russia, as they potentially can be used for the war, so we call them *Sanctioned products* even when we discuss the pre-war period. Finally, we aggregate the export data, calculate the export value in a million USD at the region(20 regions)-month(33 months)-product class(5 classes) level, and get 2899 observations as the dataset for analysis.

4.4 Language Proficiency, Media and Public opinion

4.4.1 Descriptive Statistics

Columns 3 and 4 of Table 4.1 provide descriptive statistics on language and ethnicity. The share of speaking Russian at home in our sample is 47.3% for 2021, 41.1% for 2022, and 45.3% for 2023. For all three years, the proportion of respondents who consider Russian their home language is more than twice the share of ethnic Russians. This distinction allows

Table 4.1: Descriptive statistics for surveys

		Russian (%)		War responsibility (%)			SMO is	Favorable opinion of (%)				EAEU is
Year	Obs.	Language	Ethnicity	Russia	Ukraine	West	justified (%)	Russia	US	China	Iran	beneficial (%)
2021	3,500	47.3	22.2					70.5	51.9	44.9	21.4	57.9
2022	3,020	41.1	17.9	27.5	20.8	10.8	26.3	50.9	49.4	51.5	27.0	59.5
2023	3,000	45.3	20.5	24.9	19.4	12.1	25.0	57.1	51.4	60.3	24.7	54.1

Note. The wave of May 2021 includes 2000 observations, and other waves include around 1500 observations each. All the percentages are calculated as shares in the waves of the corresponding year. In the question "SMO is justified" 1 means "Completely justified" or "Somewhat justified", and 0 means "Completely unjustified", "Somewhat unjustified" or "Refused", "Don't Know". War responsibility is denoted "West" if a respondent has chosen US, EU, or NATO. In the question "Favorable opinion of Russia/US/China/Iran" 1 means "Very favorable" or "Somewhat favorable", and 0 means "Very unfavorable", "Somewhat unfavorable" or "Refused", "Don't Know". In the question "EAEU is beneficial" 1 means "Strongly agree" or "Somewhat agree" that joining the Eurasian Economic Union has benefited Kazakhstan's national economy, and 0 means "Strongly disagree", "Somewhat disagree" or "Refused", "Don't Know".

us to separate the effects of ethnicity from those of language use.

Regarding the perception of responsibility for the conflict in Ukraine, around a quarter of respondents primarily associate it with Russia, around 20% believe that Ukraine is mainly responsible, and less than 12% of respondents think of Western countries, including the US, EU, and NATO. Moreover, three-quarters of the population do not believe that Russia's Special Military Operation (SMO) in Ukraine is justified.

More than 70% of respondents had a favorable opinion of Russia in 2021, but this share significantly dropped to about 51% in 2022 and moved back to 57% in 2023. On average, around half of the population held favorable opinions of the US and China, with almost no dynamics for the former and an improved attitude for the latter. Less than a quarter viewed Iran positively, with some improvement in dynamics. Most respondents agree that joining the Eurasian Economic Union has benefited Kazakhstan's national economy, and the attitudes are quite stable.

Table 4.2 illustrates various channels through which Russian propaganda can potentially affect respondents' beliefs. Notably, the Internet emerged as the primary source of news regarding events outside of Kazakhstan, with approximately two-thirds of the population utilizing it by 2023. Slightly more than 10% of the sample access international news primarily via local or national TV, radio, newspapers. Despite the reach of Russian TV, radio,

[13]The following 2-digits of TNVED product classification correspond to these products: electronics (85,88,90,91), machines (84,86, 87,89), weapon (93), food (01-24), cloth (41-43, 50-67), metals (72-83)

Table 4.2: Descriptive statistics for channels

	2021	2022	2023
News Internet	59.4	64.6	67
News National or Local TV, radio, newspapers	10.6	12.2	10.1
News Russian TV, radio, newspapers	7.2	5	4.4
Trust to News from Russia			36.8
Use Russia media		61.9	60.5
Use VK or OK	16.2	11.6	9.2
Use Telegram	6	8.5	10.1
Use Facebook	8.7	8.2	7.7
Use Instagram	50.8	48.8	47.4
Use Tiktok	7.7	17.1	22.5
Use Whatsapp	88.3	87.5	86.7
Absence of remittance	88.3	91.9	92.9

Note. Table shows the percentage of media usage. The row "Trust to news from Russia" shows the share respondents from September 2023 who strongly or somewhat trust news or information from Russia. The row "Absence of remittance" shows the share of the households that do not rely on remittance.

newspapers as *the main source of news* being quite marginal (4-7%), around 61% of the population consumes Russian content for news and entertainment, and nearly 37% of respondents reported trust in Russian news.

In terms of social media and messengers, the data presented in Table 4.2 highlights WhatsApp and Instagram as having the highest consumption rates. The usage of Odnoklassniki or Vkontakte has substantially dropped during 2021-2023, while platforms like TikTok and Telegram experienced increased popularity. Lastly, there was a slight decrease in the proportion of households relying on remittances after the start of the war.

Figure 4.4 shows the dynamic trends in public opinion on the justification of the Special Military Operation (SMO) and the attribution of responsibility for the situation in Ukraine, differentiated by language: Russian and non-Russian speakers. The period covers four survey waves: June 2022 (wave 11), November 2022 (wave 12), April 2023 (wave 13), and September 2023 (wave 14). Russian speakers are more likely to justify the Special Military Operation, and associate the responsibility for the situation in Ukraine with the US, EU, or NATO

countries, in comparison with respondents speaking other languages.

As shown in the left panel of Figure 4.4, the share of Russian speakers justifying the SMO declines from about 50% in June 2022 to around 35%-40% in consecutive waves. Importantly, non-Russian speakers consistently show lower justification levels, around 15% on average. Among non-Russian speakers, approximately 34% of respondents attribute responsibility for the conflict to Russia and 6% – to Western countries. Among Russian speaking, these figures are around 17% and 19%, respectively.

Figure 4.4: SMO justification and Responsibility for the situation in Ukraine in dynamics

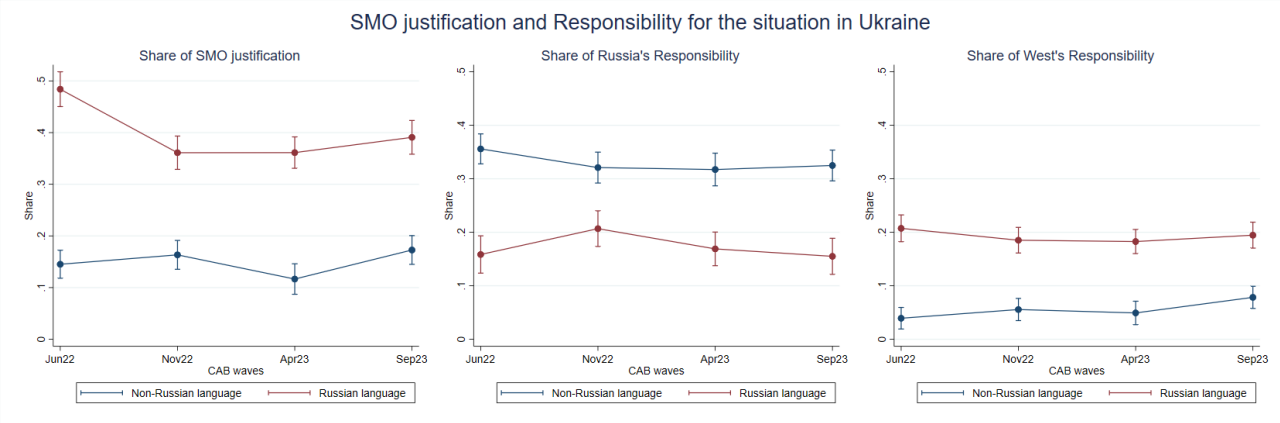
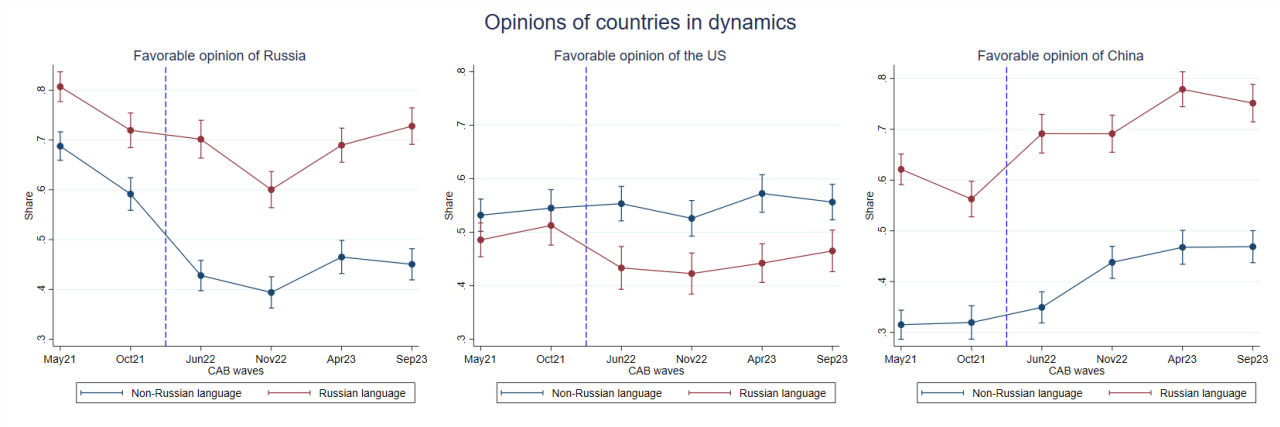


Figure 4.5 presents the dynamic trends in public opinion regarding Russian, the US, and China among Russian and non-Russian language speakers (Figure A1 shows the graphs for Iran, Turkey, and EAEU). The Russian-speaking respondents generally hold more favorable opinions of Russia and China and less favorable views of the US.

Favorable opinions of Russia declined over time among both Russian and non-Russian speakers. However, the decrease was more pronounced among respondents who did not speak Russian at home. For instance, in May 2021, the gap between the shares of favorable opinions for Russian and non-Russian speakers was 12% (with 81% of Russian speakers holding a favorable opinion of Russia compared to 69% among non-Russian speakers). By September 2023, this gap had widened significantly to 28% (with 73% of Russian speakers expressing a favorable view of Russia compared to only 45% among non-Russian speakers). This widening gap suggests a diverging trend in attitudes towards Russia based on linguistic identity over the surveyed period.

Figure 4.5: Favorable opinions of different countries



Regarding attitudes towards the United States, non-Russian speakers are generally more inclined to hold a favorable opinion of the US compared to Russian speakers. Specifically, non-Russian speakers exhibit a favorable view of the US at a rate of 55% on average during 2021-2023, whereas Russian speakers express slightly lower favorability, standing at 46%. However, there's a notable difference between Russian and non-Russian speakers in their post-war views. The share of favorable opinion of the US for Russian speakers dropped since the start of the war but has been recovering over time. On the other hand, favorability towards the US among non-Russian speakers has slightly increased since the war started.

Between May 2021 and September 2023, favorable views of China showed a gradual increase among both linguistic groups. Russian speakers consistently maintained a higher level of optimism about China compared to respondents with a different language identity. Throughout the entire period, the average share of favorable opinions of China among Russian speakers was 68%, whereas for non-Russian speakers it was notably lower at 39%. Despite this substantial gap in attitudes towards China based on linguistic identity, this gap has been persistent over time and has not been affected by the war.

4.4.2 Yandex online search

Tables 4.5 show that Russian-speaking respondents consume internet news more frequently. However, the CAB surveys do not provide information about the specific internet sites they visit. It is unclear whether the increased quantity of news consumption for Russian speakers is accompanied by a difference in news sources. We use online search statistics from Yandex

to fill this gap. Using the CAB data, we calculate the share of Russian-speaking respondents at the regional level SL_r and match this data with the regional search statistics. To see whether regions with a higher share of Russian-speaking are different from those with fewer Russian speakers in terms of the search for the war- and news-specific keywords we consider the following event-study design:

$$y_{rt} = \sum_t \beta_t \cdot SL_r \cdot \mu_t + \lambda_r + \mu_t + \epsilon_{lt}. \tag{4.1}$$

In equation 4.1, our dependent variable $y_{rt} = \frac{\text{Num. of searches in region } r \text{ in month } t}{\text{regional population in } r} \cdot 1000$, i.e. it is the number of the searcher of a keyword in a region per 1000 residents of the region. SL_r is the share of Russian speaking respondents in the region r (using CAB waves). Finally, λ_r is the region fixed effect and μ_t is year-month fixed effect. Figure 4.6 shows the results for keyword "war Ukraine". Regions with more Russian speakers experience an increase in searches compared to regions with fewer Russian speakers. Moreover, in March 2022, the increase of the share of Russian speakers from 8% to 83% in a region (i.e. the lowest and the largest shares by Kazakhstan regions) increases the search of war-related keywords by 12.3 times per 1000 of people.

The result for the keyword "war Ukraine" does not reflect any attitude to this war; it just indicates the interest. In what follows, we apply 4.1 for searches Yandex Dzen and Yandex News – the largest pro-Kremlin news aggregator since 2020, which finally was acquired by Russian state-owned social media VK [14] in 2022. Figure 4.7 shows that the effect of Russian language penetration on the search of these two sources in March 2022 is comparable to a general search "war Ukraine" and the dynamic patterns of search are also comparable, suggesting that the Yandex Dzen and Yandex News could be the most important news sources in Russian speaking regions.

Figure A8 shows the results for the most popular pro-Western news source in Kazakhstan – "Euronews". While the pattern dynamic patterns of the search are also comparable to the previous two, the magnitudes are 10 times smaller compared to the "Yandex News" or

[14] https://www.forbes.ru/tekhnologii/476829-vk-i-andeks-ob-avili-o-zakrytii-sdelki-po-dzenu-i-novostam

Figure 4.6: Event study design about the search for "war Ukraine"

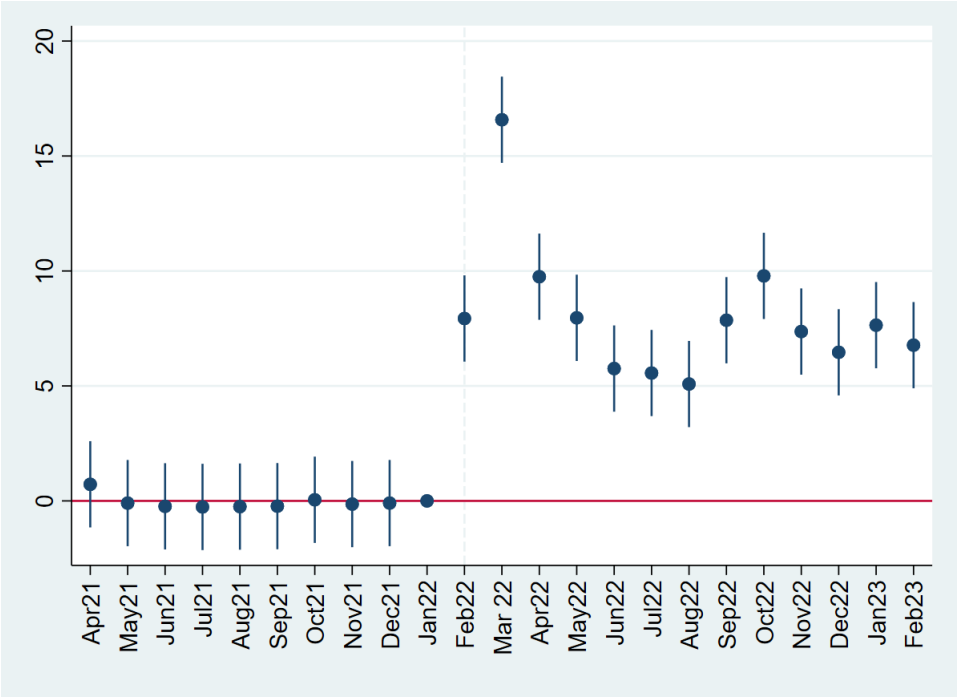
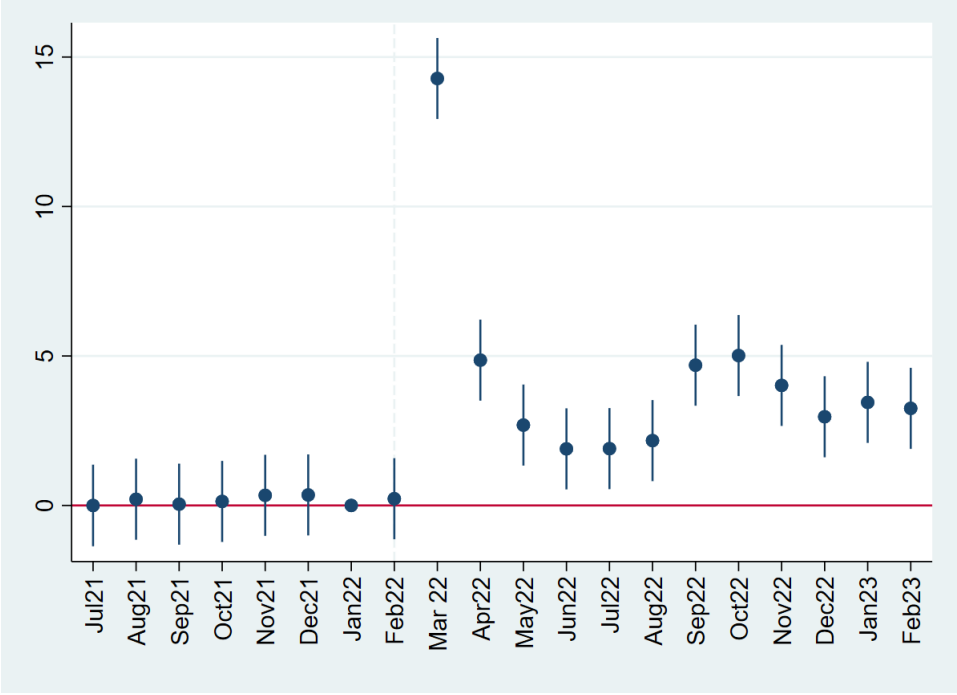


Figure 4.7: Event study design about the search for "Yandex News" or "Yandex Dzen"



"Yandex Dzen" search.

4.4.3 Language Proficiency and Justification of the War in Ukraine

As the next step, we study how linguistic proficiency in speaking Russian affects attitudes toward the war in Ukraine. We use CAB survey waves conducted since 2022, and consider the following TWFE linear equation:

$$y_{irt} = \alpha L_{irt} + \mathbf{X_{irt}}\beta + \kappa_r + \lambda_t + \epsilon_{irt}. \tag{4.2}$$

Here, i denotes respondent from region r and wave t.[15] The binary variable L_{ilt} is 1 if a respondent speaks Russian most often at home, and otherwise 0. The vector $\mathbf{X_{irt}}$ includes the set of control variables: gender, five age categories, university education (or higher), five categories of household size, urban/rural settlement, and ethnicity as a set of dummy variables (including Kazakh and Russian ethnicity), the interviewer assessment of the respondent's survey comprehension (four levels) and comfort (four levels). Noteworthy, the control for ethnicity enables us to disentangle the language effect from the ethnicity. The variables κ_r and λ_t are the region and survey wave fixed effects, respectively. As dependent variables y_{ir}, we consider answers to the following three questions on the opinion of the war: (i) the extent to which SMO is justified (binary and scaled), (ii) the respondent's statement that Russia is primarily responsible for the war, and (iii) the respondent's statement that western countries are mostly responsible for the war. Our main focus is coefficient α, which shows the association between likely exogenous language habits and war attitudes, clean of socio-demographic and ethnic components.

Table 4.3 shows the results for binary measure "SMO is justified" for the whole sample (column 1) and different sub-samples we have in CAB data: firms heads (column 2), younger (below 40) and elder (above 40) people (columns 3-4), respondents who pay a lot of attention to the information about the war (column 5), and ethnically non-Russians (column 6). The Russian-speaking respondents demonstrate high justification for the war by 0.11 pp for the total sample (Column 1), which is equivalent to 42.8% when referring to the sample average. For all the above-mentioned sub-samples, the results are similar in the direction of being

[15] Regions are three major cities, Astana, Almaty, and Shymkent, and 17 other sub-national regions.

stronger for older generations (0.16 pp) and weaker for younger (0.077 pp). Table B1 of Appendix shows the estimates for the scale variable "SMO is justified", and the results are the same, except that the coefficient for firm heads is insignificant.

Table 4.3: Russian language proficiency and war justification

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	SMO is justified (binary)					
Russian language	0.11***	0.090**	0.16***	0.077***	0.13***	0.10***
	(0.015)	(0.037)	(0.026)	(0.019)	(0.019)	(0.015)
Observations	6,020	955	2,366	3,654	3,889	4,864
R-squared	0.140	0.154	0.223	0.082	0.178	0.078
Waves	After 2022	After 2022	After 2022	After 2022	After 2022	After 2022
Sample	All	Firm head	Above 40	Below 40	Informed	No Russians
Region FE	Yes	Yes	Yes	Yes	Yes	Yes

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. The table shows the regression results (4.2) for a binary dependent variable, "SMO is justified". Robust standard errors are in parentheses.

As the next step, we consider a set of binary variables assigning responsibility for the war to three possible actors – Russia, Ukraine, or the West as the dependent variable y_{ilt} in equation (4.2). Table 4.4 shows the results with a breakdown by dependent variables and sub-groups of people. Column 1 of Panel A shows that overall, Russian-speaking respondents are less likely to blame Russia as the mainly responsible side of the war by 0.057 pp, which is equivalent to 21.7% when referring to the sample average. The effect is even stronger for the respondents who pay attention to the information about the war (Column 5 Panel A). Nevertheless, for firm heads and younger respondents, the association is insignificant (Columns 2, 4 of Panel A). Panel B shows an insignificant language gap in blaming Ukraine as the mainly responsible side, suggesting that neither of the sub-groups of respondents among Russian-speaking blame Ukraine more than non-Russian-speaking. Finally, Column 1 of Panel C shows overall, the Russian-speaking respondents blame the Western countries by 0.087 pp more than non-Russian-speaking, which is equivalent to 76% when referring to the sample average. Moreover, this effect is also strong for all other considered sub-groups of respondents, including firm heads (0.079 pp in Column 2 of Panel C) and young people

Table 4.4: Responsibility for the situation in Ukraine

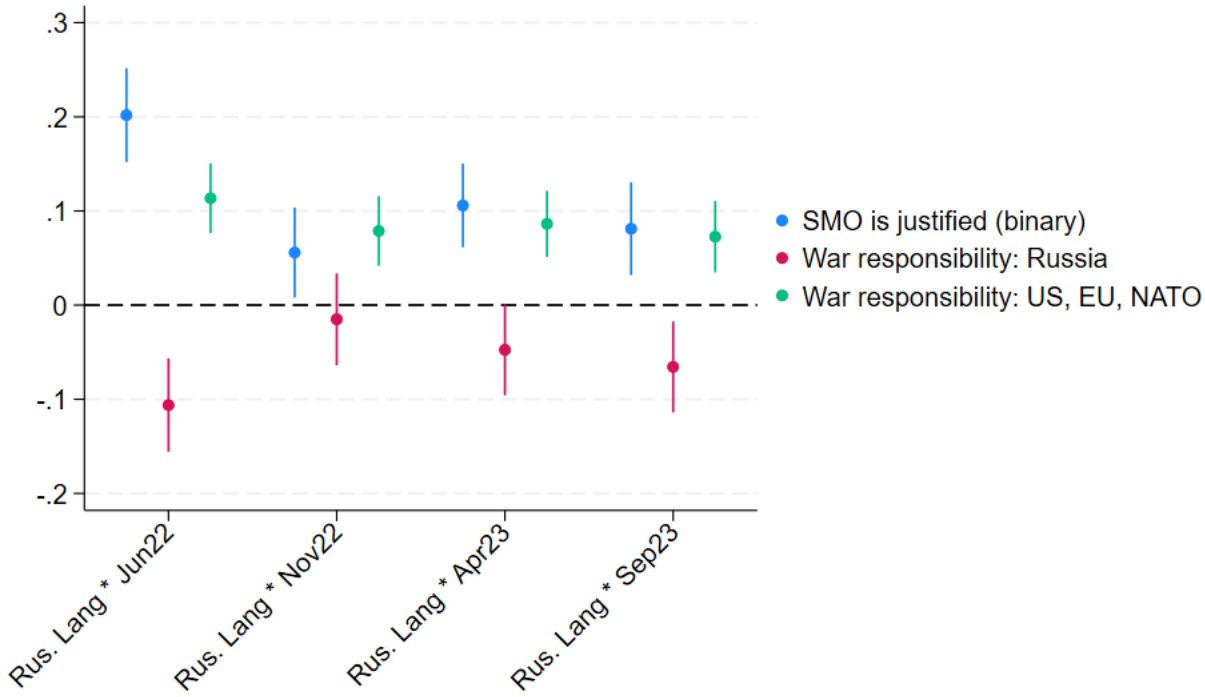
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. War responsibility: Russia					
Russian language	-0.057***	-0.025	-0.094***	-0.034	-0.070***	-0.058***
	(0.017)	(0.042)	(0.028)	(0.021)	(0.022)	(0.017)
	Panel B. War responsibility: Ukraine					
Russian language	0.013	-0.028	0.028	0.0058	0.0012	0.013
	(0.014)	(0.033)	(0.024)	(0.018)	(0.018)	(0.014)
	Panel C. War responsibility: US, EU, NATO					
Russian language	0.087***	0.079***	0.13***	0.063***	0.10***	0.083***
	(0.012)	(0.031)	(0.024)	(0.013)	(0.016)	(0.012)
Observations	6,020	955	2,366	3,654	3,889	4,864
Waves	After 2022	After 2022	After 2022	After 2022	After 2022	After 2022
Sample	All	Firm head	Above 40	Below 40	Informed	No Russians
Region FE	Yes	Yes	Yes	Yes	Yes	Yes

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. The table shows the results of regression (4.2) for binary dependent variables emphasizing the most responsible side for the war: Russia/Ukraine/Western countries. Robust standard errors are in parentheses.

(0.063 pp in Column 4 of Panel C).

Figure 4.8 shows the coefficient α from (4.2) in dynamics suggesting separate coefficient L_{irt} for each wave and including all the respondents. War justification has some dynamics, such as a stronger association between language and justification in the first wave and stabilization after that. A similar pattern is for another variable- Russia and the West's main responsibility for the war. All in all, we can conclude that Russian-speaking people justify the war more; they are less likely to blame Russia and more likely to blame Western countries for the war compared to non-Russian speakers, outlining the importance of language use and not only ethnicity in shaping political opinions.

Figure 4.8: The Language effect on SMO justification and responsibility for the situation in Ukraine in dynamics



4.4.4 Mechanisms on how does Language Proficiency affect Opinions.

In this section, we study a set of potential mechanisms on how Russian language use can result in justification of the war in Ukraine, a result that we found even for the ethnically non-Russian population.

We start by considering information channels that have traditionally played an important role in Central Asia countries - Russian traditional media (TV and Radio) and labor migration. We use them as dependent variables in our equation (4.2).

Table 4.5: Channels for Russian propaganda

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Panel A. News from Russian TV, radio, newspapers					
Russian language	0.022***	0.0074	0.037***	0.010*	0.029***	0.016***
	(0.0063)	(0.012)	(0.014)	(0.0060)	(0.0089)	(0.0061)
	Panel B. Absence of remittance					
Russian language	0.0082	0.024	0.0033	0.010	0.0032	0.013
	(0.010)	(0.025)	(0.016)	(0.014)	(0.012)	(0.011)
	Panel C. News from the Internet					
Russian language	0.038**	-0.017	0.0063	0.052***	0.046**	0.038**
	(0.016)	(0.036)	(0.030)	(0.020)	(0.019)	(0.017)
Observations	6,020	955	2,366	3,654	3,889	4,864
Waves	After 2022	After 2022	After 2022	After 2022	After 2022	After 2022
Sample	All	Firm head	Above 40	Below 40	Informed	No Russians
Region FE	Yes	Yes	Yes	Yes	Yes	Yes

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. The table shows the results for binary dependent variables regarding channels of potential Russian propaganda. Robust standard errors in parentheses.

Panel A of Table 4.5 shows that Russian-speaking respondents are more likely than non-Russian speakers to consume Russian traditional media (TV news, radio, and newspapers). This effect is specifically stronger for the older population and those who pay attention to the information about the war. There is no evidence that remittance may play any role as the channel (Panel B).

As the second step, we examine the news on the internet. Panel C of Table 4.5 shows that Russian-speaking respondents consume internet news by 0.038 pp more than non-Russian-speaking respondents, and this effect is especially prominent for the younger population. Panel A and B results suggest that Russian propaganda targets older generations via traditional media and younger generations via the Internet. Nevertheless, CAB question about international news covers only the internet per se, while it does not specify the exact source, such as websites or social media. VK and Odnoklassniki (OK) are social media owned by the Russian government. The former is more popular among younger people, while the latter is used mainly among older people. Among the above-named channels, VK, OK, and Telegram have been intensely used by Russian government propaganda, especially since February 2022. Telegram, in particular, has become a platform where many pro-Russian bloggers with often large audiences are providing information about the war in Ukraine from a Russian perspective.[16] Facebook and Instagram, on the other hand, are blocked in Russia. TikTok hosts entertainment content, and WhatsApp is mainly considered a messaging app, so these platforms are less relevant as channels of Russian propaganda. To understand the role of social media, we use the question of the most frequently used social media interacted with internet news consumption. We use them as dependent variables in our equation (4.2). Table 4.6 shows the results.

Columns 1-2 of Table 4.6 show that Russian-speaking respondents are significantly more likely to use internet news with VK-OK or Telegram than non-Russian-speaking respondents. However, the Russian-speaking respondents were less likely to use internet news with Instagram, TikTok, Facebook or WhatsApp.

4.4.5 Dynamics of Opinion Change

Apart from being influenced through social media and the internet, another potential channel that could explain our results from Tables 4.3 and 4.4 is that Russian-speaking populations in Kazakhstan have always been more pro-Russia than non-Russian speaking populations,

[16] See for example https://www.dw.com/en/russian-war-bloggers-pawns-in-a-political-game/a-64284496, or https://www.themoscowtimes.com/2022/09/14/explainer-who-are-russias-pro-war-bloggers-and-why-are-they-important-a78793

Table 4.6: Social media as internet news

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Internet news and the following Social Media:					
VARIABLES	VK or OK	Telegram	Instagram	TikTok	Facebook	WhatsApp
Russian language	0.055***	0.028***	-0.026**	-0.027***	-0.040***	-0.045***
	(0.0063)	(0.0068)	(0.012)	(0.0085)	(0.0058)	(0.013)
Observations	6,020	6,020	6,020	6,020	6,020	6,020
R-squared	0.014	0.003	0.001	0.002	0.007	0.002
Waves	After 2022	After 2022	After 2022	After 2022	After 2022	After 2022
Sample	All	All	All	All	All	All

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. The table shows the results for binary dependent variables regarding channels of potential Russian propaganda. Robust standard errors in parentheses.

as they have been sharing long-term historical and cultural ties and views with the Russian population in Russia. If this is the case, we would expect that this group always held a more favorable view of Russia and a less favorable view of the US than the group of non-Russian speakers.

To examine this argument and the dynamic evolution of opinions on Russia, the US, and China, we use the two survey waves from 2021 in combination with the survey waves after 2022 and estimate the following DID specification from Korovkin and Makarin (2023):

$$y_{irt} = \alpha L_{irt} + \gamma Post_t + \delta L_{irt} \cdot Post_t + \mathbf{X_{irt}}\beta + \kappa_r + \lambda_t + \epsilon_{irt}. \tag{4.3}$$

Here i denotes the respondent from the region r and survey wave t. The binary variable L_{irt} equals one for Russian-speaking, and the binary variable $Post_t$ equal one for the period of war in Ukraine. The vector $\mathbf{X_{irt}}$ includes the same set of control variables as in (4.2), and μ_t are the survey wave fixed effects. The design of the regression is similar to a standard DID design, though there is no pure (unaffected) control group. That is, the coefficient α shows the gap in y_{irt} between the Russian and non-Russian speaking population before the war, and the coefficient δ shows the change in this gap after the start of the war. The coefficient δ can be interpreted as the differential treatment effect. We consider the opinions of Russia,

Table 4.7: Opinion of Russia, US and China in dynamics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Favorable opinion of Russia (binary)			Favorable opinion of the US (binary)			Favorable opinion of China (binary)		
Russian language (α)	-0.00010	0.033	-0.017	0.092***	0.087***	0.093***	0.17***	0.19***	0.16***
	(0.018)	(0.030)	(0.023)	(0.018)	(0.030)	(0.023)	(0.019)	(0.030)	(0.024)
Post war (γ)	-0.13***	-0.14***	-0.13***	0.029	0.042	0.019	0.16***	0.12***	0.17***
	(0.020)	(0.033)	(0.025)	(0.020)	(0.033)	(0.025)	(0.019)	(0.032)	(0.025)
Russian language * Post war (δ)	0.13***	0.16***	0.098***	-0.078***	-0.11***	-0.060**	0.015	0.031	-0.0016
	(0.019)	(0.031)	(0.025)	(0.020)	(0.032)	(0.026)	(0.020)	(0.031)	(0.026)
Observations	9,520	3,732	5,788	9,520	3,732	5,788	9,520	3,732	5,788
R-squared	0.111	0.126	0.108	0.137	0.094	0.088	0.141	0.169	0.127
Waves	All	All	All	All	All	All	All	All	All
Sample	All	Above 40	Below 40	All	Above 40	Below 40	All	Above 40	Below 40
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. Table shows the results of regression (4.3) for binary dependent variables: "Favorable Opinion of Russia" (columns 1-3), "Favorable Opinion of the US" (columns 4-6), "Favorable Opinion of China" (columns 7-9). Robust standard errors are in parentheses.

the US, and China (binary and scaled from 1 to 4) as our dependent variables.

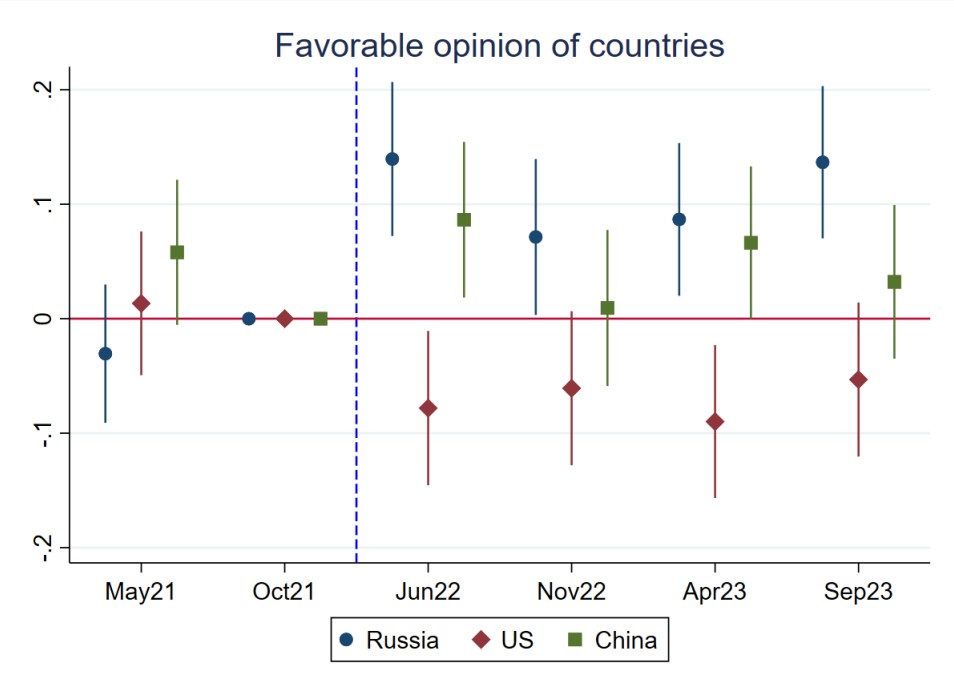
Table 4.7 shows the results. Before the full-scale invasion, the Russian-speaking respondents had the same opinion of Russia ($\alpha = 0$ in Column 1) and more favorable opinion of the US ($\alpha > 0$ in Column 4) than non-Russian-speaking respondents. These results change substantially after February 2022 (coefficient δ).

After February 2022, Russian-speaking respondents had a significantly better opinion of Russia than non-Russian-speaking respondents ($\delta > 0$), while non-Russian-speaking respondents opinion decreased substantially compared to 2021 ($\gamma < 0$) (Column 1). These results apply to younger and older populations (Columns 2, 3). Moreover, Column 4 shows that the positive attitude of Russian-speaking respondents towards the US decreased substantially stronger than for non-Russian-speaking respondents ($\delta < 0$). In contrast, the attitude toward the US of non-Russian-speaking respondents remained stable ($\gamma = 0$). These results apply to younger and older populations (Columns 5, 6). We consider China to be a neutral country (de jure) in the conflict. The Russian-speaking respondents had a more favorable opinion of China before the invasion, and there was no change in this difference after the invasion. Figure 4.9 shows that the parallel trend assumption holds for attitudes on Russia and the US so that the results of DID estimates from Table 4.7 can have the causal interpretation.

These results suggest that the argument that the Russian-speaking population in Kaza-

khstan is more supportive of the current war because they have always been closer to Russia and share historical and cultural ties and views with the Russian population does not hold.

Figure 4.9: Event study design about favorable opinion of countries



4.5 Media consumption and sanction evasion opinion

In this section, we will show how public opinion is connected to the consumption of media sources stemming from Russia. To his end, we use original survey experiment data collected in November 2023 in Kazakhstan. The survey consisted of demographic question, detailed question about media consumption, and included a list experiment (see section 4.C.2). We asked participants about their attitudes to the Russian invasion of Ukraine, as well as their opinion on whether avoiding Western sanctions and helping Russia can be 'justified'. Because these topics are politically sensitive and invite preference falsification, we concluded the survey with a list experiment in which participants could indicate their opinion without having their privacy compromised. The following section 4.5.1 will report detailed results and provide a discussion.

Table C1 presents summary statistics for all variables. The sample was 48.7% male

and 42.7 years old (SD=16.1), on average. Participants were asked on a 1 - 5 scale about their financial situation (M=2.85, SD=1.14), with the average corresponding to the answer option *We have enough money for food and clothes, but buying durable goods, such as a TV or refrigerator, is difficult.* Participants reported having received education on a scale ranging from 1 - 6 (M=4.55, SD=1.75), their residency (where 22% (SD=0.41) reported living in either of the two large cities Astana or Almaty), and being ethnically Russian (M=0.25, SD=0.43). 30.1% (SD=0.46) of participants reported consuming news sources from Russia. Weights for age and sex were applied to make the sample nationally representative.

4.5.1 Results

Table 4.8: Regression Models Direct Question

	(1)	(2)	(3)
VARIABLES	Sanction evasion	SMO/invasion Ukraine	Working 50h/w
Russian Media consumption	0.14***	0.086***	0.0012
	(0.016)	(0.015)	(0.015)
Constant	0.26***	0.19**	0.15**
	(0.076)	(0.073)	(0.070)
Observations	5,025	5,025	5,025
R-squared	0.082	0.089	0.097
Region FE	Yes	Yes	Yes
Robust SE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. Linear regression models, where the dependent variable is responding "justified" to the question *In your opinion, is* [item] *generally justified or not justified?*. Each row in the table represents a different item. Exact wording can be found in Section 4.3.3. We control for age, age squared, education, ability to speak Russian, Russian ethnicity, privacy perceptions, financial situation, and regions. All variables were used in their categorical form for the estimation. Robust standard errors in parentheses.

Table C2 shows the results from the regression models of the direct questions and Table 4.9 shows the results from the regression models of the list experiment. Each column represents a different item of interest, as indicated by the name of the column. Controls for language capabilities, ethnicity, regions, age, education, and income are included in all

models. The exact experimental setup and framing can be found in Table A2.

Table 4.9: Regression Models List Experiment

VARIABLES	(1) Sanction evasion	(2) SMO/invasion Ukraine	(3) Working 50h/w
Treatment	0.30***	0.22***	0.22***
	(0.039)	(0.038)	(0.039)
Treatment * Russian Media consumption	0.14**	0.076	0.025
	(0.072)	(0.068)	(0.070)
Observations	5,025	5,025	5,025
R-squared	0.092	0.059	0.089
Region FE	Yes	Yes	Yes
Robust SE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. Linear regression models, where the dependent variable is the number of items with which survey participants agreed. The coefficient *Treatment* captures whether participants were in a treated list, i.e., whether the item of interest was in their list or not. Correspondingly, the value of the coefficient reports the share of participants across our sample that agreed with the item of interest. Each row in the table represents a different item. Exact wording can be found in A2. We control for age, age squared, education, ability to speak Russian, Russian ethnicity, financial situation, and regions. Robust standard errors in parentheses.

In our regression models using the direct questions, we find that 26.1% of respondents think avoiding sanctions can be justified, whereas 18.6% of respondents indicate thinking that the Russian invasion of Ukraine was justified. For those consuming media sources from Russia[17], these numbers increase substantially (13.5%, $p < 0.001$) with respect to avoiding sanctions, and less pronounced (8.7%, $p < 0.001$) with respect to support for the war. In our regression models using the list experiment, we find that 30.3% of respondents think avoiding sanctions can be justified, whereas 22.4% of respondents agree that the Russian invasion of Ukraine was justified. For those consuming media sources from Russia, these numbers increase substantially (14.1%, $p = 0.049$) with respect to avoiding sanctions, and less pronounced (7.6%, $p = 0.265$) with respect to support for the war.

As indicated by the difference between the coefficients of the first row in Table C2 and the first row in Table 4.9, there is some preference falsification, which is to be expected when asking sensitive questions in an authoritarian context. Notably, both methods yield very

[17] As opposed to media sources *in* Russian

similar results, in that the magnitude of of the coefficients of Russian Media Consumption is almost the same in both estimations. By using the list experiment we can rule out that preference falsification has substantially biased our results.

These results link our previous findings to those presented in the next section. Above we established that Russian media consumption is found to be significantly higher among those who are able to speak the language and, crucially, independent of their ethnicity. Here we find evidence that the consumption of news from Russia has a strong positive effect on pro-Russian attitudes. In the next section, we provide some evidence that it is not only attitudes that are influenced, but also economic behavior in the form of sanctions evasion.

4.6 Russian propaganda and help in sanctions circumvention

In this section, we study how exposure to Russian propaganda affects the export activity of sanctioned products to Russia from Kazakhstan. Our aggregate export data from Kazakhstan to Russia includes 2899 observations at the region(20 regions)-month(33 months)-product class (5 classes) level.

Figure 4.2 shows that before the war, the Kazakhstan export to Russia of sanctioned products was relatively stable and small - around 2.4M USD per month-region (or 47M USD monthly for all regions), while the export of other products was 3.5 times larger - 8.4 M USD per month-region-product class[18] (or 645M USD monthly for all region and products). After the invasion, the export of sanctioned products became 11M USD per month-region (with a spike in Autumn 2022 reaching 20M USD per month-region) and accounting for 211M USD monthly. The export of non-sanctioned products stayed around 7.8 M USD per month-region-product class (568M USD monthly for all regions and products). In Section 4.6, we show the role of Russian propaganda in this sanctions circumvention process.

We start by estimating the average effect of war and sanctions on the export activity of sanctioned products to Russia. For this, we use the difference-in-differences, where the treatment group includes sanctioned products (e.g., electronics, machines, and weapons), the control group includes all non-sanctioned products, and February 2022 is the treatment

[18]Cloth 0.4M USD, Food 2.6M USD, Metals 12.8M USD, Other products 17M USD per month-region

date. We consider the following linear equation:

$$Value_{rtp} = \beta S_p Post_t + [\delta Capital_r S_p Post_t] + \kappa_r + \lambda_t + \mu_p + \varepsilon_{rtp}. \tag{4.4}$$

Here r denotes region, t – time (year-month) and p – product class. The outcome variable $Value_{rtp}$ is the export value (in million USD). The binary variables S_p denote the sanctioned products (electronics, machines, weapons), $Post_t$ – the period from February 2022 onwards, $Capital_r$ – Astana city (current capital) or Almaty city (former capital). Variables κ_r, λ_t, and μ_p are regional, time, and product class fixed effects, respectively. In specification (4.4)) without bracket term, coefficient β captures the average treatment effect of war and sanctions on the region-month export value of sanctioned products to Russia. Column 1 of Table 4.10 shows that due to the war and sanctions, the export transactions of sanctioned products increased by 9.7M USD per month-region, which well corresponds to Figure 4.2 and may be an indication of the re-export activity of firms in Kazakhstan.

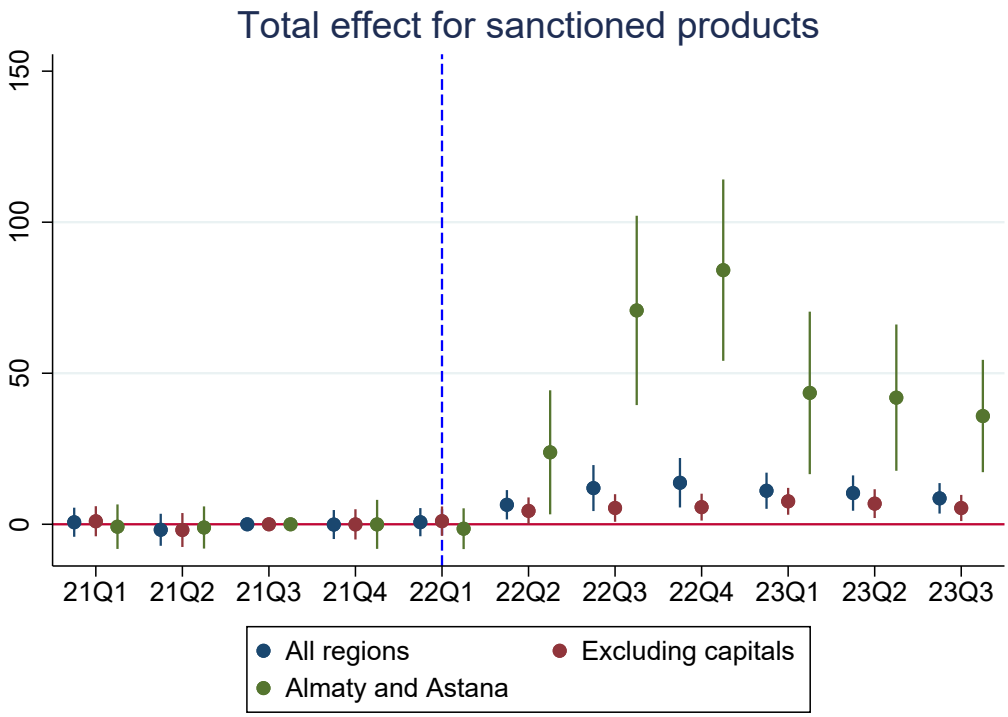
Due to the substantial concentration of business in the two largest cities – Astana and Almaty – we also show how the export activity of sanctioned products in these cities changed after the war and compare this effect with the effect for other regions. In specification (4.4)) with bracket term, coefficient δ captures this effect for Astana and Almaty, while coefficient β – for other cities. Column 2 of Table 4.10 shows that the month-region export of sanctioned products from Astana and Almaty increased by 46.5M USD, while it is substantially smaller for other regions – 4.9M USD. Figure 4.10 shows the Event Study design coefficients β_τ and δ_τ:

$$Value_{rtp} = \sum_{\tau=21Q1}^{23Q3} \beta_\tau S_p I(t=\tau) + \left[\sum_{\tau=21Q1}^{23Q3} \delta_\tau S_p I(t=\tau)\right] + \kappa_r + \lambda_t + \mu_p + \varepsilon_{rtp},$$

justifying the assumption of parallel trends for DID specifications (4.4). Due to substantial differences in the overall effects for capital cities and the rest of the regions, in the forthcoming analysis of the effect of exposure to pro-Kremlin information, we either exclude Astana and Almaty or show a standalone coefficient for them.

The analysis above showed a spike in the export value of sanctioned products to Russia. But does this spike depend on the extent of exposure to pro-Kremlin information about the

Figure 4.10: The dynamics of the gap in trade between sanctioned and other products



war? To answer this question, we consider the following triple DID specification:

$$Value_{rtp} = \alpha S_p Treat_r + \beta S_p Post_t + \gamma S_p Post_t Treat_r + [\delta Capital_r Post_t] + \kappa_r + \lambda_t + \mu_p + \varepsilon_{rtp}. \tag{4.5}$$

Variable $Treat_r$ uses CAB surveys and captures three different regional measures of treatment exposure to the pro-Kremlin information about the war as the share of (i) Russian-speaking respondents, (ii) respondents justifying the war, (iii) respondents trusting Russian news. Other variables are the same as in (4.4). When considering all regions, we include a term in brackets in (4.5); otherwise, we exclude it. When we exclude Astana and Almaty, coefficient α measures the prewar differences in the export activity of (not yet) sanctioned products depending on the treatment exposure. Coefficient β measures the effect of war and sanctions on the export activity of sanctioned products for a hypothetical region with zero treatment exposure, while the coefficient γ captures the additional contribution of treatment

exposure to this effect. When we include Astana and Almaty, the interpretation of the coefficients β and γ preserves, while δ measures the effect of the help in sanctions circumvention for these cities.

Table 4.10 presents the estimation results for model (4.5). Columns 4, 7, and 10 exclude Astana and Almaty, while Columns 3, 6, and 9 include them. The coefficient β is smaller in all these models compared to the baseline model of Column 2 and, in most cases, is either negative (Columns 9, 10) or insignificant (Columns 3, 6, 7). This suggests that in a hypothetical region with zero treatment exposure, the effect of the help of sanctions circumvention is either negative or zero. On the contrary, coefficient γ is positive and significant in all the considered models, suggesting that exposure to pro-Kremlin news stimulates firms in these regions to export more sanctioned products. Among the three measures of treatment exposure, the *Share of speaking Russian* demonstrates the weakest effect, while the *Share of trusting Russian news* – the strongest, being slightly above the *Share of justifying the war*. This evidence suggests that the main driver of the help in sanctions circumvention is Russian media, not cultural proximity. Figure 4.11 combines the information from coefficients β and γ and shows how help in sanctions circumvention depends on the treatment exposure. At the mean treatment exposure[19], the export of sanctioned products to Russia increases by 5.6M USD per region-month compared to the prewar time in non-capital regions. Figure 4.12A shows the coefficients from the Event Study design γ_τ

$$Value_{rtp} = \alpha S_p Treat_r + \beta S_p Post_t + \sum_{\tau=21Q1}^{23Q3} \gamma_\tau S_p I(t=\tau) Treat_r + \kappa_r + \lambda_t + \mu_p + \varepsilon_{rtp},$$

excluding the capital cities. It justifies the assumption of parallel trends for triple DID specifications (4.5).

Finally, we restrict our attention to sanctioned products only and implement the DID specification similar to the one in Korovkin and Makarin (2023):

$$Value_{rt} = \gamma Post_t Treat_r + \kappa_r + \lambda_t + \varepsilon_{rt}, \tag{4.6}$$

[19]Mean values: 0.42 for the Russian language, 0.27 for war justification, and 0.37 for trust in Russian media

with non-binary treatment intensities.[20] In this specification, coefficient γ identifies the differential effect of war and sanctions between regions with high and low degrees of treatment exposure. For the same reasons as before, we exclude Astana and Almaty from the sample. Columns 5, 8, and 11 of Table 4.10 show the results for the regional share of Russian speaking, justifying the war, and trusting Russian news, respectively. Similar to the previous findings, sanctions had a stronger effect on exports of sanctioned products to Russia for regions with higher treatment exposure. However, the magnitude depends on the exposure measure. An increase in the share of Russian-speaking respondents from the first quartile (.204) to the third quartile (.605) increases the additional export by 3.14M USD per month-region. Similarly, increasing the share of justifying the war and trusting the Russian news for one interquartile range (0.111 and 0.071, respectively) increases the additional export of sanctioned products by 2.4M USD and 0.86M USD per month-region. This finding also justifies the previous finding that exposure to pro-Kremlin information affects the firms' incentives to help Russia circumvent sanctions. Figure 4.12B shows the coefficients from the Event Study design γ_τ

$$Value_{rt} = \sum_{\tau=21Q1}^{23Q3} \gamma_\tau I(t = \tau) Treat_r + \kappa_r + \lambda_t + \varepsilon_{rt},$$

excluding the capital cities. It justifies the assumption of parallel trends for DID specifications (4.6).

[20] One can think of regions with low and high treatment exposure as control and treatment groups, respectively, in a binary treatment setting.

Table 4.10: The effect of Russian propaganda on sanctions circumvention

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Dependent variable: value of region-month-product export to Russia (in M USD)										
Sanctioned products x			-8.71***	-7.98***		-17.3**	-30.9***		-14.6**	-21.8***	
treatment exposure (α)			(2.42)	(2.44)		(8.09)	(7.90)		(7.37)	(7.05)	
Sanctioned products x	9.76***	4.90***	0.78	2.44**		-1.87	-1.56		-6.13**	-4.35*	
post-invasion (β)	(1.35)	(0.96)	(1.18)	(1.20)		(2.41)	(2.18)		(2.89)	(2.57)	
Sanctioned products x post-			9.56***	7.41***	7.84***	25.2***	26.2***	21.9***	29.7***	26.7***	12.1***
invasion x treatment exposure (γ)			(2.65)	(2.61)	(1.59)	(9.14)	(8.37)	(5.75)	(8.08)	(7.23)	(3.03)
Almaty or Astana x		46.5***	46.4***			47.1***			47.2***		
post-invasion (δ)		(6.18)	(6.18)			(6.24)			(6.19)		
Constant	18.1***	18.0***	18.0***	15.0***	2.21***	18.0***	15.0***	0.41	18.0***	15.0***	1.36*
	(0.88)	(0.90)	(0.90)	(0.80)	(0.38)	(0.90)	(0.80)	(0.96)	(0.90)	(0.80)	(0.72)
Observations	2,899	2,899	2,899	2,569	531	2,899	2,569	531	2,899	2,569	531
R-squared	0.289	0.352	0.353	0.298	0.612	0.353	0.298	0.602	0.353	0.298	0.590
Year-Month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Product class FE	Y	Y	Y	Y	-	Y	Y	N	Y	Y	-
Product class	All	All	All	All	Sanctioned	All	All	Sanctioned	All	All	Sanctioned
Sample	All regions	All regions	All regions	No capitals	No capitals	All regions	No capitals	No capitals	All regions	No capitals	No capitals
Regional treatment exposure	No	No	Share of speaking Russian at home			Share of justifying the war			Share of trusting Russian news		

Robust standard errors in parentheses

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. Table shows the results of regression (4.4).

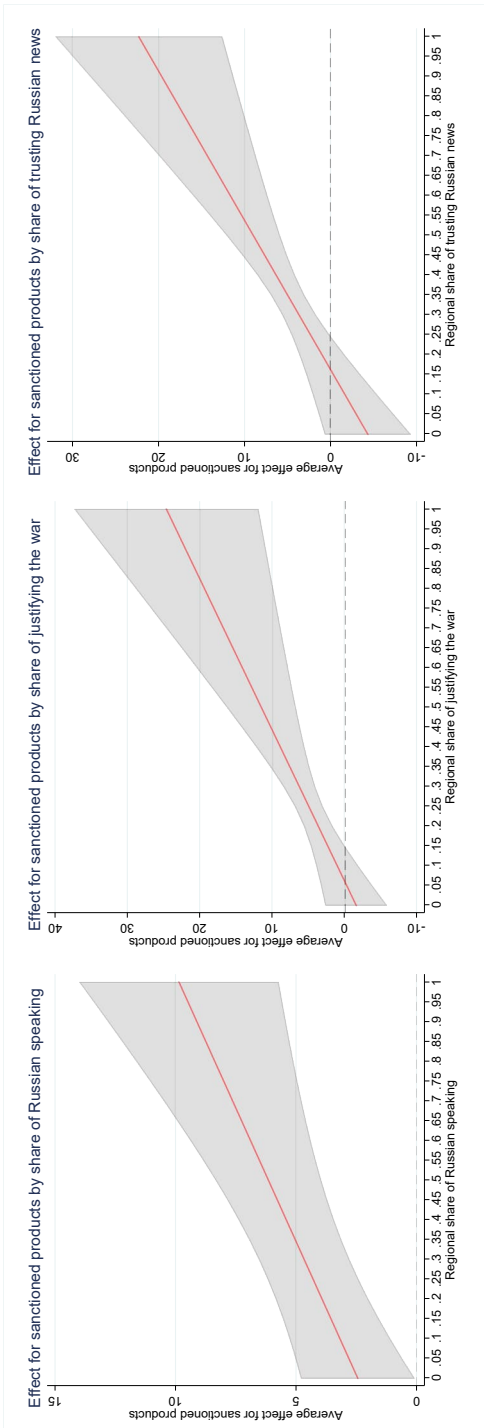


Figure 4.11: Sanctions circumvention by regional treatment exposure

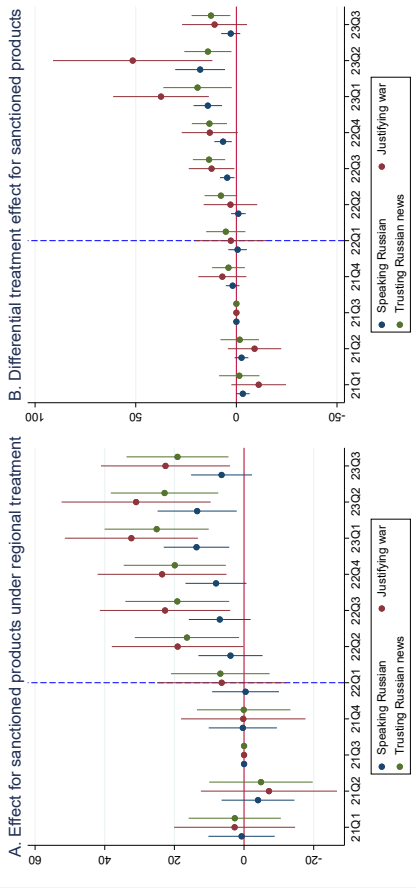


Figure 4.12: The dynamics of the effect of Russian propaganda on sanctions circumvention

4.7 Conclusion

Can language affinity facilitate the export of propaganda by authoritarian regimes during times of war? To test this question, we use data from six waves of a telephone survey conducted in Kazakhstan, and complement it with Yandex search data, our own original survey, as well as detailed trade data on exports from Kazakhstan to Russia. We find that Russia's full-scale invasion of Ukraine in February 2022 caused a spike in searches for news on the topic in Kazakhstan. Searching for information on the internet, Russian-speakers in Kazakhstan were significantly more likely to end up on pro-Kremlin media platforms than non-Russian speakers. We document how the consumption of pro-Kremlin media led to a persistent change and division in beliefs between Russian-speaking and non-Russian-speaking Kazakh citizens, a division that so did not exist before February 2022. After February 2022, Russian-speakers had a significantly more negative view of the West and a more positive view of Russia than non-Russian speakers. They were also more likely to condone Russia's invasion of Ukraine, and to justify the circumvention of economic sanctions imposed against Russia. Using detailed trade date, we show how these beliefs map into a significant increase in trade with Russia after February 2022, with the increase being particularly strong in those Kazakh regions where people are more likely to speak Russian and to consume pro-Kremlin media.

Our paper contributes to a literature that shows how the media in authoritarian regimes can have a strong influence on beliefs and behavior of the population. While the Russian case has been particularly well documented (Enikolopov et al. 2011; Zhuravskaya et al. 2020; Enikolopov et al. 2022), other cases include such diverse autocracies as Nazi Germany (Adena et al. 2015), Brazil's military dictatorship in the 1960s and 1970s (Schneider 2014), the Gulf States and Saudi Arabia (Leber and Abrahams 2019), China (Chen 2019; Weiss and Dafoe 2019; Pan et al. 2022) or North Korea (Ba et al. 2023). In these contexts, state actors are often engaged in industrialized efforts to create alternative narratives.[21] These efforts at disinformation campaigns have culminated in what some have called "post-truth" or

[21] As evidenced for example by Russia's troll "factories", https://www.nytimes.com/2018/02/18/world/europe/russia-troll-factory.html

"pseudo-realities" (Jones 2022). These false narratives take hold domestically, are replicated, and later exported globally. During the last decade, Russian-spawned narratives have taken hold in many different places around the world, influencing major events. Russian bots and user accounts interfered in both the UK's Brexit referendum and the 2016 US presidential election (Alizadeh et al. 2020; Earl et al. 2021). Recorded instances also include Russian misinformation campaigns about the Black Lives Matter movement in an attempt to polarize the public and create conflict (Earl et al. 2022). In the context of Russia's full-scale invasion of Ukraine in early 2022, Russian campaigns were launched to sell an alternative narrative, in order to create opposition to Western support for Ukraine and help circumvent Western-imposed sanctions against Russia. In this context, our paper is the first to show how this strategy is producing tangible results – both in changing public perceptions and beliefs, and in reducing public concerns about circumventing Western-imposed sanctions against Russia.

We argue in this paper that Russia's ability to do this, as well as Russia's invasion of Ukraine itself, are both legacies of empire - the Russian and the Soviet one. As evidenced elsewhere, Russia's full-scale invasion of Ukraine cannot be explained by rational choice theories of authoritarian politics alone, that depict the dictator as a rational actor (Sonin 2024). Instead, Russia's imperial past has become reason and motivation for Putin's renewed interest in territorial expansion (Schulze-Wessel 2023). Those ethnic Russians who found themselves outside of Russia after the fall of the Soviet Union have started to play a pivotal role in this respect (Zevelëv 2001; Ziegler 2006). Putin has made the Russian diaspora an essential part of his foreign policy and a way of exercising soft power (Decker 2021). The annexation of Crimea in 2014, the war in the Donbass, as well as the start of the full-scale invasion in 2022 were all at least partially justified with the need to protect Russian-speakers abroad (Pieper 2020).

In our paper, we show how the Russian diaspora has become object and actor alike in Putin's effort to rebuild the Russian empire. We identify one factor that makes populations particularly vulnerable to Russian propaganda alternatives: language. Our findings are consistent with the literature on media effects in autocracies, and underscore the importance of language and online media in this respect. Other studies have found that Russian narratives also often refer to a specific Russian identity (Spies et al. 2022; Tolz and Hutchings 2023),

which is beyond the scope of this study, but certainly warrants further investigation. However, neither identity nor language fully explain how propaganda in the form of narratives is transferred to other contexts, as the adoption of Russian propaganda in China (Hanley et al. 2024) or in English-speaking political discussions on Reddit (Hanley et al. 2023) show, suggesting additional avenues of research on this topic.

4.8 Bibliography

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., and Zhuravskaya, E. (2015). Radio and the rise of the nazis in prewar germany. *Quarterly Journal of Economics*, 130:1885–1939.

Agerberg, M. and Tannenberg, M. (2021). Dealing with measurement error in list experiments: Choosing the right control list design. *Research & Politics*, 8(2):20531680211013154.

Ahlquist, J. S. (2018). List Experiment Design, Non-Strategic Respondent Error, and Item Count Technique Estimators. *Political Analysis*, 26(1):34–53.

Alizadeh, M., Shapiro, J. N., Buntain, C., and Tucker, J. A. (2020). Content-based features predict social media influence operations. *Science Advances*, 6(30):eabb5824.

Ba, H., Greitens, S., and Kim, Y. B. (2023). Authoritarian Audiences: Theory and Evidence for Subnational Propaganda Targeting in North Korea. *Comparative Politics*, 56(1):49–72.

Blair, G., Coppock, A., and Moor, M. (2020). When to Worry About Sensitivity Bias: A Social Reference Theory and Evidence From 30 Years of List Experiments. *American Political Science Review*, 114(4):1297–1315.

Blair, G. and Imai, K. (2012). Statistical Analysis of List Experiments. *Political Analysis*, 20(1):47–77.

Blair, G., Imai, K., and Lyall, J. (2014). Comparing and Combining List and Endorsement Experiments: Evidence From Afghanistan. *American Journal of Political Science*, 58(4):1043–1063.

Bond, R., Fariss, C., Jones, J., Kramer, A., Marlow, C., Settle, J., and Fowler, J. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489:295–8.

Borozna, A. and Kochtcheeva, L. (2024). *War by Other Means: Western Sanctions on Russia and Moscow's Response*. Palgrave.

Chen, D. (2019). Political context and citizen information: Propaganda effects in china. *International Journal of Public Opinion Research*, 31(3):463–484.

Coffman, K. B., Coffman, L. C., and Ericson, K. M. M. (2013). The Size of the LGBT Population and the Magnitude of Anti-Gay Sentiment are Substantially Underestimated. Working Paper 19508, National Bureau of Economic Research.

Corstange, D. (2012). Vote Trafficking in Lebanon. *International Journal of Middle East Studies*, 44(3):483–505.

Crystal, D. (2012). *English as a GLobal Language*. Cambridge: Cambridge University Press.

Dadabaev, T. and Sonoda, S. (2023). Silence is golden? silences as strategic narratives in central asian states' response to the ukrainian crisis. *International Journal of Asian Studies*, 20(1):193–215.

Decker, P. K. (2021). "we show what is concealed": Russian soft power in germany. *Problems of Post-Communism*, 68(3):216–230.

DellaVigna, S. and Kaplan, E. (2007). The Fox News Effect: Media Bias and Voting. *The Quarterly Journal of Economics*, 122(3):1187–1234.

Durante, R. and Knight, B. (2012). Partisan Control, Media Bias, and Viewer Responses: Evidence from Berlusconi's Italy. *Journal of the European Economic Association*, 10(3):451–481.

Eady, G., Paskhalis, T., Zilinsky, J., Bonneau, R., Nagler, J., and Tucker, J. (2023). Exposure to the russian internet research agency foreign influence campaign on twitter in the 2016 us election and its relationship to attitudes and voting behavior. *Nature Communications*, 14.

Earl, J., James, R., Ramo, E., and Scovill, S. (2021). Protest, activism, and false information. In *The Routledge companion to media disinformation and populism*, pages 290–301. Routledge.

Earl, J., Maher, T., and Pan, J. (2022). The Digital Repression of Social Movements, Protest, and Activism: A Synthetic Review. *Science advances*, 8.

Enikolopov, R., Makarin, A., and Petrova, M. (2020). Social media and protest participation: Evidence from russia. *Econometrica*, 88(4):1479–1514.

Enikolopov, R., Petrova, M., and Zhuravskaya, E. (2011). Media and Political Persuasion: Evidence from Russia. *American Economic Review*, 101(7):3253–85.

Enikolopov, R., Rochlitz, M., Schoors, K., and Zakharov, N. (2022). The effect of independent online media in an autocracy. *SSRN Electronic Journal*.

Frye, T., Gehlbach, S., Marquardt, K. L., and Reuter, O. J. (2017). Is Putin's Popularity Real? *Post-Soviet Affairs*, 33(1):1–15.

Frye, T., Gehlbach, S., Marquardt, K. L., and Reuter, O. J. (2023). Is Putin's Popularity (Still) Real? A Cautionary Note on Using List Experiments to Measure Popularity in Authoritarian Regimes. *Post-Soviet Affairs*, pages 1–10.

Gerber, A. S., Karlan, D., and Bergan, D. (2009). Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions. *American Economic Journal: Applied Economics*, 1(2):35–52.

Glynn, A. N. (2013). What Can We Learn with Statistical Truth Serum?: Design and Analysis of the List Experiment. *Public Opinion Quarterly*, 77(S1):159–172.

Guiso, L. and Makarin, A. (2020). Affinity, Trust, and Information. *EIEF Working Papers Series*.

Hanley, H., Lu, Y., and Pan, J. (2024). Narratives of foreign media ecosystems in chinese social media discussions of the russo-ukrainian war. *Working paper*.

Hanley, H. W. A., Kumar, D., and Durumeric, Z. (2023). Happenstance: Utilizing semantic search to track russian state media narratives about the russo-ukrainian war on reddit. *Proceedings of the International AAAI Conference on Web and Social Media*, 17(1):327–338.

Imai, K. (2011). Multivariate regression analysis for the item count technique. *Journal of the American Statistical Association*, 106(494):407–416.

Jones, M. (2022). *Digital Authoritarianism in the Middle East: Deception, Disinformation and Social Media*. Hurst.

Korovkin, V. and Makarin, A. (2023). Conflict and intergroup trade: Evidence from the 2014 russia-ukraine crisis. *American economic review*, 113(1):34–70.

Kramon, E. and Weghorst, K. (2019). (Mis)Measuring Sensitive Attitudes with the List Experiment: Solutions to List Experiment Breakdown in Kenya. *Public Opinion Quarterly*, 83(S1):236–263.

Kuhn, P. M. and Vivyan, N. (2022). The Misreporting Trade-Off Between List Experiments and Direct Questions in Practice: Partition Validation Evidence from Two Countries. *Political Analysis*, 30(3):381–402.

Larreguy, H. and Martinez, D. (2024). Misinformation, Distrust, and Economic Outcomes: Evidence from the Effect of Russian Misinformation in Latvia. *Manuscript*, 0(0):0.

Leber, A. and Abrahams, A. S. (2019). A storm of tweets: Social media manipulation during the gulf crisis. *Review of Middle East Studies*, 53:241–258.

Lehtisaari, K., Burkhanov, A., Turdubaeva, E., and Pietiläinen, J. (2018). Framing of Crimean Annexation and Eastern Ukraine Conflict in Newspapers of Kazakhstan and Kyrgyzstan in 2014. In Resende, E., Budrytė, D., and Buhari-Gulmez, D., editors, *Crisis and Change in Post-Cold War Global Politics*, chapter 7, pages 181–201. Palgrave Macmillan.

Leyris, T. (2022). Building an Imperial Broadcasting Network as the Empire Disintegrated – The Birth of Radio in the French sub-Saharan African Colonies during Decolonization. *Journal of Radio & Audio Media*, 29(1):120–135.

Martin, D. A., Shapiro, J. N., and Nedashkovskaya, M. (2019). Recent trends in online foreign influence efforts. *Journal of Information Warfare*, 18(3):15–48.

Mattingly, D. C. and Yao, E. (2022). How soft propaganda persuades. *Comparative Political Studies*, 55(9):1569–1594.

Morozov, E. (2011). Liberation technology: Whither internet control? *Journal of Democracy*, 22(2):62–74.

Neathery-Castro, J. and Rousseau, M. (2005). Does French Matter? France and Francophonie in the Age of Globalization. *The French Review*, 78(4):678–693.

Oleinik, A. (2024). Telegram channels covering russia's invasion of ukraine: a comparative analysis of large multilingual corpora. *Journal of COmputational Social Science*, 0:0.

Pan, J., Shao, Z., and Xu, Y. (2022). How government-controlled media shifts policy attitudes through framing. *Political Science Research and Methods*, 10(2):317–332.

Peisakhin, L. and Rozenas, A. (2018). Electoral Effects of Biased Media: Russian Television in Ukraine. *American Journal of Political Science*, 62(3):535–550.

Phillipson, R. (2008). The Linguistic Imperialism of Neoliberal Empire. *Critical Inquiry in Language Studies*, 5(1):1–43.

Pieper, M. (2020). Russkiy mir: The geopolitics of russian compatriots abroad. *Geopolitics*, 25(3):756–779.

Roberts, M. E. (2018). *Censored*. Princeton University Press, Princeton.

Robinson, D. and Tannenberg, M. (2019). Self-censorship of Regime Support in Authoritarian States: Evidence From List Experiments in China. *Research & Politics*, 6(3):2053168019856449.

Rosenfeld, B., Imai, K., and Shapiro, J. N. (2016). An Empirical Validation Study of Popular Survey Methodologies for Sensitive Questions. *American Journal of Political Science*, 60(3):783–802.

Sablina, L. (2023). The russian-language pro-war movement in germany amid the invasion of ukraine. *Nations and Nationalism*, 29(4):1422–1438.

Schneider, N. (2014). *Brazilian Propaganda: Legitimizing an Authoritarian Regime*. University Press of Florida.

Schulze-Wessel, M. (2023). *Der Fluch des Imperiums: Die Ukraine, Polen und der Irrweg in der russischen Geschichte*. C.H.Beck.

Snyder, T. (2018). *The Road To Unfreedom*. Crown, New York.

Sonin, K. (2024). Contemporary Imperialism and the Main Reasons for the Russo-Ukrainian War. *Essay, University of Chicago*.

Spies, D. C., Mayer, S. J., Elis, J., and Goerres, A. (2022). Why do immigrants support an anti-immigrant party? russian-germans and the alternative for germany. *West european politics*, 46(2):275–299.

Tolz, V. and Hutchings, S. (2023). Truth with a z: disinformation, war in ukraine, and russia's contradictory discourse of imperial identity. *Post-Soviet Affairs*, 39(5):347–365.

Vaillant, D. (2017). *Across the Waves: How the United States and France Shaped the International Age of Radio*. Champaign: University of Illinois Press.

Vigouroux, C. (2013). Francophonie. *Annual Review of Anthropology*, 42:379–397.

Weiss, J. and Dafoe, A. (2019). Authoritarian audiences, rhetoric, and propaganda in international crises: Evidence from china. *International Studies Quarterly*, 63(4):963–973.

Yanagizawa-Drott, D. (2014). Propaganda and Conflict: Evidence from the Rwandan Genocide. *The Quarterly Journal of Economics*, 129(4):1947–1994.

Zevelëv, I. (2001). *Russia and Its New Diasporas*. United States Institute of Peace.

Zhuravskaya, E., Petrova, M., and Enikolopov, R. (2020). Political effects of the internet and social media. *Annual Review of Economics*, 12(1):415–438.

Ziegler, C. E. (2006). The russian diaspora in central asia: Russian compatriots and moscow's foreign policy. *Demokratizatsiya*, 14(1).

Zizzo, D. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13(1):75–98.

4.A Appendix

Figure A1: Opinions of different countries and the Eurasian Economic Union (EAEU) in dynamics

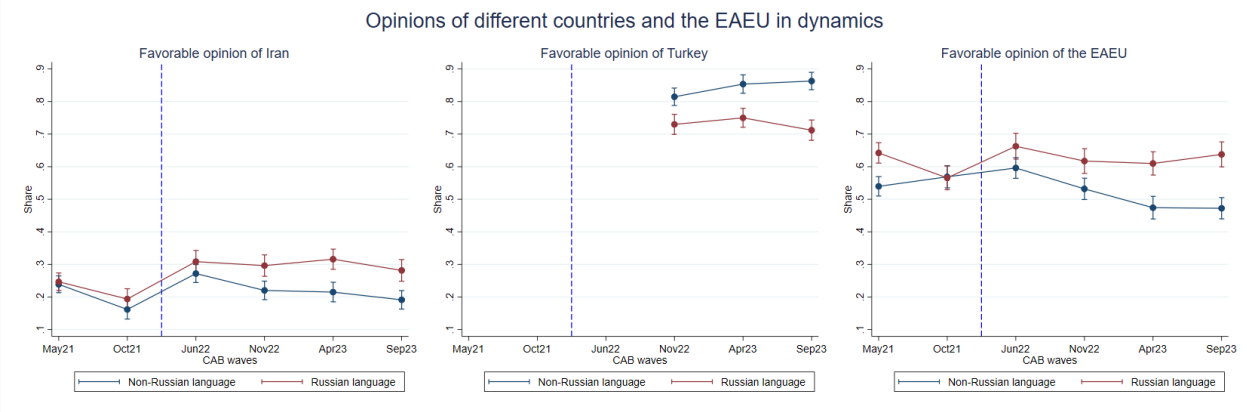


Figure A2: Main source of news about events outside Kazakhstan in dynamics

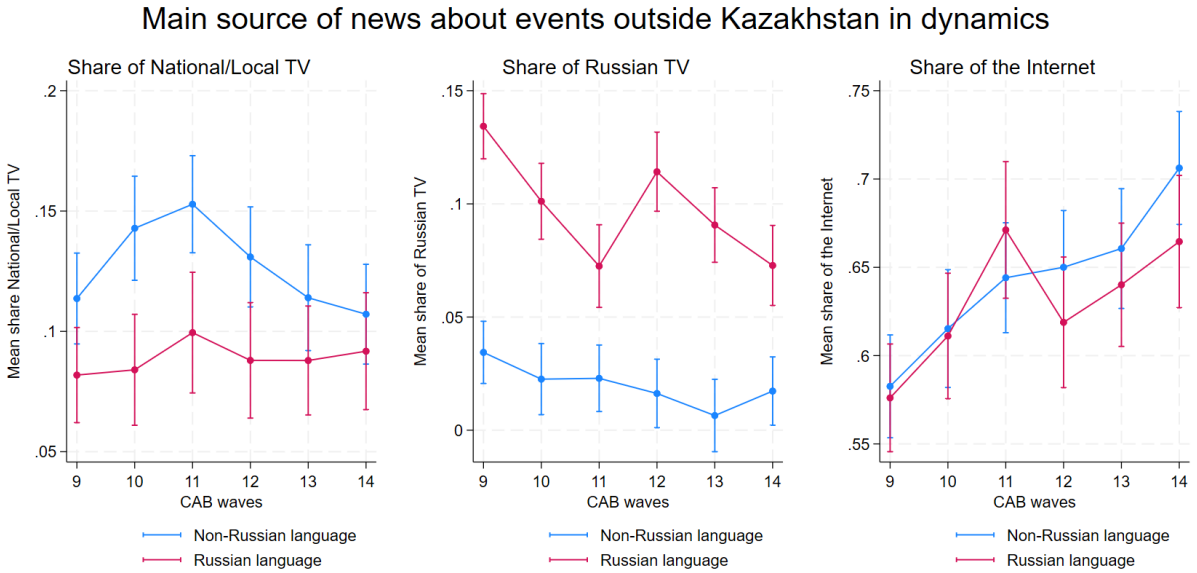


Figure A3: The usage of social media and messengers in dynamics

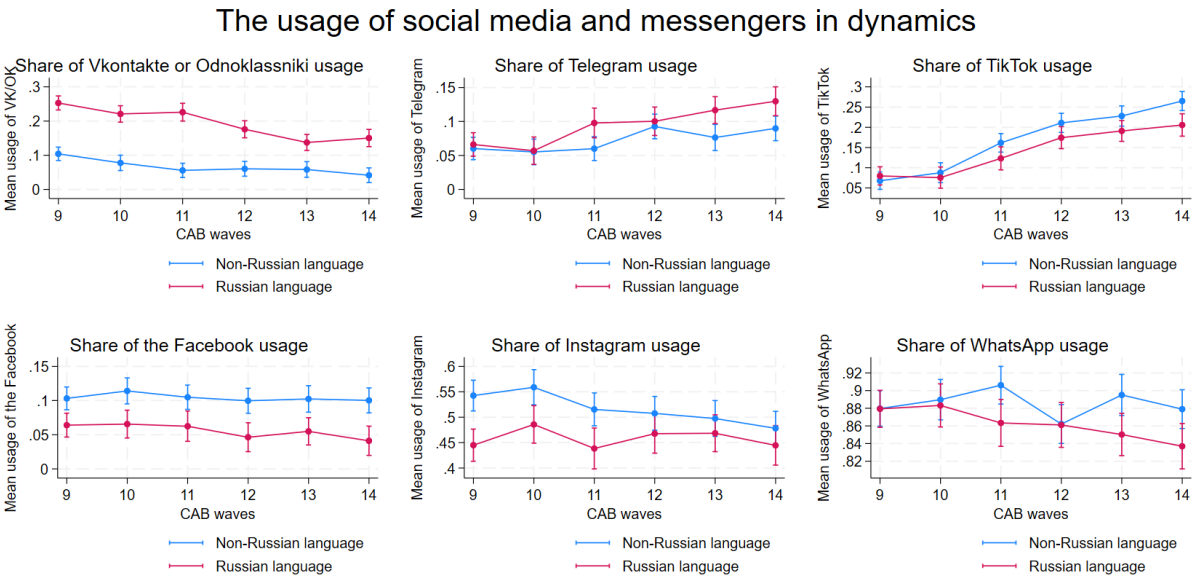


Figure A4: Media consumption for news or entertainment in dynamics

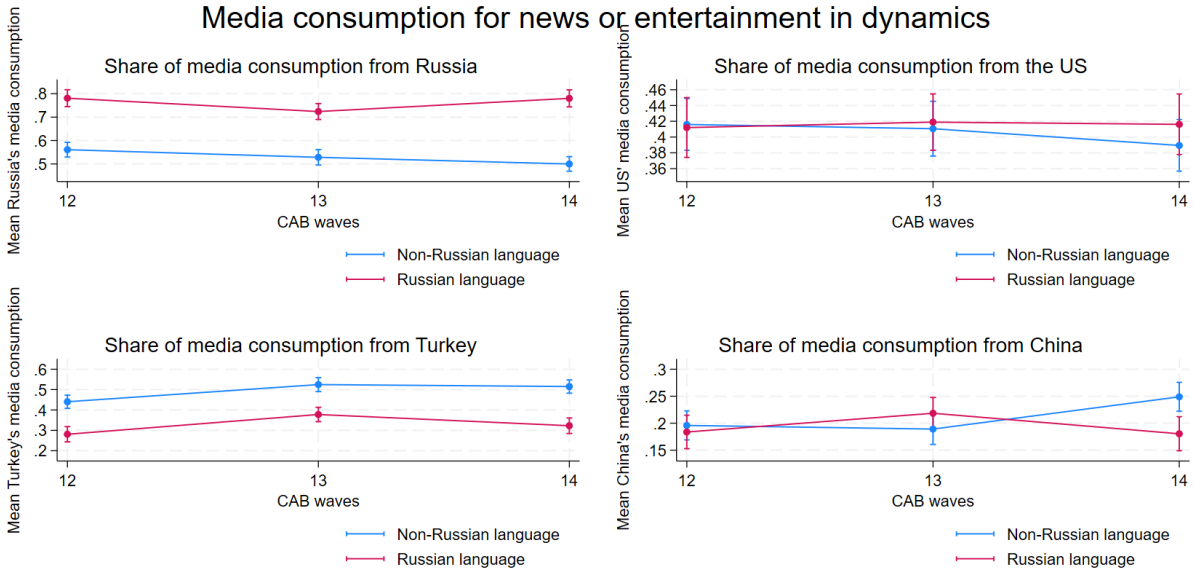


Figure A5: Search for "Yandex Dzen" on Yandex

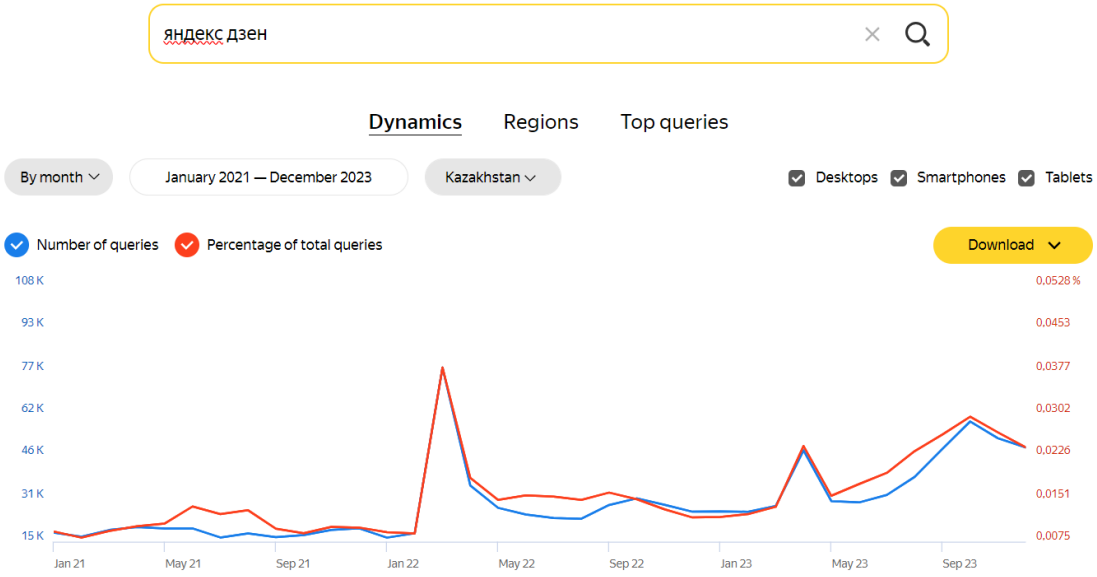


Figure A6: Search for "Yandex News" on Yandex

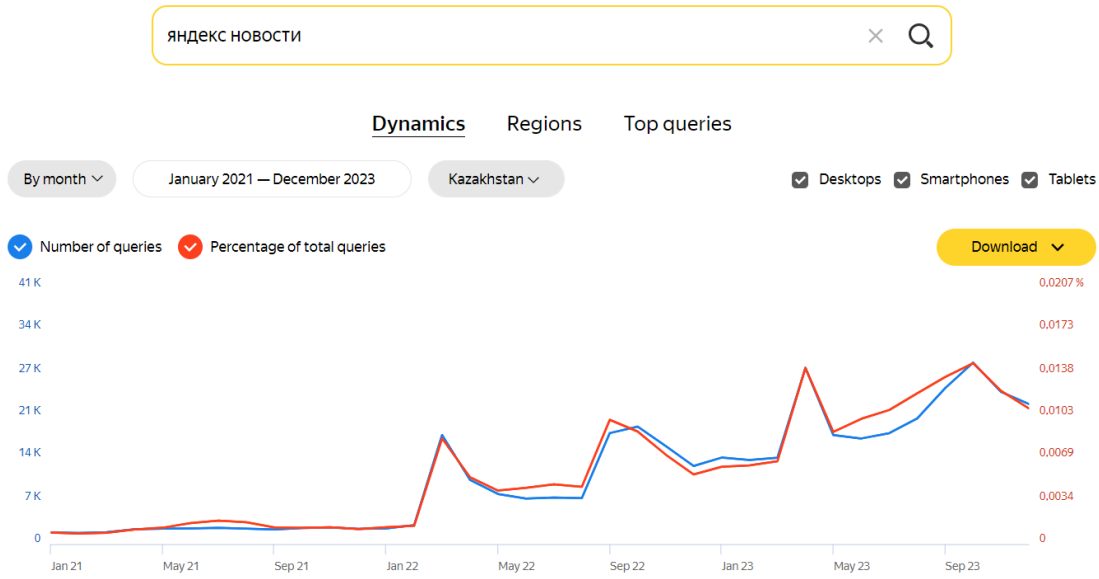


Figure A7: Search for "Euronews" on Yandex

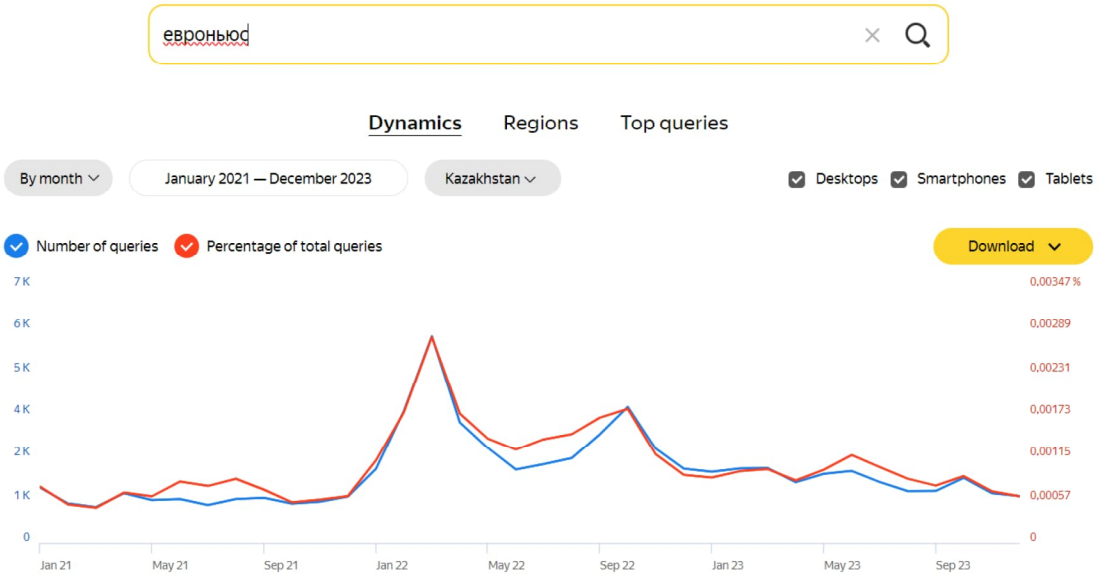
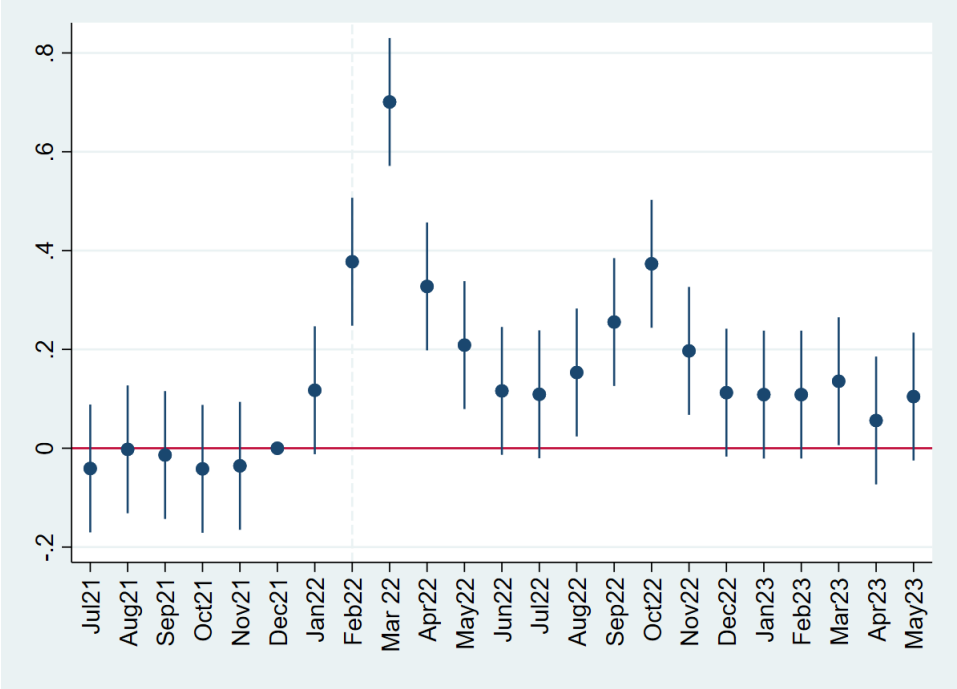


Figure A8: Event study design about the search for "Euronews"



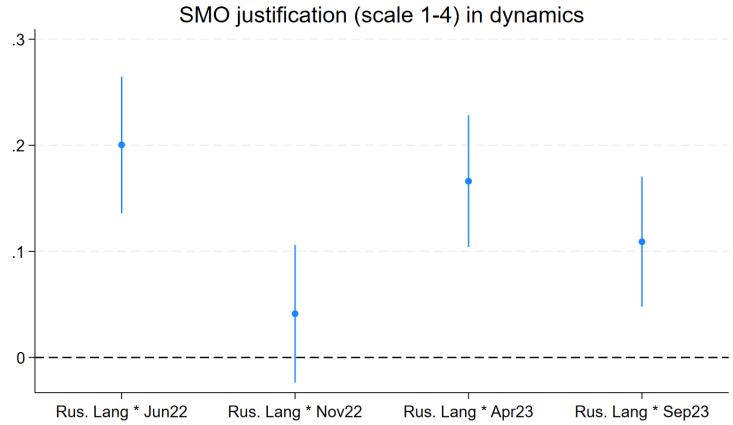
4.B Appendix

Table B1: SMO justification (scale from 1 to 4)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	SMO is justified (scale)					
Russian language	0.27***	0.13	0.44***	0.16***	0.31***	0.26***
	(0.042)	(0.11)	(0.074)	(0.051)	(0.051)	(0.043)
Observations	4,458	740	1,818	2,640	3,200	3,566
R-squared	0.207	0.227	0.309	0.137	0.241	0.133
Waves	After 2022	After 2022	After 2022	After 2022	After 2022	After 2022
Sample	All	Firm head	Above 40	Below 40	Informed	No Russians

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. The table shows the regression results (4.2) for a discrete dependent variable measured from 1 "SMO is completely unjustified" to 4 "SMO is completely justified". Robust standard errors are in parentheses.

Figure B1: The effect on SMO justification in dynamics



Note. Coefficients are divided to the mean of the SMO justification (scale 1-4) variable, so they can be interpreted in the percentage of the mean.

Table B2: Daily use of the Internet via a phone (binary)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Daily Internet Use (binary)					
Russian language	0.042***	0.017	0.040*	0.043***	0.047***	0.040***
	(0.012)	(0.024)	(0.023)	(0.013)	(0.014)	(0.012)
Observations	6,020	955	2,366	3,654	3,889	4,864
R-squared	0.076	0.089	0.090	0.039	0.094	0.083
Waves	After 2022	After 2022	After 2022	After 2022	After 2022	After 2022
Sample	All	Firm head	Above 40	Below 40	Informed	No Russians

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. Table shows the results of regression (4.2) for a binary dependent variable, Daily use of the Internet via a phone. Robust standard errors in parentheses.

Table B3: Frequency of Internet use via a phone (scale 1-5)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Frequency of Internet Use (from 1 daily to 5 never)					
Russian language	-0.069***	-0.038	-0.082	-0.061**	-0.080***	-0.065**
	(0.026)	(0.043)	(0.054)	(0.027)	(0.029)	(0.026)
Observations	6,020	955	2,366	3,654	3,889	4,864
R-squared	0.080	0.090	0.089	0.038	0.093	0.089
Waves	After 2022	After 2022	After 2022	After 2022	After 2022	After 2022
Sample	All	Firm head	Above 40	Below 40	Informed	No Russians

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. Table shows the results of regression (4.2) for a discrete dependent variable with values from 1 (daily) to 5 (never), Frequency of using the Internet via a phone. Robust standard errors in parentheses.

Table B4: Channels for Russian propaganda export: National or Local TV, radio, newspapers

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	News from National or Local TV, radio, newspapers					
Russian language	-0.036***	-0.012	-0.035	-0.035***	-0.046***	-0.033***
	(0.010)	(0.016)	(0.023)	(0.0097)	(0.013)	(0.010)
Observations	6,020	955	2,366	3,654	3,889	4,864
R-squared	0.072	0.061	0.063	0.038	0.089	0.087
Waves	After 2022	After 2022	After 2022	After 2022	After 2022	After 2022
Sample	All	Firm head	Above 40	Below 40	Informed	No Russians
Region FE	Yes	Yes	Yes	Yes	Yes	Yes

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. Table shows the results of regression (4.2) for a binary dependent variable, News from National or Local TV, radio, newspapers. Robust standard errors in parentheses.

Figure B2: The effect on the main source of news and the use of social media in dynamics

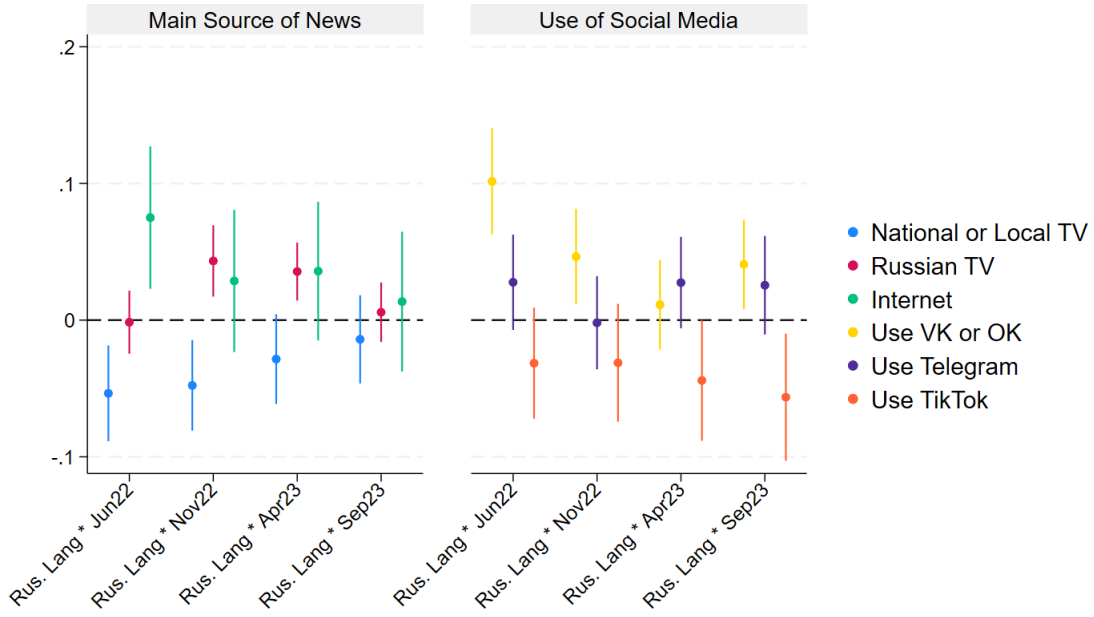


Table B5: Media consumption for news or entertainment from Russia

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Use Media from Russia					
Russian language	0.11***	0.039	0.11***	0.11***	0.13***	0.11***
	(0.020)	(0.049)	(0.033)	(0.026)	(0.024)	(0.021)
VARIABLES	Use Media from the US					
Russian language	0.054***	0.080	-0.036	0.098***	0.047*	0.060***
	(0.021)	(0.050)	(0.035)	(0.026)	(0.026)	(0.021)
VARIABLES	Use Media from Turkey					
Russian language	-0.11***	-0.14***	-0.13***	-0.099***	-0.15***	-0.11***
	(0.021)	(0.049)	(0.036)	(0.026)	(0.027)	(0.021)
VARIABLES	Use Media from China					
Russian language	-0.018	-0.060	-0.020	-0.018	-0.031	-0.015
	(0.017)	(0.043)	(0.031)	(0.022)	(0.023)	(0.018)
Observations	4,511	741	1,768	2,743	2,872	3,611
Waves	After 2022	After 2022	After 2022	After 2022	After 2022	After 2022
Sample	All	Firm head	Above 40	Below 40	Informed	No Russians
Region FE	Yes	Yes	Yes	Yes	Yes	Yes

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. Table shows the results of regression (4.2) for a binary dependent variable, Use Media for news or entertainment from Russia. Robust standard errors in parentheses.

Table B6: Trust News or Information from Russia

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Trust News or Information from Russia (binary)					
Russian language	0.070*	-0.10	0.10	0.052	0.066	0.076**
	(0.036)	(0.087)	(0.063)	(0.045)	(0.048)	(0.037)
VARIABLES	Trust News or Information from the US (binary)					
Russian language	0.017	-0.088	-0.013	0.032	0.032	0.019
	(0.034)	(0.082)	(0.053)	(0.045)	(0.046)	(0.035)
VARIABLES	Trust News or Information from Turkey (binary)					
Russian language	-0.090**	-0.31***	-0.14**	-0.062	-0.12**	-0.089**
	(0.039)	(0.093)	(0.065)	(0.049)	(0.051)	(0.040)
VARIABLES	Trust News or Information from China (binary)					
Russian language	0.068**	-0.051	0.12**	0.037	0.088**	0.060*
	(0.031)	(0.073)	(0.049)	(0.040)	(0.043)	(0.032)
Observations	1,500	245	621	879	963	1,209
R-squared	0.076	0.174	0.138	0.076	0.100	0.062
Waves	After 2022	After 2022	After 2022	After 2022	After 2022	After 2022
Sample	All	Firm head	Above 40	Below 40	Informed	No Russians
Region FE	Yes	Yes	Yes	Yes	Yes	Yes

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. Table shows the results of regression (4.2) for a binary dependent variable, Trust News or Information from Russia. Robust standard errors in parentheses.

Figure B3: The effect on the media consumption in dynamics

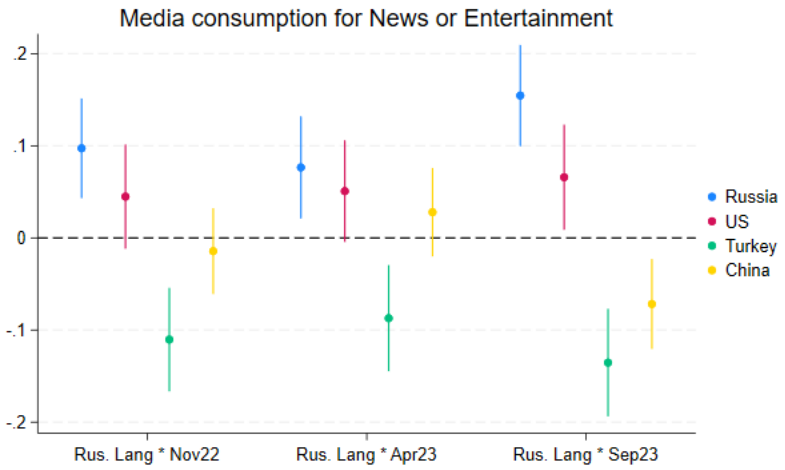


Figure B4: Main source of news about events outside Kazakhstan: event study

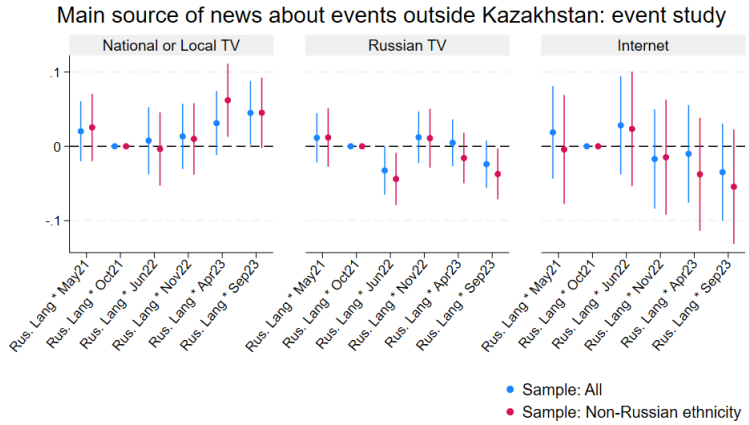


Figure B5: The usage of social media and messengers I: event study

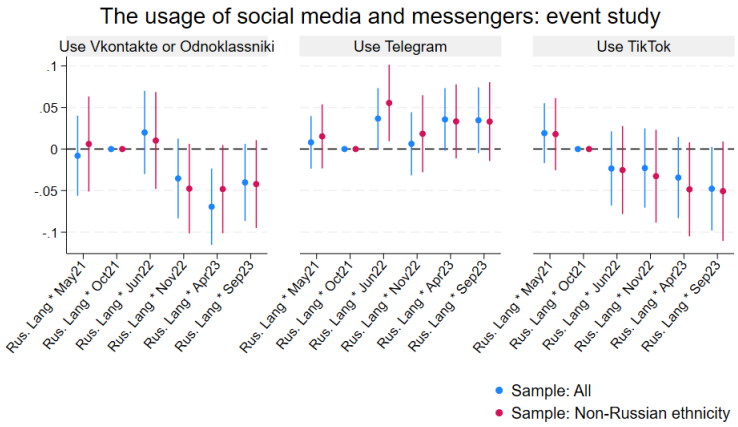


Figure B6: The usage of social media and messengers II: event study

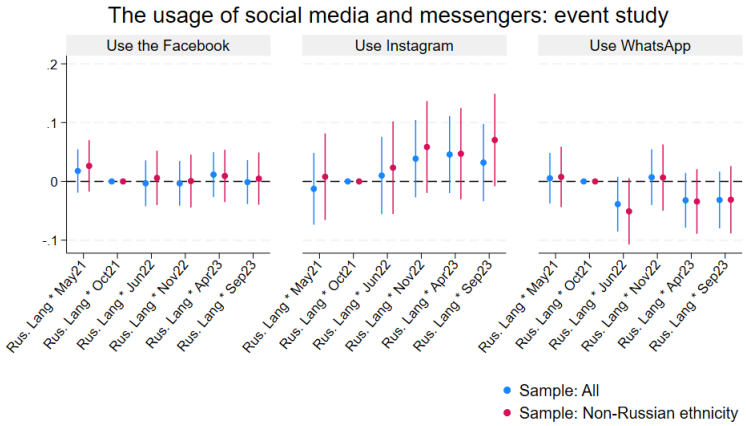
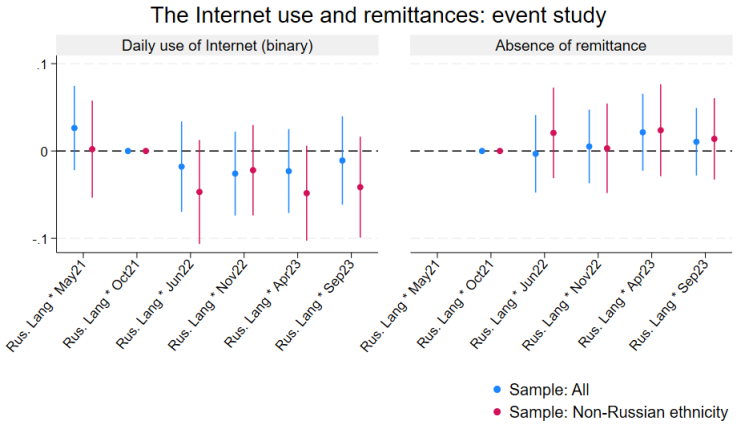


Figure B7: The Internet use and remittances: event study



4.C Appendix

4.C.1 List experiment

Table C1: Summary statistics

Variable	N	Mean	SD
age	5025	42.7	16.1
education			
... Complete higher education	5025	0.232	0.422
... Complete secondary education/ Complete school	5025	0.298	0.457
... Complete Vocational education (vocational school, lyceum, college, technical school, etc.)	5025	0.208	0.406
... Incomplete higher education	5025	0.0756	0.264
... Incomplete secondary education/ Incomplete school	5025	0.0796	0.271
... Incomplete Vocational education (vocational school, lyceum, college, technical school, etc.)	5025	0.0903	0.287
... No education	5025	0.0161	0.126
ethnicity			
... Kazakh	5025	0.722	0.448
... Other	5025	0.033	0.179
... Russian	5025	0.245	0.43
financial_situation			
... We are not experiencing financial difficulties and, if necessary, can buy anything	5025	0.0625	0.242
... We don't have enough money even for food	5025	0.148	0.355
... We have enough money for everything except very expensive purchases like a car or apartment	5025	0.247	0.431
... We have enough money for food and clothing, but it's difficult to buy durable goods like a TV or refrigerator	5025	0.319	0.466
... We have enough money for food, but not for clothing	5025	0.224	0.417
male	5025	0.487	0.5
region			
... Abai	5025	0.0306	0.172
... Akmola	5025	0.0312	0.174
... Aktobe	5025	0.0494	0.217
... Almaty city	5025	0.135	0.341
... Almaty region	5025	0.0446	0.206
... Astana city	5025	0.085	0.279
... Atyrau	5025	0.035	0.184
... East Kazakhstan	5025	0.0269	0.162
... Jambyl	5025	0.0396	0.195
... Jetisu	5025	0.0189	0.136
... Karaganda	5025	0.0529	0.224
... Kostanay	5025	0.0283	0.166
... Kyzylorda	5025	0.109	0.312
... Mangystau	5025	0.0386	0.193
... North Kazakhstan	5025	0.0271	0.162
... Pavlodar	5025	0.0593	0.236
... Shymkent	5025	0.0605	0.238
... Turkistan	5025	0.089	0.285
... Ulytau	5025	0.0109	0.104
... West Kazakhstan	5025	0.0283	0.166
russian_media_high	5025	0.301	0.459

Table C2: Regression Models Direct Question

VARIABLES	(1) Sanction evasion	(2) SMO/invasion Ukraine	(3) Working 50h/w
Russian Media consumption	0.14***	0.086***	0.0012
	(0.016)	(0.015)	(0.015)
treatment_n = 2, privacy	0.012	0.0023	0.0087
	(0.016)	(0.015)	(0.016)
treatment_n = 3, surveillance	-0.0090	0.0049	0.0099
	(0.016)	(0.015)	(0.016)
framing_invasion==SMO		0.0089	
		(0.012)	
ethnicity_russian	0.10***	0.0069	-0.093***
	(0.018)	(0.017)	(0.017)
language_russian	-0.15***	-0.14***	-0.14***
	(0.017)	(0.017)	(0.018)
age_sqr	0.000053*	0.000026	-0.000022
	(0.000030)	(0.000030)	(0.000029)
age	-0.0030	-0.00050	0.0028
	(0.0028)	(0.0027)	(0.0027)
education_scale = 1, No education	0.11*	0.13**	0.077
	(0.060)	(0.059)	(0.061)
education_scale = 2, Incomplete secondary education	0.11***	0.21***	0.083***
	(0.031)	(0.030)	(0.030)
education_scale = 3, Complete secondary education	0.034*	0.098***	0.036*
	(0.019)	(0.017)	(0.019)
education_scale = 4, Incomplete vocational education	0.045	0.066**	0.0087
	(0.027)	(0.026)	(0.026)
education_scale = 5, Complete vocational education	0.025	0.039**	0.015
	(0.020)	(0.018)	(0.019)
education_scale = 6, Incomplete higher education	0.011	0.011	0.024
	(0.028)	(0.027)	(0.029)
financial_situation_scale = 1, Not enough money even for food	0.081**	0.039	0.035
	(0.034)	(0.033)	(0.034)
financial_situation_scale = 2, Enough money for food but not clothes	0.053*	-0.012	0.038
	(0.031)	(0.030)	(0.031)
financial_situation_scale = 3, Enough money for food and clothes but not TV or refrigerator	0.043	-0.011	0.0097
	(0.030)	(0.029)	(0.030)
financial_situation_scale = 4, Enough money for everything but car or apartment	-0.032	-0.053*	0.0019
	(0.031)	(0.030)	(0.030)
Constant	0.26***	0.19**	0.15**
	(0.076)	(0.073)	(0.070)
Observations	5,025	5,025	5,025
R-squared	0.082	0.089	0.097
Region FE	Yes	Yes	Yes
Robust SE	Yes	Yes	Yes

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. Linear regression models, where the dependent variable is responding "justified" to the question *In your opinion, is* [item] *generally justified or not justified?*. Each row in the table represents a different item. Exact wording can be found in Section 4.3.3. We control for age, age squared, education, ability to speak Russian, Russian ethnicity, privacy perceptions, financial situation, and regions. All variables were used in their categorical form for the estimation. The treatment privacy and treatment surveillance variables correspond to text-based reminders about the privacy in the experiment and lack thereof, respectively. Robust standard errors in parentheses.

4.C.2 List experiment

List experiments – also known as the item count technique – have been successfully used, for example, to study support for authoritarian leaders (Blair et al. 2014; Frye et al. 2017; Robinson and Tannenberg 2019; Frye et al. 2023), estimating the size of LGBT population (Coffman et al. 2013), and vote trafficking in Lebanon (Corstange 2012).

Participants are exposed to either J or $J + 1$ items and then asked to count the number of items that apply to them, with the additional (J + 1th) item being the sensitive item of interest.[22] The premise of list experiments is that when a sensitive question is asked indirectly, respondents are more likely to give a truthful answer, even if social norms encourage them to answer the question in a particular way (Blair and Imai 2012). Fear of being judged or punished by others leads to a change in behavior best known as social desirability bias, a subset of what is known as sensitivity bias (Blair et al. 2020). More sophisticated statistical methods allow analysis beyond mean comparisons so that sensitivity bias can be shown, but also which sociodemographic factors and personality traits play a role (Imai 2011; Blair and Imai 2012).

In list experiments, inevitable *nonstrategic* misreporting can be be kept at a minimum by specific design choices (Kuhn and Vivyan 2022). Nonstrategic measurement errors are, for example, floor and ceiling effects (Ahlquist 2018), and complexity (Kramon and Weghorst 2019). Floor and ceiling effects appear when either all or none of the statements in the list are true, and survey participants thus must reveal their preference over the sensitive item in order to remain giving truthful answers. In this case, one would have to expect strategic misreporting, i.e., lying. List experiments thus help to elicit true preferences, but, by design, suffer from overwhelming or confusing study participants. By careful consideration of the design, these measurement errors can be minimized (Glynn 2013; Blair et al. 2020), for example by choosing a suitable control list (Agerberg and Tannenberg 2021). Is it furthermore common practice to validate experiments before going into the field (Rosenfeld et al. 2016).

Our list experiment takes this literature into account and is designed accordingly. Non-

[22] J is commonly equal to 3 or 4. Higher values substantially increase complexity, while lower values are risking to compromise the privacy of the participants (floor effect).

strategic measurement errors were tried to be minimized according to suggestions from the literature:

First, floor and ceiling effects were avoided by (1) choosing items where the prevalence among the population was known beforehand, and (2) choosing control list items that are negatively correlated. For example, the items 'homosexuality' and 'holding on to religious values' are strongly negatively correlated. It is very unlikely that respondents agree with both items. Furthermore, it is unlikely that none of the items are true, statistically. Items were taken and adapted from the World Value Survey and Central Asian Barometer. Thus, ceiling effects and floor effects could, by design, be minimized. Second, sensitive control list items were chosen, in order to conceal the item of interest and not evoke demand effects (Zizzo 2010). Third, a placebo item is used in order to test whether there are design effects (Frye et al. 2017, 2023). Fourth, a pre-test was run with 400 subjects in order to validate the lists. Last, a double list experiment was employed in order to enhance statistical power. In a double list experiment the sample is split into two groups, and each half acts as a control group for the other half, see Table A2.

Table C3: Balance Table List Experiment

Variable	Level (for categorical variables)	Group A	Group B	Difference
Framing	SMO	0.492	0.504	0.012
		(0.500)	(0.500)	(0.014)
	Invasion of Ukraine	0.508	0.496	-0.012
		(0.500)	(0.500)	(0.014)
Russian media consumption	Yes	0.295	0.307	0.012
		(0.456)	(0.461)	(0.013)
Ethnicity	Russian	0.237	0.253	0.016
		(0.425)	(0.435)	(0.012)
Language skills	Russian	0.628	0.616	-0.011
		(0.484)	(0.486)	(0.014)
Age		42.603	42.708	0.106
		(16.171)	(16.084)	(0.455)
Education	No education	0.018	0.014	-0.003
		(0.132)	(0.120)	(0.004)
	Incomplete secondary education	0.084	0.075	-0.009
		(0.277)	(0.264)	(0.008)
	Complete secondary education	0.313	0.283	-0.030**
		(0.464)	(0.450)	(0.013)
	Incomplete vocational education	0.085	0.096	0.012
		(0.278)	(0.295)	(0.008)
	Complete vocational education	0.198	0.219	0.022*
		(0.398)	(0.414)	(0.011)
	Incomplete higher education	0.076	0.075	-0.001
		(0.265)	(0.264)	(0.007)
	Complete higher education	0.227	0.237	0.009
		(0.419)	(0.425)	(0.012)
Financial Situation	Not enough money even for food	0.151	0.144	-0.007
		(0.358)	(0.351)	(0.010)
	Enough money for food but not clothes	0.225	0.224	-0.001
		(0.418)	(0.417)	(0.012)
	Enough money for food and clothes but not TV or fridge	0.312	0.325	0.013
		(0.463)	(0.469)	(0.013)
	Enough money for everything but car or apartment	0.250	0.244	-0.006
		(0.433)	(0.430)	(0.012)
	Enough money for everything	0.062	0.063	0.001
		(0.242)	(0.243)	(0.007)

Note. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. Group A corresponds to the left column and Group B to the right column of Table A2.

5 State Surveillance and Self-Censorship

Essay title: State Surveillance and Self-Censorship

Authors: David Karpa

Abstract

How do authoritarian states maintain their hegemony over public opinion beyond state-run media and outright repression of journalists? Theory and previous research suggest that surveillance practices discourage citizens from engaging in legitimate digital communication behaviors, such as expressing opinions online. Drawing on an original survey experiment conducted in Kazakhstan in November 2023 (N=5,025), this study is able to show that citizens exposed to a text-based surveillance treatment reduce their response rate to sensitive questions by 2.5% to 4%, while this effect is not triggered for non-sensitive questions. By comparing subgroups, heterogeneous treatment effects are identified, showing that treatment effects are much larger (9%) or non-existent for certain groups. In particular, older citizens, those living in rural areas, those with a high intensity of internet use, and those consuming media from Russia are prone to self-censorship. By comparing indirect and direct questions within a list experiment, the baseline level of self-censorship is calculated by which allows to distinguish pre-existing self-censorship from treatment effects. This study contributes to the literature on digital authoritarianism by showing how state surveillance practices undermine political discourse on the Internet, which in turn contributes to authoritarian stability.

Keywords: big data, surveillance, privacy, political repression, democracy, autocracy

Publication: Under Review at Political Behaviour (since 19. August 2024).

5.1 Digital mass surveillance

Over 75 countries worldwide use surveillance tools that are associated with artificial intelligence, including over 50% of advanced democracies (Feldstein 2019a). For example, Ethiopia, with its long-standing network of in-person surveillance, was a quick adopter and transitioned to digital surveillance despite initially having a low percentage of the population with access to the internet (Feldstein 2021). Many have raised the need to critically reflect on surveillance practices in contemporary societies, because of ongoing human rights violations.[1] Beyond ethical and human rights concerns, mass surveillance has been shown to have an effect on human behaviour by undermining autonomy and well-being, and inducing self-censorship (Büchi et al. 2022). Surveillance practices lead to a "spiral of silence", where people are deterred from exchanging opinions (online), particularly concerning sensitive topics (Stoycheff 2016). The rise of pre-emptive and conformist behaviour is in direct conflict with the essential components of deliberative democratic frameworks and represents a significant challenge to the healthy functioning of democratic societies (Penney 2022; Kappeler et al. 2023).

In contrast, within an autocratic context, the autocrat seeks anticipatory and obedient behaviors. Scholars have argued that autocrats refrain from directly repressing their population because of its net negative consequences (Guriev and Treisman 2019), and instead try to control the informational environment by co-opting the elite and media (Guriev and Treisman 2020). However, many of the long ruling autocrats like Russia's Putin or Turkey's Erdoğan have increasingly resorted to using violence on protesters, repressing dissidents, and imprisoning journalists, as a means to consolidating power (Pan and Siegel 2020; Egorov and Sonin 2024). Once feared, dictators strategically signal their surveillance and repression capabilities in order to enforce self-disciplining behaviour (Gohdes 2023). This self-disciplining behavior can come in many forms, but first and foremost, it results in self-censorship concerning political topics (Roberts 2018). Surveillance thus contributes to undermining democratic deliberation processes, democratic backsliding, and to authoritarian stability (Carothers and

[1] https://www.ohchr.org/en/press-releases/2022/09/spyware-and-surveillance-threats-privacy-and-human-rights-growing-un-report

Press 2022).

The main aim of this paper is to test these theories and investigate whether surveillance practices indeed induce self-censorship among citizens, to which magnitude, and what factors, if any, moderate this effect. To this end, a survey experiment with 5,025 participants was conducted in Kazakhstan, a country where the government has repeatedly deployed mass surveillance technology at the internet service provider level (Raman et al. 2020). Participants in the study were asked sensitive questions on domestic and geopolitical topics, after exposure to either a control, surveillance or privacy condition. The main results of this study are that participants in the surveillance condition indeed self-censor, 4% on items concerning domestic politics and between 2.5 and 3.2% on geopolitical topics, whereas exposure to the privacy treatment had no effects. Furthermore, strong heterogeneity in the surveillance treatment effects was detected, with effect sizes increasing up to three times the size, or diminishing entirely, for some demographic groups. This study adds to the literature, by following the call of Büchi et al. (2022) in experimentally investigating self-censorship induced by digital surveillance and estimating its magnitude. In addition, this study contributes to public opinion research by estimating a usually undetected baseline of self-censorship that leads to an overestimation of politically desirable attitudes in autocracies (Corstange 2012; Frye et al. 2017, 2023; Robinson and Tannenberg 2019; Tannenberg 2022). Finally, it adds to the literature on (digital) authoritarianism by showing how autocrats control the informational environment with digital tools (King et al. 2017; Roberts 2018; Guriev and Treisman 2019, 2020; Feldstein 2021; Gohdes 2023; Egorov and Sonin 2024). The following Section provides an overview of the relevant literature from which the hypotheses are derived. Section 5.3 embeds the hypotheses in the research design and elaborates on the methodological details of the study. Section 5.4 presents the results, while Section 5.5 concludes with a discussion of the results.

5.2 Literature

Social scientists who study digital surveillance sometimes call it *covert repression* (Earl et al. 2022), *dataveillance* (Festic 2022; Büchi et al. 2022; Kappeler et al. 2023; Lee 2023), *fear-based censorship* (Roberts 2018, 2020), or embed it into a broader discussion of *digital authori-*

tarianism (Feldstein 2019b, 2021; Jones 2022; Gohdes 2023). The literature distinguishes between research on digital surveillance in different types of regimes, because there is an important difference. In theory, government surveillance in democracies is an unintended side effect, a necessary evil of anti-terror or COVID measures. Independent institutions are supposed to monitor each other and keep power in check to protect civil liberties and individual rights. In the literature on autocracies, surveillance is a crucial tool in the state's repertoire of survival strategies, to the extent that it is strategically signalled to the population (Roberts 2018; Gohdes 2023). Accordingly, research on digital surveillance in autocracies tends to understand it as a form of state repression strategically deployed by autocrats to stay in power. This research is complemented by a political economy perspective that focuses on the mutual benefits of a private-public partnership in the development of surveillance technologies in autocracies (Liu 2019; Beraja et al. 2023b,c,a; Huang et al. 2022).

A literature agnostic to the institutional background revives Bentham's and Foucault's metaphors of the panopticon (Manokha 2018; Stoycheff et al. 2019), in which Bentham (2011) paints a picture of a prison called *Panopticon*, where a central tower oversees cells in the form of a ring around the tower in the centre. In essence, the theory of the panopticon involves three main assumptions (Manokha 2018): First, the omnipresence of the inspector, guaranteed by his total invisibility; second, the universal visibility of the objects of surveillance; and third, the *assumption* of constant observation of those being watched. Under this regime, inmates infer that they are under constant surveillance and thus exercise self-control and self-discipline. Coercion becomes unnecessary, except for a few rare instances of disobedience. Based on this design, Foucault (2012) developed his theory of self-discipline through *assumed* surveillance. The metaphor of the panopticon helps understanding the effects of state surveillance, as it involves a centralised entity (the watchtower/state) watching over the inmates/citizens. Even if in practice the actual surveillance is much more distributed than the metaphor suggests, the perception of a centralised state may dominate because of the at least partially opaque and sometimes highly technical processes involved. This perception of surveillance, similar to the original design of Bentham's prison, encourages self-discipline. Indeed, empirical studies explicitly used the metaphor of the panopticon in the context of digital environments (Stoycheff et al. 2019), while others implicitly described the mecha-

nism of self-discipline (mostly self-censorship) due to the *fear* of repression (Roberts 2018; Manokha 2018; Tannenberg 2022; Stoycheff 2022; Oz and Yanik 2022).

5.2.1 Repression, Fear and Chilling Effects

The importance of surveillance in authoritarian states can also be explained by the information dilemma of the authoritarian government. As a result of censorship, media control, and the absence or manipulation of elections, the regime does not know the true sentiments of its citizens (Edmond 2013; Xu 2021; Egorov and Sonin 2024). As a result, the efficient allocation of resources to co-opt regime opponents remains impossible, as the regime is uncertain about which actors require co-optation and which actors can be better controlled through repression. Such targeted co-optation or repression is necessary, however, because large-scale mass repression is rarely used in contemporary dictatorships (Guriev and Treisman 2019; Xu 2021), partly because of the disadvantages of international backlash in a globalized economy, but also because visible repression can signal regime weakness (Guriev and Treisman 2020). Surveillance of social media helps to identify protests early and monitor local governments and officials (Qin et al. 2017).

When dissidents were identified through surveillance, targeted repression of regime dissidents discourages and deters the participation of larger segments of the population (Roberts 2018; Xu 2021; Gohdes 2023). In autocracies, political expression and discussion are possible but very limited (King et al. 2017). By *taxing information* through propaganda, distraction, and censorship, free debate on political issues is hindered (Roberts 2018). Thus, political participation takes the form of protests or revolts because of the absence of meaningful elections and the censorship of grievances. More surveillance can lead to more repression since the authorities can act on the collected information (Earl et al. 2022). In sum, surveillance enables targeted repression, and the mere possibility of repression, in turn, induces *fear*, which leads to self-censorship (Roberts 2018).

In the discourse on surveillance in democracies, a related phenomenon has been referred to as *chilling effect*. Chilling effects – the deterrence of lawful behavior out of fear that it is suspect – have been studied by several scholars (Schauer 1978; Penney 2016, 2017; Stoycheff 2016; Stoycheff et al. 2019; Büchi et al. 2022). The core of democracy can be considered to

be the freedom to hold and express any political views. The discussion of political issues has increasingly moved to online spaces such as social media and text messengers, and while in online environments these expressions and debates of political opinion are vulnerable to surveillance. Theoretical studies of digital surveillance argue that *salience shocks*[2] of digital surveillance lead to inhibited digital communication behavior (Büchi et al. 2022). Recent research has suggested a common denominator in research on surveillance in autocracies and democracies: surveillance induces self-discipline (mostly self-censorship) due to the *fear* of repression (Roberts 2018; Manokha 2018; Tannenberg 2022; Stoycheff 2022; Oz and Yanik 2022). Citizens – when aware of surveillance practices – have an increased expectation of negative outcomes and will self-censor. In this vein, the first hypothesis is formulated as:

Hypothesis 1: Digital surveillance induces self-disciplining behavior in the form of self-censorship in politically sensitive topics.

5.2.2 Mass surveillance in Kazakhstan

Kazakhstan is a resource-rich Central Asian country bordering China and Russia. After the collapse of the Soviet Union, of which Kazakhstan was a part, the country gained independence and was ruled authoritatively for nearly three decades by former Party Secretary Nursultan Nazarbayev. Nazarbayev followed the model of the modern autocrat of the late 20th century, who didn't oppress his people with brutal force, but rather told the story of a man of the people while ensuring an acceptable minimum of living conditions (Guriev and Treisman 2019). In 2019, the country's leadership changed as Nazarbayev appointed a predecessor, Kassym-Jomart Tokayev. While this transition of power was initially successful, Tokayev eventually struggled with perceptions of illegitimacy (Kudaibergenova and Laruelle 2022; Silvan 2024). Growing protests culminated in the so-called "Bloody January" of 2022 – mass protests against corruption and economic inequality on an unprecedented scale were followed by a state of emergency and fighting between the military and protesters, with thousands arrested and hundreds killed (FreedomHouse 2023a). There have been reports of torture of protesters, activists, and journalists.[3]

[2]One such shock was Edward Snowden's revelations about the NSA's ongoing surveillance of US citizens.

[3]https://www.hrw.org/news/2024/01/31/longing-justice-kazakhstan

The government has broad powers to control the digital infrastructure, deriving its authority from laws and weak legal resistance. From controlling the content of websites through legal pressure to outright blocking of websites, to punishing journalists, there is widespread censorship (FreedomHouse 2023b). In addition, laws make anonymity online impossible, VPNs are cracked down on, and SIM cards – the access point to the internet for most of the population – must be registered with an ID. In 2019, Kazakhstan became the first country to force its population to install a custom root certificate capable of decrypting content running through the country's largest internet service provider. These surveillance capabilities have primarily targeted social media and communications services, making them seemingly a political rather than a security endeavor (Raman et al. 2020). While the root certificate was only active for about three weeks, it set a precedent and signaled the government's capabilities to the population. In addition to mass surveillance on the internet service provider level, government agencies monitor social media and communication apps targeting journalists, dissidents, and minorities (FreedomHouse 2023b). All this culminates in self-censorship on a large scale, especially when it comes to the two most important political issues – the "Bloody January" and Russia's invasion of Ukraine.

Other studies suggest that behavioural adaptations to surveillance include increased use of privacy-preserving technologies to cope with surveillance (Büchi et al. 2022; Kappeler et al. 2023). Censorship in the form of blocked websites is being bypassed with circumvention tools, leading to renewed access by citizens and increased interest in blocked content (Hobbs and Roberts 2018). In the same way, effective encryption mechanisms should *recover* digital communication behaviour. Given the baseline of digital surveillance in contemporary societies, particularly in Kazakhstan, the potential for recovering digital communication behaviour is significant (see Section 5.4.3). Correspondingly, the second hypothesis proposes that:

Hypothesis 2: Privacy-enhancing technology reduces self-disciplining behavior in the form of self-censorship in politically sensitive topics.

This study draws on this literature and investigates (1) whether digital surveillance induces self-censorship, (2) whether this effect can be reversed by a privacy-preserving tech-

nology, and (3) which role demographic factors play in moderating these effects.

5.3 Method and data

To test the hypotheses, an online survey experiment with 5,025 respondents was conducted in November 2023 in Kazakhstan. The survey was pre-registered[4] and carried out by NAC Analytica, a leading Kazakh sociological and public opinion research organization.[5] Participants were recruited through advertisements in social media, and a weighting-scheme was applied to make the sample nationally representative.

Before being randomly assigned to either a control group or one of the two treatment conditions, participants answered a range of demographic questions. The treatments were text-based information on the security of participants data. The treatment conditions differ with the control condition in that they either point out the possibility of the government being able to access information on online activity (surveillance condition) or ensure confidentiality by encryption (privacy condition). Section 5.A.1 in the Appendix presents the control and treatment scenarios. The control condition consists only of a standard experimental instruction without additional information.

After having faced either treatment, participants were asked four questions in random order, three of which are politically sensitive, and one that is not sensitive and acts as a placebo. The sensitive questions concerned domestic politics (*In your opinion, is participating in protests for political change generally justified or not justified?*) and geopolitics (*In your opinion, is helping Russia avoid Western sanctions generally justified or not justified?* and *In your opinion, is Russia's Special Military Operation/ invasion of Ukraine generally justified or not justified?*). The framing *Special Military Operation* and *invasion of Ukraine* was assigned at random, in order to balance invoked framing effects. Arguably, the way one describes Russia's invasion of Ukraine gives away their view on this war and thus invokes demand effects and social desirability bias. A neutral stance between the two mutually exclusive narratives of an illegitimate invasion or a 'Special Military Operation' is hard to find. Figure A1 in the appendix shows the difference between the two framings in

[4] https://aspredicted.org/BVT_9Z3

[5] https://nacanalytica.com/en/

terms of the outcome variable. While there are differences in the proportions, the dynamics of self-censorship develop analogously across treatments. Question 4 acted as a placebo, in order to control for design effects (*In your opinion, is working more than 50 hours per week generally justified or not justified?*). Answer options for the outcome variables were *Justified*, *Not justified*, and *Prefer not to answer*.

Quality controls included attention checks (two questions on respondents age had to match), speeding filters (minimum of 200 seconds), allowing only two completes per IP address, and allowing phone numbers to participate only once (payment was carried out by phone number). Out of 28,201 participants, 5,025 completed the survey, passed quality checks, were unique respondents, and were compensated 700 Tenge (approx. 1.50 USD). 25 respondents left the experiments after having faced the control (7), surveillance (7) or privacy (11) condition, respectively. Most of the participants that left the survey before finishing did so in the very first pages of the survey.

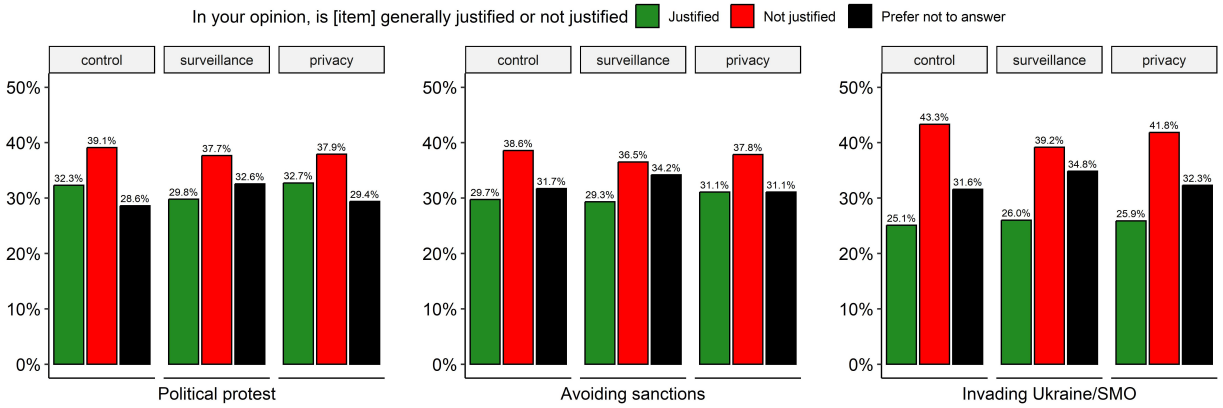
Table A1 presents summary statistics for all variables. Categorical variables were transformed to scales or dummies. The sample was 48.7% male and 42.7 years old (SD=16.1), on average. Participants were asked on a 1 - 5 scale about their financial situation (M=2.85, SD=1.14), with the mean corresponding to the answer option *We have enough money for food and clothes, but buying durable goods, such as a TV or refrigerator, is difficult*. Participants reported having received education on a scale ranging from 1 - 6 (M=4.55, SD=1.75), their residency (where 22% (SD=0.41) reported living in either of the two large cities Astana or Almaty), and being ethnically Kazakh (M=0.72, SD=0.45). 30.1% (SD=0.46) of participants reported consuming news sources from Russia. Participants were asked on a 1 - 4 scale about their trust in government institutions (M=2.99, SD=0.94) and whether they think most people can be trusted (M=2.57, SD=0.87). 10.9% (SD=0.31) of the participants work for some government organization and 10.9% (SD=0.31) use a virtual private network (vpn). Using a 1 - 5 scale, participants were asked about the frequency of using the internet (M=4.47, SD=1.8), and about the intensity of religious practices (M=3.77, SD=1.25). Weights for age and sex were applied to make the sample nationally representative.

5.4 Results

5.4.1 Average treatment effects

Figure 5.1 shows the proportion of responses in percent by treatment condition. For the first item, *participating in protests for political change* 32.3% responded *justified*, 39.1% *not justified*, and 28.6% *prefer not to answer*. In the surveillance treatment, these numbers changed by -2.51%, -1.44% and 3.95% and in the privacy treatment by 0.41%, -1.19%, and 0.78%, respectively. The second item, *helping Russia to avoid Western sanctions* has a justification rate of 29.72%, whereas 38.57% responded *not justified*, and 31.71% *prefer not to answer*. In the surveillance treatment, these numbers changed by -0.41%, -2.05% and 2.46% and in the privacy treatment by 1.36%, -0.72%, and -0.63%, respectively. The third item, whether *Russia's Special Military Operation/ invasion of Ukraine* was justified, found 25.09% of supporters, whereas 43.32% responded *not justified*, and 31.59% *prefer not to answer*. In the surveillance treatment, these numbers changed by 0.92%, -4.16% and 3.24% and in the privacy treatment by 0.79%, -1.48%, and 0.7%, respectively. Generally, self-censorship was the lowest in the question revolving around avoiding sanctions, and the highest in the question corresponding to domestic politics.

Figure 5.1: Responses to dependent variables



In order to assess whether these differences are meaningful statistical deviations, i.e., whether treatments have significant effects, multinomial logistic regression models were estimated with the three response options as unordered dependent variables, and treatment

dummies as dependent variables. Weights for age and gender were applied to make the sample representative of the population. Figure 5.2 shows the average treatment effects (ATE) resulting from these models. A treatment effect, as defined here, would cause a change in proportions in the response options between treatments. The sensitive items show treatment effects being significant at the 95% level. The surveillance treatment results in an increase in the 'prefer not to answer' option for sensitive items, by 2.46 to 3.95%. The differences are significant at the 95% level in the questions involving participating in political protests and Russia's invasion of Ukraine. Additionally, the surveillance treatment decreased the rate of the 'not justified' option by 1.44 to 4.16%, on average. This difference is significant at the 95% level in the question involving Russia's invasion of Ukraine. The average differences between control and treatment groups are modest but hint at systematic self-censorship, thus lending some support for hypothesis 1. The differences in the option 'justified' in the surveillance condition were not significant at the 95% level.

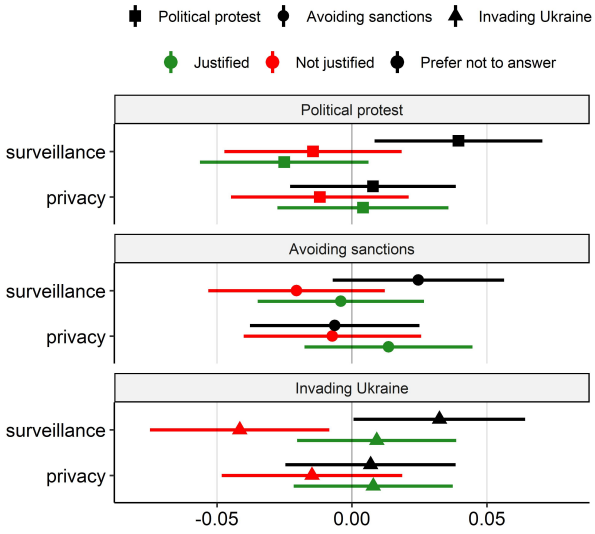
The privacy condition did not yield any statistically meaningful comparison. Noteworthy, however, is that treatment effects seem to be systematic, in that in every question the number of respondents who answered *justified* increased (between 0.41 to 1.36%) and the number of those who answered *not justified* decreased (between 0.72 to 1.48%). This study is perhaps not powered to detect effects this small (Kane 2024), hence not providing evidence in favour of the second hypothesis.

5.4.2 Conditional average treatment effects

Following the pre-registered analysis plan, conditional average treatment effects are calculated for the following moderators: age, education, trust in government, ethnicity, government employment, trust in government, income, Internet use intensity, place of residence, gender, religiosity, Russian media consumption, social trust, and VPN use. Some of these are discussed in detail here, while for readability the reader is referred to Figure A2 in the appendix for an overview. Scales were split at the median, whereas binary variables were analysed following their natural split at 1 and 0. Fully visible point estimates are significant at the 95% level, while insignificant comparisons are transparent.

First, considering results in Figure A2, and consistent with the results in 5.4.1, a compar-

Figure 5.2: Average treatment effects

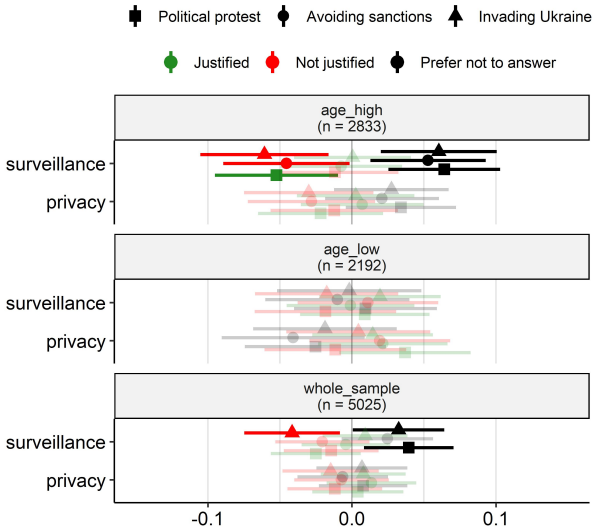


Note: Comparisons of control and treatment groups; multinomial logistic regression model point estimates with 95% level confidence intervals.

ison of the estimates between the control and privacy conditions yields only two significant comparisons (out of 98 comparisons). The confidence intervals are not larger compared to the surveillance condition, which could indicate greater uncertainty or nonlinearity in the effects. In sum, there is insufficient evidence to conclude a reduction in self-censorship or, conversely, an increase in preference disclosure due to the treatment of respondents in the privacy condition. Therefore, hypothesis 2 is not accepted.

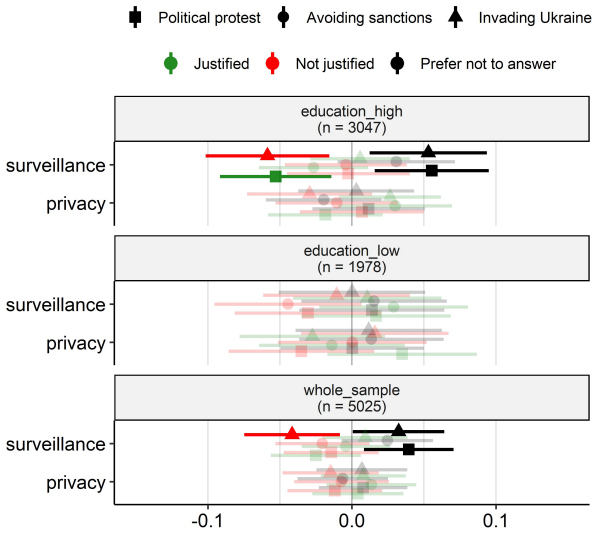
Second, across the majority of splits, the surveillance condition increased self-censorship on sensitive items relative to the control group, mostly through an increase in the proportion of *prefer not to answer* responses on sensitive items. Interestingly, this effect seems to depend on both item and individual characteristics, as different patterns emerge on both dimensions. Older participants self-censor by switching from saying *participating in a political protest* can be justified to not revealing their opinion (see Figure 5.3). Regarding the second and third items corresponding to the invasion of Ukraine, older subjects tend to hide their preferences by avoiding disagreement and not revealing their opinion. Younger participants do not seem to be sensitive to the surveillance treatment at all, as reflected by very small estimates and insignificant comparisons. This finding challenges previous empirical studies measuring self-

Figure 5.3: Conditional treatment effects: age



censorship in autocracies finding self-censorship to be greater among younger respondents (Robinson and Tannenberg 2019) or, less dramatically, suggests different effects depending on the context.

Figure 5.4: Conditional treatment effects: education



This pattern is repeated for more educated individuals, as they when compared to those with less education, exert a higher level of self-censorship. Precisely, the proportions of respondents answering *prefer not to answer* to the sensitive questions in the whole population

of 3.9%, 2.5%, and 3.2% increase to 5.6%, 3.1%, and 5.3%, respectively, for the subgroup of more educated respondents.

Figure 5.5: Conditional treatment effects: place of residence

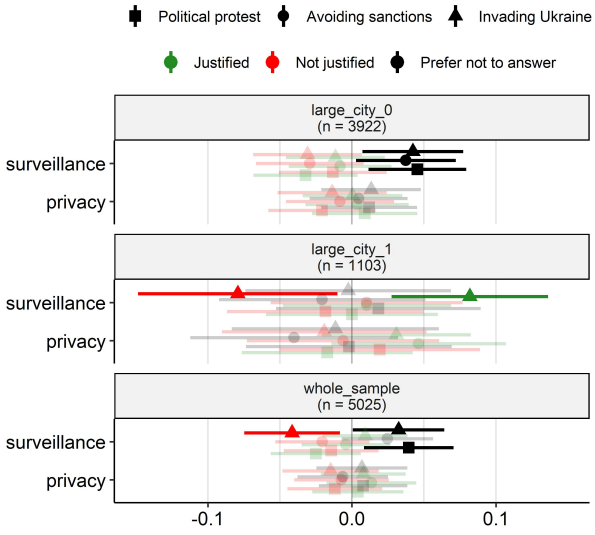
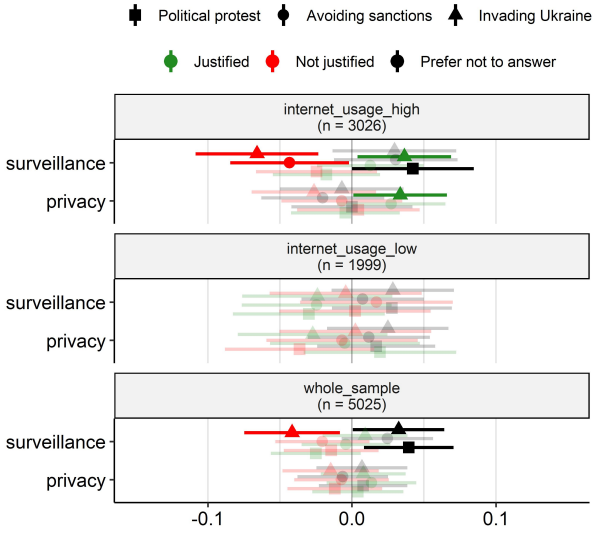


Figure 5.5 shows comparisons where the sample is split in respondents residing in Almaty or Astana (*large_city_1*), or not. In the rural population (not living in either metropolitan area), an increase in the *prefer not to answer* response was found, not in stating an opinion to this question. On the contrary, respondents from either of the two large cities – when comparing the surveillance to the control group – responded *justified* more often (8.2%) and *not justified* less often (-7.9%) to the question of whether invading Ukraine was justified. This behavior is unique in that it was only found for the subgroups of large city residents and high-intensity internet users (see Figure 5.6),[6] and for the item corresponding to Russia's invasion of Ukraine. Similarly large effects were found for respondents who consumed media sources where they indicated Russia as the country of origin, see Figure 5.7. In the surveillance condition, respondents increased the proportion of *prefer not to answer* by between 8.5 and 8.9% when they also consumed media sources from Russia. Conversely, those who did not consume any media sources from Russia chose *prefer not to answer* between -0.04% and 1.9% less or more often, without statistical significance. This finding raises the question of what

[6]There is a large overlap between the two subgroups: 847 out of 1103 respondents (77%) in the large city group were also high-intensity internet users.

it is that makes respondents who choose to consume Russian media also self-censor strongly. Or, perhaps, what it is about consuming Russian media that makes people self-censor more intensively. Other studies found strong effects of domestic propaganda consumption on approval rates of surveillance technology in China (Xu et al. 2022) and Russia (Karpa and Rochlitz 2023).[7] This study adds to this by finding what appears to be a 'spill-over' effect of Russian propaganda on the Kazakh population.[8]

Figure 5.6: Conditional treatment effects: internet use

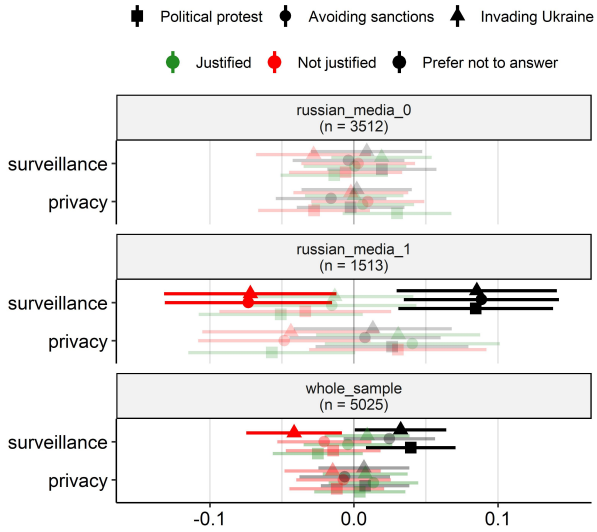


When focusing on the second item, whether to *help Russia avoid Western sanctions*, results show that older subjects tend to hide their preferences by avoiding to disagree and instead not revealing their opinion, which seems to also be the case for heavy internet users, those exhibiting a lower trust in the government, and highly educated subjects. The same pattern emerges for the third item, whether the *Russian invasion of Ukraine/SMO* was justified. Older, more educated, and those with low social trust hide their preferences by avoiding to disagree and instead not revealing their opinion. For those who earn better, heavy internet users and those with high social trust the same pattern of response shifts emerge,

[7] All of these studies including this one, however, provide only correlational evidence.

[8] Out of 1513 respondents who indicated consuming Russia media, 939 are ethnically Kazakh, 494 Russian, and 80 preferred not to reply to this question.

Figure 5.7: Conditional treatment effects: Russian media consumption

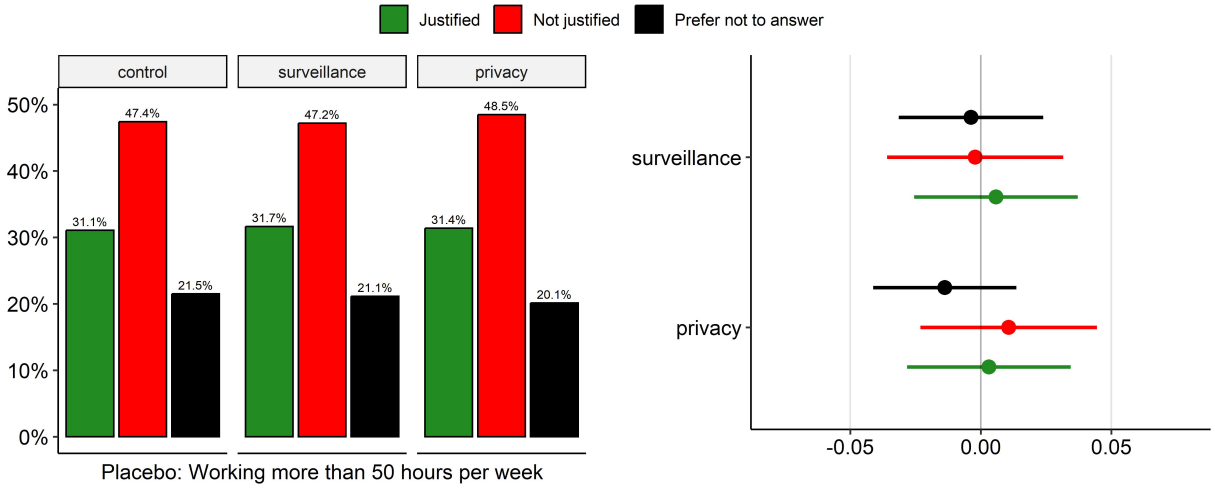


albeit appearing less often with statistically significant estimates. In general, the surveillance treatment made respondents significantly less likely to say *not justified* when asked whether Russia's invasion of Ukraine was justified. The only two subgroups that increased their approval of the invasion in the surveillance condition were respondents employed by the government (3.6%) and ethnically Russian respondents (0.2%), both of which effects are statistically insignificant. Other studies found that people are less likely to *say* they would like to trade civil liberties for security when they are disadvantaged compared to peers (Davis and Silver 2004; Dietrich and Crabtree 2019; Alsan et al. 2023). In this study, those with lower income *act* close to average, whereas those with lower education seem unresponsive to the surveillance treatment.

In summary, the main results correspond to: (1) the surveillance treatment mostly leads those to who agree or disagree with an item to instead chose *prefer not to answer*, and (2) far less often it induces preference falsification by participants (dis)agreeing with the opposite of the stated opinion in the control group. Furthermore, (3) the privacy treatment had small and thus undetectable effects. Finally, (4) effect sizes dramatically increase for some subgroups (up to three times the size) or diminish entirely, clearly indicating strong heterogeneity in treatment effects.

5.4.3 Robustness tests and baseline self-censorship

Figure 5.8: Placebo Item

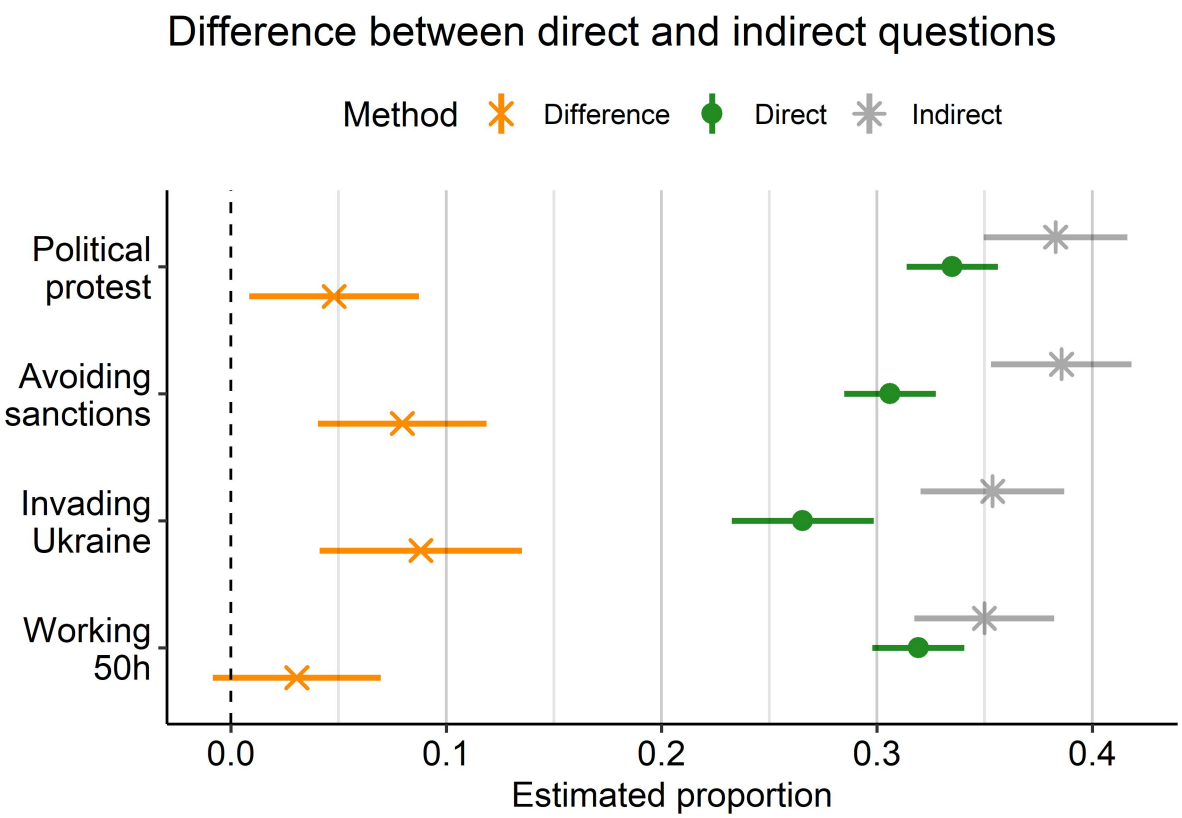


Note: Left: Responses to dependent variable by treatment condition. Right: Comparisons of control and treatment groups; multinomial model point estimates with 95% level confidence intervals.

In order to assess the quality of effects described above different robustness checks were performed. First, Figure 5.8 shows the results for a placebo item (*In your opinion, is working more than 50 hours per week generally justified?*) that was run beside the three sensitive items. First, the rate of non-responses, i.e. subjects who chose "prefer not to answer", is about 30% for the sensitive items (figure 5.1), while it is about 20% for the placebo item. On average, preference falsification is higher in the politically sensitive items, which lends support for the assumption that participants perceived the questions as sensitive and adapted accordingly. Second, the estimates for treatment effects in both conditions for the placebo question are very small, do not show systematic variation, and are insignificant at the 95% level. This serves as an indication that the study does not suffer from this specific design effect, which would correspond to self-censorship regardless of the sensitivity of the item.

How does the context in which this study was conducted affect its results? When asking about politically sensitive issues – e.g., whether political protests are justified – the context itself is a powerful treatment. Other studies have shown that there is substantial self-censorship on such questions in autocracies, with up to a quarter of respondents falsify-

Figure 5.9: Baseline self-censorship



Note: The shape of the point estimate indicates whether the response was given to an indirect question (list experiment) or whether it was given to a direct question with the options *Justified/Not justified/Prefer not to answer* (logistic regression on "Justified"). Confidence intervals are obtained using Monte Carlo simulations and are given at the 95% level. Fitted values were obtained using control variables including region fixed-effects, following Blair and Imai (2012). For more methodological details on the list experiment, see Section 5.A.4

ing their preferences when asked directly (Robinson and Tannenberg 2019). Arguably, the treatment effects in this study's experiment are conservative because participants are *already* treated, simply by using the Kazakh Internet. Thus, what was measured in 5.4.1 and 5.4.2 are the *additional* treatment effects, and these should be smaller, assuming a diminishing marginal effect.[9]

To account for this, direct and indirect questioning techniques are combined and a baseline of self-censorship is calculated and shown in Figure 5.9. The differences between the

[9]While the assumption of diminishing marginal effects seems reasonable, it cannot be ruled out that participants become unresponsive to an additional treatment or that there are increasing marginal effects, nor would it make sense to assume homogeneous effects across different demographic groups.

direct and indirect questions are, from top to bottom, 4.8%, 8.0%, 8.8% and 3.0%, with the difference being significant at the 95% level for the sensitive items, but not for the placebo item, again supporting the assumption of no design effects. In the indirect question, non-response was both not possible and not strategically necessary, as the indirect question does not compromise the respondent's privacy. Similar to the treatment effects in Figure 5.8, there is no also no *baseline* self-censorship in the placebo condition. The differences between the direct and indirect questions thus include the amount of respondents who hid their preferences behind *prefer not to answer* and, arguably about 3% of participants who chose the option because it was less cognitively demanding. While seemingly modest compared to other studies (Robinson and Tannenberg 2019), these numbers increased when respondents were treated with the surveillance condition and, notably, did not decrease when participants were treated with the privacy condition. In other words, while people tend to increase their self-censorship in the face of salient surveillance practices, no decrease in existing self-censorship in the face of encryption technology could be found. This study thus provides evidence for the theoretical prediction of Büchi et al. (2022), which suggests an erosion of digital communication behavior over time, with an increasing aggregate chilling effect and imperfect recovery. More specifically, it was shown that the potential for *immediate* recovery is very low, if not non-existent, and that the only recovery possible is one in which the salience of surveillance practices declines over time.

An alternative interpretation is that, since the loss of privacy reduces communication behavior much more than the gain of privacy increases it, seemingly, citizens are loss averse concerning privacy. In other words, losses of privacy affect citizens more than gains in privacy do, as expressed in their behavioral adaptations. Assuming symmetry in the strength of the experimental treatments, this asymmetry in measured effects suggests asymmetric preferences, corresponding to what is known as loss-aversion (Schmidt and Zank 2005). This also means that – for policies that aim to enhance the political discourse – privacy-preserving technologies are no solution for increasing surveillance capabilities, first because they are costly and access is unequally distributed, and second because they are simply not as effective – because of the aforementioned loss-aversion.

5.5 Concluding discussion

This study contributes to the literature on digital authoritarianism by showing how surveillance reduces digital communication behavior. Self-censoring citizens do not express their opinions on political issues, which contributes to the chilling of political discussions and the further depoliticization of individuals, or in other words, to the stabilization of the hegemonic power of the state over public opinion. Without knowledge of peers' preferences on political issues, political opposition to incumbents has difficulty organizing, a key reason why autocrats resort to censorship (King et al. 2017). New surveillance technologies can thus directly bolster the autocrat's power before unrest forms, which in turn can be suppressed through the use of facial recognition surveillance technology (Beraja et al. 2023b).

Previous research has focused on the acceptance of new (surveillance) technologies (Kostka 2019; Kostka and Antoine 2020; Kostka et al. 2021; Kostka and Habich-Sobiegalla 2022; Xu et al. 2022; Kalmus et al. 2022; Karpa and Rochlitz 2023; Kostka et al. 2023), measurements of *opinion* towards surveillance (Davis and Silver 2004; Dietrich and Crabtree 2019; Alsan et al. 2023), or behavioral *intentions* in order to cope with surveillance (Stoycheff 2016; Stoycheff et al. 2019; Stoycheff 2022; Büchi et al. 2022; Xu 2022). The correlation between *approval* or *intentions* towards a specific technology and *behavioral adaptations* because of this exact technology might not be linear nor homogeneous. More specifically, approval or tolerance for state surveillance does not singularly translate into no self-censorship, or conversely, high self-censorship. In China, there are exceptionally high approval rates of state surveillance (Su et al. 2022), while there are also high rates of self-censorship Robinson and Tannenberg (2019). In Kazakhstan, the approval towards state surveillance is much lower,[10] and self-censorship rates are also smaller, yet substantial. Furthermore, in this study, the average treatment effect was driven by older respondents and previous research on mass surveillance found a high tolerance among the older population, when it comes to state surveillance in post-soviet countries (Kalmus et al. 2022). It appears as if approving or tolerating state surveillance might be a coping mechanism to deal with the cognitive and emotional stress

[10] 31.6% of Kazakh people say the government should definitely or probably have the right to monitor all emails and any other information exchanged on the Internet, whereas this number is 60.6% in China. Source: World value survey wave 7.

of surveillance, an argument also suggested in the context of China (Ollier-Malaterre 2023). As Ollier-Malaterre (2023) documents, living with digital surveillance intertwines cultural, psycho-social, and economic factors, resulting in multifaceted behavior not free of contradictions.

The complexity of behavioral adaptations concerning self-censorship is reflected by finding *directed* and *undirected* self-censorship in this study. Directed self-censorship corresponds to falsifying preferences by giving a specific answer where socially desirable behavior is known. Conversely, undirected self-censorship relates to denying to state any opinion where socially desirable behavior is not known or not deductible. This can result from a lack of information or a lack of political literacy to evaluate relevant information. Citizens who are intensively using the internet showed directed self-censorship, whereas generally undirected self-censorship in the form of non-responses was found. This finding begs the question of how respondents inferred that they *should* answer that the invasion was justified and answered accordingly in the surveillance condition. Those with a higher degree of education also resort to *preferring not to answer*, while those characterized by a lower level of education tend to behave unresponsive under the effect of the treatments, on average. Assuming that political literacy correlates with higher levels of education, it seems that it is the informational difference rather than a difference in political literacy that matters when it comes to inferring socially desirable behavior. Thus, frequent use of the Internet and living in one of Kazakhstan's two large cities produce an informational difference relative to others that translates into different behavioral responses. Others found higher self-censorship among urban citizens (Robinson and Tannenberg 2019), a finding that is complemented here by showing a more nuanced differentiation in self-censorship patterns. In other words, the complexity of response options as well as the complexity of questions does seem to matter. When socially desired behavior is not easily deductible from the context, responses mirror this complexity in that they tend to be non-responses (i.e., *prefer not to answer*), instead of what is believed to be the correct response. In a similar vein, others have argued that chilling effects induced by surveillance "can best be understood as an act that conforms to, or is in compliance with, social norms in that context" (Penney 2022, p.1520).

Finally, there are cognitive components behind behavioural adaptations that remain

opaque to the design of this study. The present study identified average behavioral responses and further investigated which groups are more sensitive to self-censoring as a behavioral response, but by design neglected an investigation of cognitive mechanisms. There are different promising offers in the literature providing avenues for further research; the economics of privacy literature suggests the involvement of an evolutionary 'sense' of privacy related to congenital processes of impression management (Acquisti et al. 2022), or, the literature on chilling effects of dataveillance, which suggests including "dataveillance imaginaries", i.e., the cognitive understanding of humans subject to (data) surveillance processes, which substantially shape behavioral responses (Kappeler et al. 2023). If anything, this study has contributed to shed light on the necessity for qualitative studies or mixed-method designs that complement and enhance results of quantitative studies.

5.6 Bibliography

Acquisti, A., Brandimarte, L., and Hancock, J. (2022). How privacy's past may shape its future. *Science*, 375(6578):270–272.

Agerberg, M. and Tannenberg, M. (2021). Dealing with measurement error in list experiments: Choosing the right control list design. *Research & Politics*, 8(2):20531680211013154.

Ahlquist, J. S. (2018). List Experiment Design, Non-Strategic Respondent Error, and Item Count Technique Estimators. *Political Analysis*, 26(1):34–53.

Alsan, M., Braghieri, L., Eichmeyer, S., Kim, M. J., Stantcheva, S., and Yang, D. Y. (2023). Civil liberties in times of crisis. *American Economic Journal: Applied Economics*, 15(4):389–421.

Bentham, J. (2011). *The Panopticon Writings*. Radical Thinkers. Verso Books.

Beraja, M., Kao, A., Yang, D., and Yuchtman, N. (2023a). Exporting the Surveillance State via Trade in AI. *NBER Working Paper*.

Beraja, M., Kao, A., Yang, D. Y., and Yuchtman, N. (2023b). AI-tocracy. *The Quarterly Journal of Economics*, 138(3):1349–1402.

Beraja, M., Yang, D. Y., and Yuchtman, N. (2023c). Data-intensive innovation and the state: evidence from ai firms in china. *The Review of Economic Studies*, 90(4):1701–1723.

Blair, G., Coppock, A., and Moor, M. (2020). When to Worry About Sensitivity Bias: A Social Reference Theory and Evidence From 30 Years of List Experiments. *American Political Science Review*, 114(4):1297–1315.

Blair, G. and Imai, K. (2012). Statistical Analysis of List Experiments. *Political Analysis*, 20(1):47–77.

Blair, G., Imai, K., and Lyall, J. (2014). Comparing and Combining List and Endorsement Experiments: Evidence From Afghanistan. *American Journal of Political Science*, 58(4):1043–1063.

Büchi, M., Festic, N., and Latzer, M. (2022). The Chilling Effects of Digital Dataveillance: A Theoretical Model and an Empirical Research Agenda. *Big Data & Society*, 9(1):205395172110653.

Carothers, T. and Press, B. (2022). Understanding and responding to global democratic backsliding. *CEIP: Carnegie Endowment for International Peace*.

Coffman, K. B., Coffman, L. C., and Ericson, K. M. M. (2013). The Size of the LGBT Population and the Magnitude of Anti-Gay Sentiment are Substantially Underestimated. Working Paper 19508, National Bureau of Economic Research.

Corstange, D. (2012). Vote Trafficking in Lebanon. *International Journal of Middle East Studies*, 44(3):483–505.

Davis, D. and Silver, B. (2004). Civil Liberties vs. Security: Public Opinion in the Context of the Terrorist Attacks on America. *American Journal of Political Science*, 48(1):28–46.

Dietrich, N. and Crabtree, C. (2019). Domestic Demand for Human Rights: Free Speech and the Freedom-Security Trade-Off. *International Studies Quarterly*, 63(2):346–353.

Earl, J., Maher, T., and Pan, J. (2022). The Digital Repression of Social Movements, Protest, and Activism: A Synthetic Review. *Science advances*, 8.

Edmond, C. (2013). Information Manipulation, Coordination, and Regime Change. *The Review of Economic Studies*, 80(4 (285)):1422–1458.

Egorov, G. and Sonin, K. (2024). The political economics of non-democracy. *Journal of Economic Literature*, 62(2):594–636.

Feldstein, S. (2019a). The Global Expansion of AI Surveillance. *Carnegie Endowment for International Peace*.

Feldstein, S. (2019b). The Road to Digital Unfreedom: How Artificial Intelligence Is Reshaping Repression. *Journal of Democracy*, 30(1):40–52.

Feldstein, S. (2021). *The Rise of Digital Repression: How Technology Is Reshaping Power, Politics, and Resistance*. Oxford: Oxford University Press.

Festic, N. (2022). Same, same, but different! qualitative evidence on how algorithmic selection applications govern different life domains. *Regulation & Governance*, 16(1):85–101.

Foucault, M. (2012). *Discipline and Punish: The Birth of the Prison*. Vintage. Knopf Doubleday Publishing Group.

FreedomHouse (2023a). Kazakhstan: Freedom in the world 2023 country report.

FreedomHouse (2023b). Kazakhstan: Freedom on the net 2023 country report.

Frye, T., Gehlbach, S., Marquardt, K. L., and Reuter, O. J. (2017). Is Putin's Popularity Real? *Post-Soviet Affairs*, 33(1):1–15.

Frye, T., Gehlbach, S., Marquardt, K. L., and Reuter, O. J. (2023). Is Putin's Popularity (Still) Real? A Cautionary Note on Using List Experiments to Measure Popularity in Authoritarian Regimes. *Post-Soviet Affairs*, pages 1–10.

Glynn, A. N. (2013). What Can We Learn with Statistical Truth Serum?: Design and Analysis of the List Experiment. *Public Opinion Quarterly*, 77(S1):159–172.

Gohdes, A. (2023). *Repression in the Digital Age: Surveillance, Censorship, and the Dynamics of State Violence*. Disruptive Technology and International Security Series. Oxford University Press, Incorporated.

Guriev, S. and Treisman, D. (2019). Informational Autocrats. *Journal of Economic Perspectives*, 33(4):100–127.

Guriev, S. and Treisman, D. (2020). A Theory of Informational Autocracy. *Journal of Public Economics*, 186(104158).

Hobbs, W. R. and Roberts, M. E. (2018). How sudden censorship can increase access to information. *American Political Science Review*, 112(3):621–636.

Huang, G., Hu, A., and Chen, W. (2022). Privacy at Risk? Understanding the Perceived Privacy Protection of Health Code Apps in China. *Big Data & Society*, 9(2).

Imai, K. (2011). Multivariate regression analysis for the item count technique. *Journal of the American Statistical Association*, 106(494):407–416.

Jones, M. (2022). *Digital Authoritarianism in the Middle East: Deception, Disinformation and Social Media*. Hurst.

Kalmus, V., Bolin, G., and Figueiras, R. (2022). Who is Afraid of Dataveillance? Attitudes Toward Online Surveillance in a Cross-cultural and Generational Perspective. *New Media & Society*.

Kane, J. V. (2024). More than meets the itt: A guide for anticipating and investigating nonsignificant results in survey experiments. *Journal of Experimental Political Science*, page 1–16.

Kappeler, K., Festic, N., and Latzer, M. (2023). Dataveillance imaginaries and their role in chilling effects online. *International Journal of Human-Computer Studies*, 179:103120.

Karpa, D. and Rochlitz, M. (2023). Authoritarian Surveillance and Public Support for Digital Governance Solutions. *SSRN*, 4454627.

King, G., Pan, J., and Roberts, M. E. (2017). How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, not Engaged Argument. *American Political Science Review*, 111(3):484–501.

Kostka, G. (2019). China's Social Credit Systems And Public Opinion: Explaining High Levels of Approval. *New Media & Society*, 21(7):1565–1593.

Kostka, G. and Antoine, L. (2020). Fostering Model Citizenship: Behavioral Responses to China's Emerging Social Credit Systems. *Policy & Internet*, 12(3):256–289.

Kostka, G. and Habich-Sobiegalla, S. (2022). In Times of Crisis: Public Perceptions Toward COVID-19 Contact Tracing Apps in China, Germany, and the United States. *New Media & Society*, 0(0).

Kostka, G., Steinacker, L., and Meckel, M. (2021). Between Security and Convenience: Facial Recognition Technology in the eyes of citizens in China, Germany, the United Kingdom, and the United States. *Public Understanding of Science*, 30(6):671–690.

Kostka, G., Steinacker, L., and Meckel, M. (2023). Under Big Brother's Watchful Eye: Cross-country Attitudes Toward Facial Recognition Technology. *Government Information Quarterly*, 40(1):101761.

Kramon, E. and Weghorst, K. (2019). (Mis)Measuring Sensitive Attitudes with the List Experiment: Solutions to List Experiment Breakdown in Kenya. *Public Opinion Quarterly*, 83(S1):236–263.

Kudaibergenova, D. T. and Laruelle, M. (2022). Making sense of the january 2022 protests in kazakhstan: failing legitimacy, culture of protests, and elite readjustments. *Post-Soviet Affairs*, 38(6):441–459.

Kuhn, P. M. and Vivyan, N. (2022). The Misreporting Trade-Off Between List Experiments and Direct Questions in Practice: Partition Validation Evidence from Two Countries. *Political Analysis*, 30(3):381–402.

Lee, H. J. (2023). 'i've left enough data': Relations between people and data and the production of surveillance. *Big Data & Society*, 10(1):20539517231173904.

Liu, K. Z. (2019). Commercial-state Empire: A Political Economy Perspective on Social Surveillance in Contemporary China. *The Political Economy of Communication*, 7(1).

Manokha, I. (2018). Surveillance, Panopticism, and Self-Discipline in the Digital Age. *Surveillance & Society*, 16:219–237.

Ollier-Malaterre, A. (2023). *Living with Digital Surveillance in China: Citizens' Narratives on Technology, Privacy, and Governance*. Routledge studies in surveillance. Routledge.

Oz, M. and Yanik, A. (2022). Fear of surveillance: Examining turkish social media users' perception of surveillance and willingness to express opinions on social media. *Mediterranean Politics*, 0(0):1–25.

Pan, J. and Siegel, A. A. (2020). How saudi crackdowns fail to silence online dissent. *American Political Science Review*, 114(1):109–125.

Penney, J. W. (2016). Chilling Effects: Online Surveillance and Wikipedia Use. *Berkeley Technology Law Journal*, 31(1):117–182.

Penney, J. W. (2017). Internet Surveillance, Regulation, and Chilling Effects Online: a Comparative Case Study. *Internet Policy Review*, 2(6).

Penney, J. W. (2022). Understanding chilling effects. *Minnesota Law Review*, 106:1451.

Qin, B., Strömberg, D., and Wu, Y. (2017). Why Does China Allow Freer Social Media? Protests versus Surveillance and Propaganda. *Journal of Economic Perspectives*, 31(1):117–40.

Raman, R. S., Evdokimov, L., Wurstrow, E., Halderman, J. A., and Ensafi, R. (2020). Investigating large scale https interception in kazakhstan. In *Proceedings of the ACM Internet Measurement Conference*, IMC '20, page 125–132, New York, NY, USA. Association for Computing Machinery.

Roberts, M. E. (2018). *Censored: Distraction and Diversion Inside China's Great Firewall*. Princeton: Princeton University Press.

Roberts, M. E. (2020). Resilience to Online Censorship. *Annual Review of Political Science*, 23(1):401–419.

Robinson, D. and Tannenberg, M. (2019). Self-censorship of Regime Support in Authoritarian States: Evidence From List Experiments in China. *Research & Politics*, 6(3):2053168019856449.

Rosenfeld, B., Imai, K., and Shapiro, J. N. (2016). An Empirical Validation Study of Popular Survey Methodologies for Sensitive Questions. *American Journal of Political Science*, 60(3):783–802.

Schauer, F. (1978). Fear, Risk, and the First Amendment: Unraveling the "Chilling Effect.". *Boston University Law Review*, 58(0):685—732.

Schmidt, U. and Zank, H. (2005). What is loss aversion? *Journal of Risk and Uncertainty*, 30(2):157–167.

Silvan, K. (2024). Three Levels of Authoritarian Legitimacy: Successor Designation and Peaceful to Non-Peaceful Leadership Transition in Kazakhstan. *Communist and Post-Communist Studies*, pages 1–23.

Stoycheff, E. (2016). Under Surveillance. *Journalism & Mass Communication Quarterly*, 93(2):296–311.

Stoycheff, E. (2022). Cookies and Content Moderation: Affective Chilling Effects of Internet Surveillance and Censorship. *Journal of Information Technology & Politics*, pages 1–12.

Stoycheff, E., Liu, J., Xu, K., and Wibowo, K. (2019). Privacy and the Panopticon: Online Mass Surveillance's Deterrence and Chilling Effects. *New Media & Society*, 21(3):602–619.

Su, Z., Xu, X., and Cao, X. (2022). What Explains Popular Support for Government Monitoring in China? *Journal of Information Technology & Politics*, 19(4):377–392.

Tannenberg, M. (2022). The Autocratic Bias: Self-censorship of Regime Support. *Democratization*, 29(4):591–610.

Xu, X. (2021). To Repress or to Co-opt? Authoritarian Control in the Age of Digital Surveillance. *American Journal of Political Science*, 65(2):309–325.

Xu, X. (2022). The Unintrusive Nature of Digital Surveillance and Its Social Consequences. *Working Paper*.

Xu, X., Kostka, G., and Cao, X. (2022). Information Control and Public Support for Social Credit Systems in China. *Journal of Politics*, 84(2).

5.A Appendix

5.A.1 Treatment design

control	privacy	surveillance
In the next section, you will be asked your opinion on economic and political issues directly. Your answers will remain confidential.	In the next section, you will be asked your opinion on economic and political issues directly. Your answers will remain confidential. Our encryption mechanisms make it completely impossible to track your data.	In the next section, you will be asked your opinion on economic and political issues directly. Your answers will remain confidential. However, as you may be aware, the government of Kazakhstan may access information about your online activity directly from your Internet Service Provider.

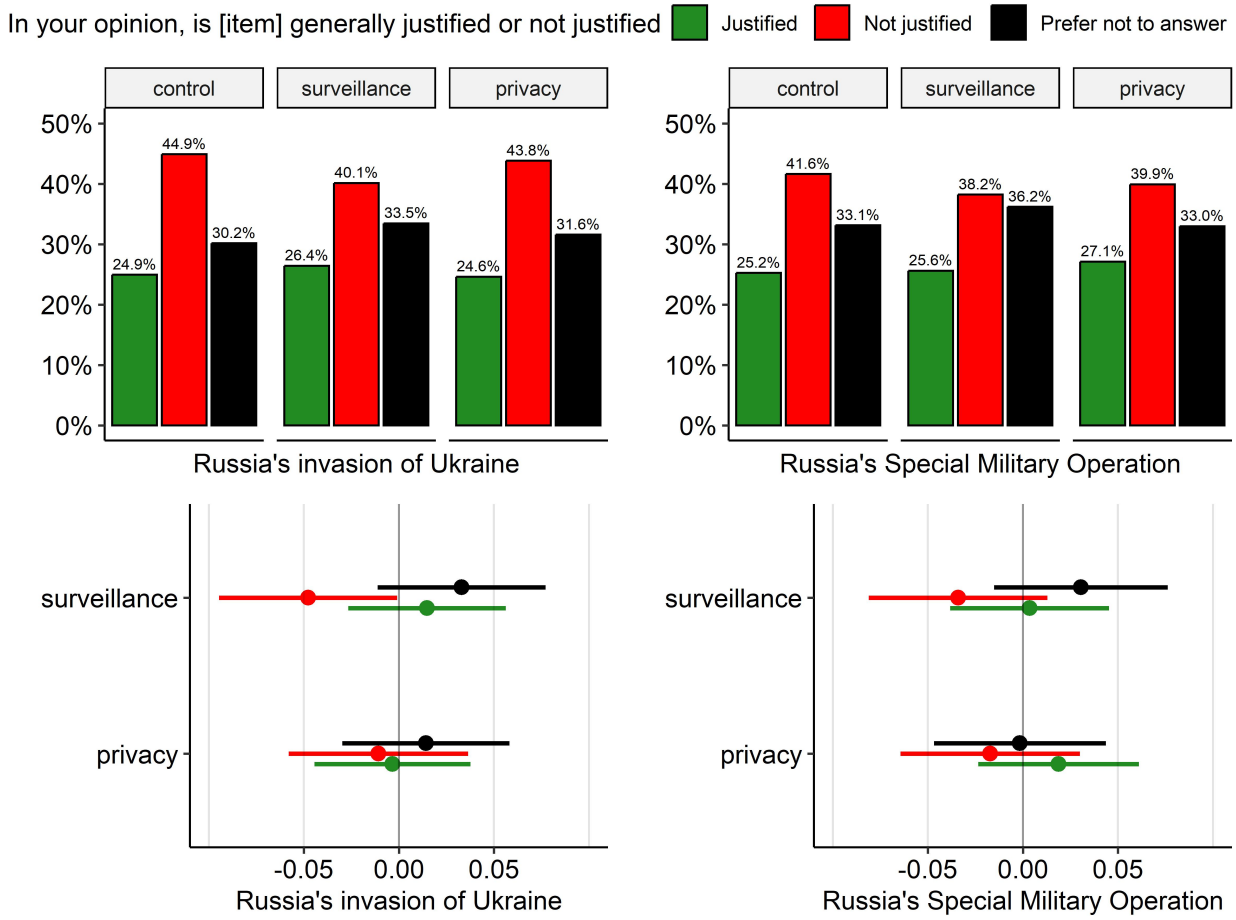
5.A.2 Additional Tables

Table A1: Summary statistics

Variable	N	Mean	SD
age	5025	42.7	16.1
education_scale	5025	4.55	1.75
ethnicity_kazakh	5025	0.722	0.448
government_employee	5025	0.109	0.312
government_trust_scale	4340	2.99	0.937
financial_situation_scale	5025	2.85	1.14
internet_scale	5025	4.47	1.8
large_city	5025	0.22	0.414
male	5025	0.487	0.5
religiosity_scale	3816	3.77	1.25
russian_media_consumption	5025	0.301	0.459
social_trust_scale	4485	2.57	0.87
vpn_user	5025	0.109	0.312
framing_invasion			
... invasion	5025	0.502	0.5
... SMO	5025	0.498	0.5

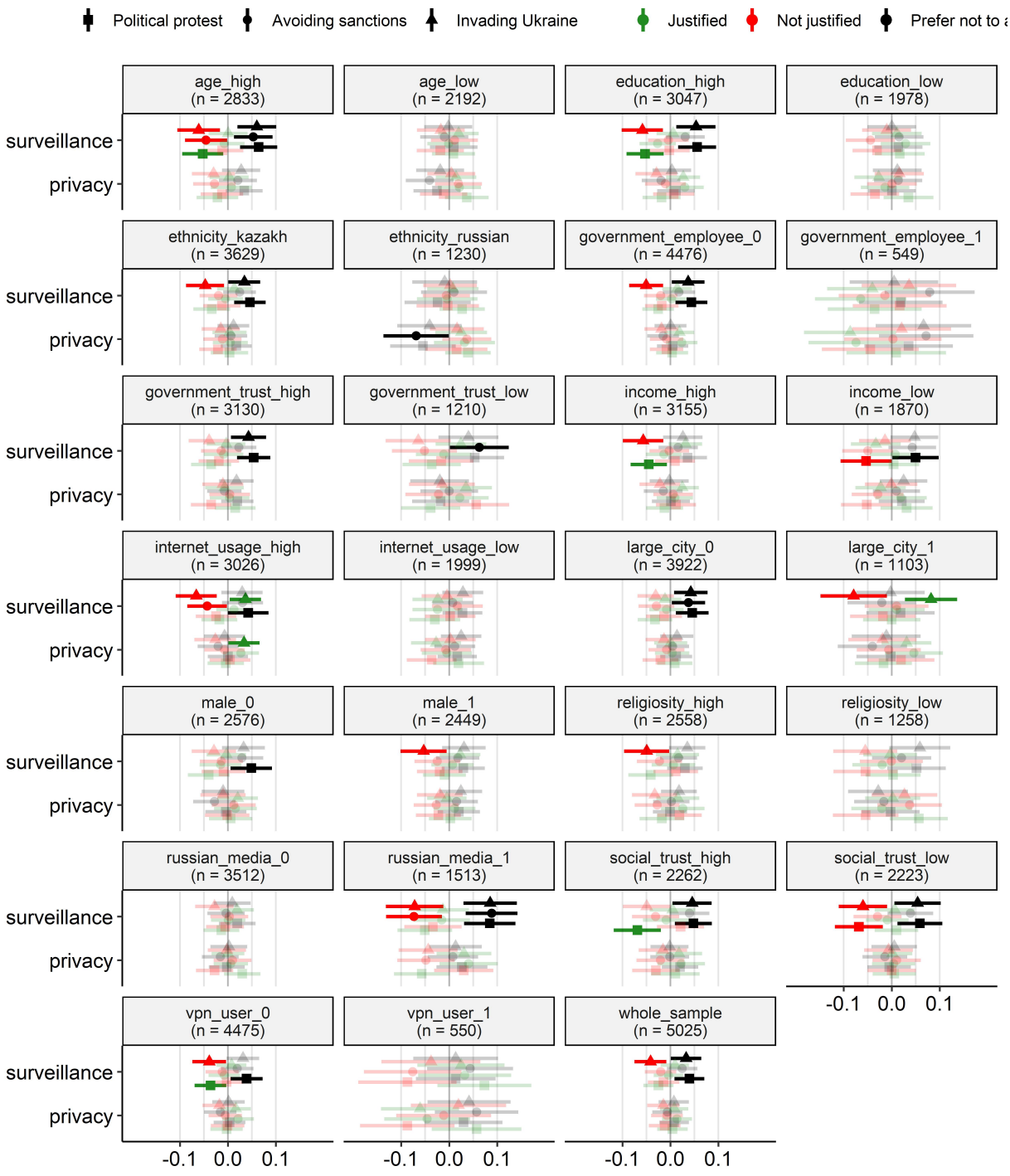
5.A.3 Additional Figures

Figure A1: Differences between framings of the invasion of Ukraine



Note: Top row: Responses to dependent variable by treatment condition. Bottom row: Comparisons of control and treatment groups; multinomial model point estimates with 95% level confidence intervals.

Figure A2: Conditional average treatment effects



Note: The shape of the points indicates the item posing as the dependent variable of the multinomial model, while the colors indicate the response options. Fully visible point estimates with confidence intervals are statistically significant different from the control condition at the 95% level, while transparent coefficients are not. Splits are made at the median or between 1 and 0 for binary variables.

5.A.4 List experiment design

List experiments – also known as the item count technique – have been successfully used, for example, to study support for authoritarian leaders (Blair et al. 2014; Frye et al. 2017; Robinson and Tannenberg 2019; Frye et al. 2023), estimating the size of LGBT population (Coffman et al. 2013), and vote trafficking in Lebanon (Corstange 2012).

Participants are exposed to either J or $J + 1$ items and then asked to count the number of items that apply to them, with the additional (J + 1th) item being the sensitive item of interest.[11] The premise of list experiments is that when a sensitive question is asked indirectly, respondents are more likely to give a truthful answer, even if social norms encourage them to answer the question in a particular way (Blair and Imai 2012). Fear of being judged or punished by others leads to a change in behavior best known as social desirability bias, a subset of what is known as sensitivity bias (Blair et al. 2020). Using list experiments as well as direct questions, the difference between honest and self-censored responses can be estimated (Robinson and Tannenberg 2019). In other words, comparing the means of direct and indirect responses can reveal strategic misreporting. More sophisticated statistical methods allow analysis beyond mean comparisons so that sensitivity bias can be shown, but also which sociodemographic factors and personality traits play a role (Imai 2011; Blair and Imai 2012).

In list experiments, there is a trade-off between being able to circumvent or measure *strategic* misreporting, and inevitable *nonstrategic* misreporting (Kuhn and Vivyan 2022). Nonstrategic measurement errors are, for example, floor and ceiling effects (Ahlquist 2018), and complexity (Kramon and Weghorst 2019). Floor and ceiling effects appear when either all or none of the statements in the list are true, and survey participants thus must reveal their preference over the sensitive item in order to remain giving truthful answers. In this case, one would have to expect strategic misreporting, i.e., lying. List experiments thus help to elicit true preferences, but, by design, suffer from overwhelming or confusing study participants. By careful consideration of the design, these measurement errors can be minimized (Glynn 2013; Blair et al. 2020), for example by choosing a suitable control list (Agerberg

[11] J is commonly equal to 3 or 4. Higher values substantially increase complexity, while lower values are risking to compromise the privacy of the participants (floor effect).

and Tannenberg 2021). Is it furthermore common practice to validate experiments before going into the field (Rosenfeld et al. 2016).

This study's list experiment takes this literature into account and is designed accordingly. The experiment tries to measure self-censorship by comparing the means of direct and indirect questioning, with respect to three different questions on politically sensitive items and one placebo item, which can be found below in Table A2. Non-strategic measurement errors were tried to be minimized according to suggestions from the literature:

First, floor and ceiling effects were avoided by (1) choosing items where the prevalence among the population was known beforehand, and (2) choosing control list items that are negatively correlated. For example, the items "homosexuality" and "holding on to religious values" are strongly negatively correlated. It is very unlikely that respondents agree with both items. Furthermore, it is unlikely that none of the items are true, statistically. Items were taken and adapted from the World Value Survey and Central Asian Barometer. Thus, ceiling effects and floor effects could, by design, be minimized. Second, sensitive control list items were chosen, in order to conceal the item of interest and not evoke demand effects. Third, a placebo item is used in order to test whether there are design effects (Frye et al. 2017, 2023). Fourth, a pre-test was run with 400 subjects in order to validate the lists. Last, a double list experiment was employed in order to enhance statistical power. In a double list experiment the sample is split into two groups, and each half acts as a control group for the other half, see Table A2.

Table A2: List experiment design

In the next section, you will be presented with a number of items. Please read the list of items carefully. After you have read the items, please indicate HOW MANY of the items (things or actions) below are generally justified in your opinion. We will not be able to know which items are justified for you.	
[randomization into one of the two columns below with equal probability]	
In your opinion, HOW MANY of the things or actions below are generally justified? [selection, randomization of list item order] Prioritizing industrial development above environmental safety Protecting the environment even if costs are high Prostitution Participating in protests for political change 0 1 2 3 4	In your opinion, HOW MANY of the things or actions below are generally justified? [selection, randomization of list item order] Prioritizing industrial development above environmental safety Protecting the environment even if costs are high Prostitution 0 1 2 3
In your opinion, HOW MANY of the things or actions below are generally justified? [selection, randomization of list item order] Homosexuality Helping Russia avoid Western sanctions Full time work for women Holding on to religious values 0 1 2 3 4	In your opinion, HOW MANY of the things or actions below are generally justified? [selection, randomization of list item order] Homosexuality Full time work for women Holding on to religious values 0 1 2 3
In your opinion, HOW MANY of the things or actions below are generally justified? [selection, randomization of list item order] Suicide Being proud of national traditions Aspiring to Western values 0 1 2 3	In your opinion, HOW MANY of the things or actions below are generally justified? [selection, randomization of list item order, randomization of framing: SMO or invasion - variable framing34] Suicide Being proud of national traditions Russia's Special Military Operation in Ukraine/ Russia's invasion of Ukraine Aspiring to Western values 0 1 2 3 4
In your opinion, HOW MANY of the things or actions below are generally justified? [selection, randomization of list item order] Death penalty Violating traffic rules Banning smoking in public places 0 1 2 3	In your opinion, HOW MANY of the things or actions below are generally justified? [selection, randomization of list item order, randomization of framing: SMO or invasion - variable framing34] Death penalty Violating traffic rules Working more than 50 hours per week Banning smoking in public places 0 1 2 3 4

Statement of Contribution

Surveillance, Innovation, and Growth

This paper is joint work with Torben Klarl, Michael Rochlitz, and Matheus Leusin. Torben Klarl mainly contributed to conceptualization and provided the theoretical model and writing. Michael Rochlitz and I mainly contributed to conceptualization, literature research, and writing. Matheus Leusin and I were involved in supplementing the paper with data and analysis, most of which has been kept out of the paper so far for conciseness. The data used in the paper so far is based on my work.

Digital Governance Solutions

This paper is joint work with Michael Rochlitz. Michael Rochlitz and I contributed similarly to roughly each step in the process, including conceptualization, methodology, and writing. I contributed to all steps in the process and especially the empirical analysis was mostly conducted by me.

Propaganda and Sanction Evasion

This paper is a joint work with Andrey Tkachenko, Michael Rochlitz, Galiya Sagyndykova, and Meruyert Tatkeyeva. Andrey Tkachenko, Michael Rochlitz, and I were mostly involved with literature research, and writing. The paper consists of three parts that are based on different data sources. Andrey Tkachenko and Meruyert Tatkeyeva analysed part one and three of the data, whereas I analysed part two. All of the authors contributed to the conceptualization of the paper and to data collection.

State Surveillance and Self-Censorship

This is a singled-authored paper which is based exclusively on my own work.

Erklärung

Erklärung über die Anfertigung der Dissertation ohne unerlaubte Hilfsmittel

Ich erkläre hiermit, dass diese Arbeit ohne unerlaubte Hilfe angefertigt worden ist und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt wurden.

Ich erkläre ferner, dass die den benutzten Werken wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht wurden.

Eine Überprüfung der Dissertation mit qualifizierter Software im Rahmen der Untersuchung von Plagiatsvorwürfen ist gestattet.

Bremen, August 27, 2024 David Karpa