Essays on the Economics of Social Norms and Identity

Emilio Esguerra

Dissertation

LMU Munich

2023

Essays on the Economics of Social Norms and Identity

Inaugural-Dissertation
zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2023

vorgelegt von
Emilio Esguerra

Referent: Prof. Davide Cantoni, Ph.D.
Korreferent: Prof. Fabian Waldinger, Ph.D.
Promotionsabschlussberatung: 12. Juli 2023

Tag der mündlichen Prüfung: 30. Juni 2023
Namen der Berichterstatter: Davide Cantoni, Fabian Waldinger, Fabian Kosse

Für Christina.

Acknowledgments

Completing this dissertation would not have been possible without the support and encouragement of a great number of advisors, colleagues, friends, and family.

First and foremost, I would like to thank my supervisor Davide Cantoni, without whom I would have never embarked on this great endeavor. Ever since attending his lectures during my undergraduate studies, Davide has been an intellectual source of inspiration, making it a privilege to complete my doctoral thesis under his guidance. Furthermore, I could not have wished for a better mentor, entrusting me with the space, time, and resources to freely explore my research and encouraging me to continuously improve on it. Davide always kept a literal open door to discuss my questions and doubts or to just have an insightful and humorous conversation about politics, history, and life in general, obviously accompanied by a good cup of *caffè*.

I would also like to extend my thanks to the other members of my committee: Fabian Waldinger, for his continuing encouragement of my projects and his critical yet always constructive eye for details, and Fabian Kosse, for fully trusting me with my teaching responsibilities and making me feel as an equal academic peer from the very beginning of my PhD.

I gratefully acknowledge the generous funding from a variety of sources, without which conducting the research presented in this dissertation would not have been possible: the German Research Foundation (DFG) through the GRK 1928 *Microeconomic Determinants of Labor Productivity* doctoral program as well as the CRC-TRR 190 *Rationality and Competition* collaborative research center, the Egon Sohmen Graduate Center, and the Joachim Herz Foundation.

These last five years would not have been half as enjoyable without the great atmosphere created by my fantastic colleagues at the *Lehrstuhl Cantoni*: Leonie Bielefeld, Mathias Bühler, Felix Hagemeister, Kalle Kappner, Cathrin Mohr, Daniela Miehling, Navid Sabet, Leo Vollmer, and Johannes Wimmer. Our obligatory daily coffee breaks and regular Italian lunches made the 4^{th} floor of Ludwigstraße 33 and our *Econhist Bib* feel like a second – at times even first – home. A particularly heartfelt thanks is reserved for Britta Pohr, whose administrative management skills are unmatched and who always had an open ear for us doctoral students.

This dissertation would not be the same without the support and hard work of my fantastic co-authors. To begin with, I certainly could not have wished for a better collaboration than with Leo Vollmer and Johannes Wimmer. Being one year my academic

seniors, they guided me from the beginning and served as great role models on how to tackle the many challenges of writing a dissertation. Together, we embarked on a very productive, sometimes challenging, but always fun-filled intellectual *tour de force* that even the outbreak of a global pandemic could not halt. I could not be prouder of the success our joint research project has enjoyed and, above all, of having found such good and reliable friends.

Next, I thank Felix Hagemeister, Tim Leffler and Julian Heid for a great collaboration that saw a small course project turn into a fully-fledged research paper. I very much enjoyed our productive and humorous working sessions, with our joint conference trip to Dublin being a particularly memorable experience.

Finally, I would like to thank Sebastian Hager for jointly exploring with me the fascinating world of Habsburg and South Tyrolean history. Your creativity, wit, and the inevitable Austrian pastries contributed to turning this into an exciting and fruitful collaboration. I would also like to express my gratitude to Patrick Cuka for his invaluable and diligent work as a research assistant and to Alexia Lochmann for generously sharing her data on emigration records.

Two colleagues, both of whom I met on my very first day, deserve a special mention: Luisa Wallossek, whose availability for well-deserved breaks at any time, steady supply with Tweets and sweets and, most importantly, constant personal encouragement kept me uplifted and motivated throughout. Leonie Bielefeld, my constant companion who was always there to storm into the office next door and discuss any issue at hand, providing us with mutual support to bite our way through until the successful completion of our theses. I immensely admire their intellect, determination, and perseverance and count myself fortunate to have found such loyal friends.

My deepest gratitude goes to my family, and in particular my parents, whose unconditional love, support, and sacrifice only enabled my long educational journey and whom I could count on all these years. They always remained curious and enthusiastic about my work for which I cannot thank them enough. ¡Gracias, Mama, Papá, Elisa, Ceci y Oma!

The last, and without any doubt most important, thank you is reserved for Christina, my partner, soulmate, and best friend. Thank you for having walked with me this long journey from beginning to end; for having endured me all these years – even at times it got a bit stressful; and for always reminding me that there are more important things in life. I could not have done it without your love and support. This thesis is dedicated to you.

EMILIO ESGUERRA

Contents

Introduction

Human beings are by nature social creatures.[1] They care immensely about their relations to others, in particular about how they are seen by others and how they see themselves in this regard (*social identity*). These notions are, in turn, strongly influenced by the sets of beliefs and rules about how members of a group or society should behave (*social norms*). While sociologists, psychologists, and other social scientists have long been using these concepts, they have only recently (re-)captured the interest of economists. In their seminal work, Akerlof and Kranton (2000, 2010) have introduced a novel analytical framework on "identity economics", incorporating social identity and norms into modern economics. In essence, while most economists had previously viewed preferences and tastes as individual characteristics, identity and norms crucially depend on the social context in which economic agents operate (Akerlof and Kranton 2010, p. 6).

These insights have sparked an active research agenda with theoretical extensions on the role of social norms and identity for behavior (Bénabou and Tirole 2011; Charness and Chen 2020), trade policy (Grossman and Helpman 2021), or redistribution and political conflict (Bonomi, Gennaioli, and Tabellini 2021; Shayo 2009, 2020). Increasingly, this has been complemented by empirical work, e.g., highlighting how people choose their identities (Atkin, Colson-Sihra, and Shayo 2021) and how social norms can quickly change (Bursztyn, Egorov, and Fiorin 2020).

My dissertation was inspired by this foundational work and aims to contribute to an on-going research agenda. While the four chapters cover a wide range of settings, they all evolve around a central idea: how the behavior of individuals is influenced by their social identity and the norms of the societies they live in. More precisely, this thesis will first explore the role of *externalities* and *changes* in social norms, before turning to investigating the *roots* and behavioral *consequences* of social identity.

Another common theme of this dissertation is that all papers share an inherently interdisciplinary empirical approach. First, I build on existing work in political science, social psychology, and history to derive quantitatively testable hypotheses. This corresponds to my deep conviction that modern economics can only provide convincing answers to scientific questions when taking existing research in neighboring fields seriously. Second, I employ both natural and field experiments to empirically test these

[1] This idea can be traced back to Aristotle's *Politics* in which he famously describes man as a *zoon physei politikon* – a political or social animal by nature – living in a society that precedes it by definition (see e.g., Schwaabe 2018, p. 59).

hypotheses using novel individual-level data.

CHAPTER 1, which is joint work with Leonhard Vollmer and Johannes Wimmer, investigates why being observed by others influences our own behavior. Recent field experiments have established how observability affects individuals' behavior in a range of settings, from educational investments to charitable giving and voting (see e.g., Bursztyn and Jensen 2015, 2017; DellaVigna, List, and Malmendier 2012; DellaVigna et al. 2017). Typically, such behavioral responses have been attributed to *social pressure*, i.e., an anticipation of being judged by peers and a desire to conform to social norms. We propose an alternative explanation based on *influence motives*, i.e., an anticipation and desire that our own behavior may influence the actions taken by others. In the framework of identity economics, this constitutes an *externality* as individuals' utility may not only depend on their own choices but also on those of others (Akerlof and Kranton 2010, p. 18).

We empirically establish influence motives in the context of the COVID-19 pandemic in Germany. In a survey-based field experiment conducted with a general population sample, we study individuals' willingness to register for a COVID-19 vaccination with the state-monopoly vaccination supplier during an early phase of the pandemic. We group survey participants into pairs, where one participant takes the role of a Sender and the other of a Receiver. We vary whether the Sender's registration decision is shared with the Receiver to introduce social pressure effects. To then isolate influence motives in the Sender's decision, we further randomize whether Senders learn that their decision will be reported to the Receiver *before* or *after* their partner makes his own choice. We find three main results: first, introducing influence motives almost doubles the share of Senders who verifiably registered for a COVID-19 vaccination. This finding is unrelated to a number of alternative explanations including social pressure and experimenter demand. Second, our main effect is driven by participants with higher ex-ante trust in the vaccine. Third, despite anticipating to influence their peers, Senders are not successful in influencing their partners within the experiment. Our results establish influence motives as a complementary channel underlying social signaling which can explain why individuals behave differently under observability even in absence of social pressure effects.

In CHAPTER 2, co-authored with Felix Hagemeister, Julian Heid, and Tim Leffler, we study the *change* of social norms. For example, Bursztyn, Egorov, and Fiorin (2020) have recently shown how social norms can quickly erode in response to new information on the views of others, such as the electoral success of populist politicians. We investigate the role of language – one of the most important channels of social interaction – for the spread of extreme ideas. More concretely, we ask whether being exposed to right-wing (populist) rhetoric exerts a contagious effect on the speech of

other politicians.

Our paper studies the first-time entry of a right-wing populist party, the *Alternative für Deutschland* (AfD), to the German parliament. Using techniques from natural language processing, we measure the cosine similarity of parliamentarians' speeches to the right-wing rhetoric employed by the AfD. To induce individual-level variation in exposure to AfD politicians, we exploit a quasi-exogenous allocation rule for committee members in the German parliament. We show that such higher day-to-day exposure to right-wing populists makes mainstream politicians adopt a more distinctively right-wing language. Comparing a politician with the highest to one with the lowest relative AfD exposure, increases the cosine similarity to right-wing speech by 0.1 of a standard deviation. We corroborate these findings with alternative measures of right-wing rhetoric based on a populism dictionary and on extremist speeches at far-right rallies. Our results seem specific to right-wing populism and suggest strategic motives related to local electoral competition behind rhetorical changes by individual politicians. We attest the spread of right-wing language *within* the political elite, potentially setting a precedent for the subsequent normalization and dissemination to a wider audience.

The remaining two chapters of this dissertation explore the *roots* and *consequences* of one of history's most powerful expressions of social identity: the feeling of belonging to a nation. While nations can be powerful "imagined communities" among strangers (Anderson 1983), in its worst form, nationalism has been linked to violence, war, and the rise of totalitarianism (Arendt 1962; Hutchinson 2017; Wimmer 2013). Both chapters are located in a historical context that is ideally suited for studying questions of national identity, yet has received comparatively little attention among economic historians: the collapse of the multi-ethnic Austro-Hungarian Empire after World War I and its aftermath in the first half of the 20th century.

In CHAPTER 3, which is jointly written with Sebastian Hager,[2] we study how personal grievances affect individuals' willingness to comply with nationalist policy. In 1939, all residents of South Tyrol, a German-speaking and formerly Austrian province of Italy, were forced to choose between emigrating to Germany and assimilating into Italian culture. A generation earlier, South Tyroleans were fighting as Austro-Hungarian soldiers against Italy in World War I (WWI). We investigate whether grievances resulting from WWI impact individuals' willingness to comply with this nationalist mass mobilization. To this end, we match digitized historical emigration records with WWI casualty lists to identify households in which a family member experienced a casualty during WWI. We use exogenous variation in soldiers' front experience for identification of enemy-specific grievances and compare individuals with Italy-specific

[2] I furthermore thank Alexia Lochmann for generously sharing her data on emigration records.

grievances to individuals whose grievances are directed at another enemy. We find a causal effect of grievances on nationalist mobilization: South Tyroleans whose war grievance is directed at Italy were 10 to 15 percentage points more likely to emigrate and choose German citizenship, relative to those who lost someone on another front. We show that deep-seated resentment and grievances can have powerful implications for the success and failure of nation-building policies.

Finally, CHAPTER 4 examines one of the most consequential decisions individuals can be forced to take: whether to risk their life in fighting a war for their nation. While one would not expect rational economic actors to willingly accept injury or death, identity economics can explain such potentially self-destructive behavior (Akerlof and Kranton 2000, p. 717). More precisely, in this chapter I ask whether loyalty and the willingness to fight in war differs by soldiers' (ethnic) identity. I shed light on this question by studying the combat motivation of Italian minority soldiers from Tyrol fighting for Austria-Hungary during World War I. Using a novel individual-level dataset from the Habsburg Army in WWI, I first employ language classification methods to predict whether a soldier was German or Italian. I then find that Italian minority soldiers are significantly less likely to die or get seriously wounded, but more likely to end up in enemy captivity. Crucially, this differential is driven by soldiers with less exposure to the German out-group and materializes only after Italy entered the war against Austria-Hungary in 1915. My results are in line with an interpretation of increasing disloyalty among minority soldiers due to ethnic mistrust and negative nationalism as posited in Alesina, Reich, and Riboni (2020). They underline the powerful impact social identities and feelings of belonging and exclusion can have on the behavior of individuals.

The four essays forming this dissertation are self-contained and can be explored independently. Each chapter is followed by appendices with additional material. A consolidated bibliography is presented at the end of the thesis.

CHAPTER 1

Influence Motives in Social Signaling: Evidence from COVID-19 Vaccinations in Germany*

1.1 Introduction

A prominent literature in economics has documented that being observed by others influences our actions in various domains, including educational investments, financial decision-making, and voting.[1] Typically, behavioral responses to observability are attributed to an anticipation of being judged by our peers and a desire to conform to social norms (*social pressure*). However, with observability also comes the opportunity to influence our peers.

In this paper, we provide evidence from a high-stakes setting that individuals are also motivated by a desire to influence others and separate such *influence motives* conceptually and empirically from social pressure effects. We analyze individuals' decisions whether to register for a COVID-19 vaccination with the state-monopoly vaccination supplier using a pre-registered survey-based field experiment with a general population sample from the German state of Bavaria.[2] Studying COVID-19 vaccinations offers us an ideal setting in which individuals' decisions have substantial and salient externalities and where (prosocial) influence motives likely differ by individuals' trust in vaccine safety and efficacy.

Our study faces a key empirical challenge: both influence motives and social pressure effects arise if behavior is observable and are thus hard to distinguish in observational data. To overcome this challenge, we build random pairs of survey participants, where one takes the role of "Sender" (she) and the other of "Receiver" (he). In a baseline condition *"not informing partner"*, we tell Senders that their decision whether to register

*This chapter is a reproduction of the following published article: Emilio Esguerra, Leonhard Vollmer, and Johannes Wimmer (2023). "Influence Motives in Social Signaling: Evidence from COVID-19 Vaccinations in Germany." *American Economic Review: Insights*, 5 (2), 275–291. © American Economic Association; reproduced with permission.

[1] Bursztyn and Jensen (2017) offer an extensive review of experimental evidence on how observability shapes our behavior. Bénabou and Tirole (2006) provide the seminal theoretical exposition of social signaling in the context of prosocial behavior.

[2] We pre-registered all experimental details at the AEA RCT Registry with ID AEARCTR-0007437 (Esguerra, Vollmer, and Wimmer 2021). In this pre-analysis plan, we referred to the behavioral motive of interest as "anticipated peer effects"; we have since changed our terminology to "influence motives".

for a COVID-19 vaccination will *not* be reported to the Receiver. In the *"informing partner after"* condition, Senders learn that their decision will be shared with the Receiver, but only *after* the Receiver has already decided whether to register. Finally, in the *"informing partner before"* condition, we tell Senders that the Receiver will be informed *before* the Receiver's own registration decision. We expect influence motives to affect Senders' behavior in the *"informing partner before"* condition, while the *"informing partner after"* condition serves as a control group which holds other motives – in particular social pressure and private benefits and costs of signing up – constant.

Subsequently, we provide participants with the opportunity to register for a vaccination via the official state-wide online registration and appointment allocation system for COVID-19 vaccinations in Bavaria, which constituted the only pathway to obtaining the vaccine at the time of the experiment in spring 2021 when vaccines first became available. Once participants had the opportunity to complete their registration, they returned to our survey where we verified their registration status. This institutional setting allowed us to elicit a revealed-preference measure of individuals' willingness to register for a vaccination.

We present three main results. First, influence motives double Senders' likelihood to sign up for a COVID-19 vaccination: we estimate that Senders in the *"informing partner before"* condition are 4.12 percentage points more likely to register for a vaccination, relative to a baseline of 3.79 percent in the *"informing partner after"* condition. Thus, we document that individuals anticipate and act upon the externalities they may have on others' decisions when choosing their privately optimal behavior.

Second, our main result varies strongly with Senders' ex-ante trust in the vaccine: influence motives increase the registration likelihood of Senders by 11.53 percentage points in the top-trust tertile, but they do not significantly affect the behavior of Senders in the mid and bottom tertiles. This suggests that Senders who can influence their partner are more likely to register because they want to send a signal about vaccine quality to their partner.

Third, Senders are not successful in influencing their partners within the experiment. Receivers who had been informed about their Sender's registration decision before they could decide themselves were no more likely to make the same choice as their partner than those who learned about their Sender's choice only after their own decision. This may be due to several reasons, e.g., an absence of social pressure on the Receivers' side as Senders did not observe their decision; Receivers' behavior being affected only with a delay, i.e., after we can observe them in the experiment; or Senders overestimating their impact on others. We present evidence consistent with the latter interpretation by showing that vaccine-trusting Senders who strongly respond to our treatment also shift their perceived probability of influencing their partner substan-

tially upwards.

Our experimental setting allows us to rule out several alternative explanations for our main findings: first, we can hold social pressure constant since we identify influence motives from the comparison of Senders in the *"informing partner before"* with those in the *"informing partner after"* conditions. While the decision of both types is observable and subject to social pressure, only Senders in the *"informing partner before"* condition should infer that they can influence their partner during the experiment.[3] Second, we provide evidence that experimenter demand is unlikely to affect our results: our treatment effects are largest for the revealed-preference measure of *verified* registrations, while we do not find significant effects for *self-reported* alternatives which exhibit no significant cost and are thus more prone to experimenter demand. Third, Senders are not just signaling to their partner that they are willing to register; they actually follow through. Finally, we demonstrate that Senders do not differentially try to cheat and sidestep our registration verification process.

Our paper speaks to a prominent literature in economics studying observability-dependent motives underlying (prosocial) behavior, to which we contribute in two main dimensions: first, we add a complementary mechanism linking observability and behavior, operating even in absence of social pressure; the latter's relevance has been established both theoretically and experimentally (Bénabou and Tirole 2006; Bursztyn and Jensen 2017; DellaVigna et al. 2017; Perez-Truglia and Cruces 2017). Second, we offer an explanation as to why individuals may be willing to deliberately publicize stigmatized information about themselves: while social pressure motives predict that individuals would hesitate to tell others about, e.g., their mental health condition, influence motives would suggest that individuals choose to do so to set a good example.

Influence motives have been an implicit theme in a variety of papers in economics. Previous research has centered on investigating whether and under which conditions leading by example is successful in increasing contributions in public goods games played in the lab.[4] For example, Potters, Sefton, and Vesterlund (2007) show that leading by example increases public goods contribution under asymmetric information, i.e., when an informed leader can signal information about the value of contributing to an uninformed follower. This result is consistent with earlier theoretical work (e.g., Hermalin 1998) as well as with our finding that influence motives vary with Senders' trust in vaccines.

[3]When contrasting influence motives with social pressure we remain agnostic about the extent to which influence motives interact with social- or self-image concerns. Influence motives may well be a facet of self or social signaling – deriving utility from seeing oneself or being seen by others as a leader.

[4]Important contributions include Arbak and Villeval (2013), Cappelen et al. (2016), Dannenberg (2015), Gächter et al. (2012), Drouvelis and Nosenzo (2013), Gächter and Renner (2018), Güth et al. (2007), and Haigner and Wakolbinger (2010).

More closely related are lab experiments studying influence motives in prosocial behavior by Andreoni and Petrie (2004), Karlan and McConnell (2014), and Reinstein and Riener (2012). In contrast to our results, however, neither of these papers find evidence for the existence of influence motives in prosocial settings.[5] As implied by Foerster and van der Weele (2021), one reason for these divergent results might be that influence motives are most relevant where social pressure effects are small – as is the case in our experiment, where participants remain anonymous, but not in the aforementioned previous studies.

Our paper furthermore contributes to an extensive literature examining determinants of vaccination uptake, such as information availability, cues and nudges, or peer effects (Alatas et al. 2022; Alsan and Eichmeyer 2021; Banerjee et al. 2010; Brewer et al. 2017; Milkman et al. 2011) which have recently been re-examined in the course of COVID-19 immunization campaigns (Bartoš et al. 2022; Dai et al. 2021; Milkman et al. 2021). A few studies have investigated the role of social signaling in vaccination choice so far: Karing (2023) shows that parents in Sierra Leone are more likely to have their children vaccinated if they can signal their decision to others. Sasaki, Saito, and Ohtake (2022) report an increase in self-reported intent to take a COVID-19 vaccine among Japanese survey participants when informed that their decision might motivate others to follow suit. We advance these findings by separating the role of influence motives in vaccination decisions using a field experiment.

Finally, our paper offers a methodological contribution by combining key advantages of survey experiments, particularly large general population samples (Alesina, Miano, and Stantcheva 2023; Bursztyn et al. 2022; Roth, Settele, and Wohlfart 2022), with important revealed-preference decisions (Cantoni et al. 2019; DellaVigna, List, and Malmendier 2012; Perez-Truglia and Cruces 2017).

1.2 Experimental Setup

1.2.1 Setting and Sample

We examine Bavarian residents' willingness to register for their first COVID-19 vaccination via the central appointment allocation system BayIMCO, which at the time of our experiment in April 2021 constituted the only pathway for obtaining a vaccination in Bavaria.[6] Owing to supply shortages, access to vaccination appointments was prioritized based on residents' age, pre-existing health conditions, and occupation.

[5] Our setting is prosocial in the sense that it allows us to focus on a behavioral motive that centers on the opportunity of generating externalities for others, while first-order private costs and benefits as well as social pressure are held constant. Analogously to settings where warm-glow altruism motivates generating positive externalities for others, we do not require that influence motives in a prosocial setting are solely motivated by "pure" altruism (Andreoni 1989).

[6] Later on, this was complemented by a decentralized system relying on local doctors' offices.

However, all Bavarian residents were given the opportunity to register online from January 2021 onward, regardless of their prioritization status. Once vaccine supply and their prioritization status allowed, registered residents automatically received a vaccination appointment through BayIMCO.

We recruit a sample of Bavarian residents who had neither registered for nor received a vaccination.[7] In total, 1,857 participants completed our survey, for which we report summary statistics in Appendix Table 1.C.1: in our sample, the share of participants willing to get vaccinated is 51 percent, which is – due to our exclusion of already vaccinated and registered individuals – somewhat lower than the vaccination willingness of 65 percent elicited in a nationally representative study at the same point in time (Betsch, Wieler, and Habersaat 2020; COSMO 2021). With respect to other key characteristics such as gender, age, and income, our study participants are suitably representative of the Bavarian population.[8]

1.2.2 Experimental Design

1. Introduction Prior to starting the survey, all participants are informed that they are about to take part in a scientific study on their attitudes towards COVID-19 vaccinations. We then begin by collecting demographic information as well as a rich set of attitudes, beliefs, and preferences related to the vaccination (e.g., beliefs about the safety and efficacy of vaccines).[9]

2. Joint problem solving task Next, we build teams consisting of two randomly paired participants who collaborate on a joint problem solving task adopted from Goette and Tripodi (2021) to induce social proximity between partners as well as to convince participants that they are interacting with a human subject and not a chatbot.

3. Social proximity Subsequently, we again follow Goette and Tripodi (2021) and elicit perceived social proximity between partners using the *"oneness"* scale (Cialdini et al. 1997) as a fast and simple way of measuring relationship closeness. Gächter, Starmer, and Tufano (2015) confirm the original results underlying the *"oneness"* scale in a large general-population sample and conclude that it is a useful tool to meaningfully measure social proximity without the need to draw on more elaborate relation-

[7] At the time of the experiment, 15 percent of the Bavarian population had received at least one vaccination and a further 30 percent had registered at BayIMCO.

[8] Roughly half of our sample is female; mean age and monthly net income are 40.9 years and €2,907, respectively, compared to official state averages of 43.7 years in 2017 (Bayerisches Landesamt für Statistik 2019) and €2,549 in 2018 (GESIS – Leibniz-Institut für Sozialwissenschaften 2019).

[9] We provide a more detailed description of the experimental design in Appendix 1.A as well as the complete survey instrument in Appendix 1.E.

ship inventories.[10]

4. Treatment Next, teams enter the treatment stage, where we vary (1) whether the Sender's decision to register for a COVID-19 vaccination is reported to the Receiver and (2) whether this happens before or after the Receiver's decision. Across experimental conditions, Senders always decide first and are informed that they will not learn about Receivers' decisions.[11]

In the *"not informing partner"* condition, we inform Senders that their registration decision will not be reported to their partner, providing no scope for neither influence motives nor social pressure effects. In the *"informing partner after"* condition, Senders learn that their registration decision will be shared with their partner.[12] However, we highlight to Senders that their partner will only be informed about their decision once their partner has already taken his registration decision. Therefore, while social pressure effects may arise, Senders cannot influence their partner's decision within the experiment and, consequently, influence motives should play no role in this condition. In the third condition, *"informing partner before"*, we inform Senders that their partner will learn about their decision *before* their partner takes his registration decision. As in the previous condition, Senders are subject to social pressure effects. On top, Senders should now infer that they can influence their partner's registration decision within the experiment, opening room for influence motives. Hence, by comparing Senders' willingness to sign up for a vaccination between Senders who can and those who cannot influence their partner, we can hold social pressure constant and thereby causally assess the importance of influence motives in social signaling.

Irrespective of experimental condition, all Receivers are informed that they will decide after the Sender but they do not learn ex ante whether and when they will be informed about their partner's decision. While Senders take their decision whether to sign up, Receivers are directed to a waiting page where they remain until their Sender has decided or 60 seconds have elapsed. Subsequently, Receivers in the *"informed before"* condition are informed about their partner's decision, whereas Receivers in the *"informed after"* condition only receive this information after we provided them with the opportunity to sign up for a vaccination.

[10] We provide a screenshot of how we elicited the *"oneness"* scale in the survey in Appendix Figure 1.D.2. In Appendix Table 1.C.6, we show that this scale correlates as expected with plausible predictors of social proximity. Importantly, partner (dis-)similarity in age and in gender – the two key characteristics participants were told about their partner – are the strongest predictors of our social proximity measure.

[11] In Appendix Figure 1.B.1, we illustrate the main intuition of our design and report the treatment messages shown to Senders.

[12] At the time when Senders learn that their partner will be informed about their decision, they do not yet know that we plan to verify their self-reported registration. Thus, in the experimental instructions, we refer to the Sender's registration decision without further specifying whether we share her self-reported intent to sign up, her self-reported registration status, or her verified registration.

5. First stage Subsequently, we ask Senders how likely they think it is that they can influence their partner's registration decision using a slider ranging from 0 to 100. We use this metric to obtain an estimate of the first-stage effect of our manipulations on Senders' beliefs about their impact on their partner's registration decision.

6. Main outcome Next, we elicit our main outcome by asking participants whether they wish to register for a COVID-19 vaccination right away. If participants answer "yes", they are forwarded to the official registration website (BayIMCO) outside of our survey.[13] Participants responding "no" are forwarded to the next module of our survey. After completing the BayIMCO form – which took participants on average six minutes – participants obtain an email from BayIMCO officially confirming their registration. We use this confirmation email to verify whether participants indeed registered for a vaccination by asking them to report the sending address and the subject line to us. We incentivized participants by informing them that reporting both pieces of information correctly would qualify them for one of 30 additional €20 Amazon vouchers. After participants enter their information, their responses are checked by our system. If both answers are correct, a lottery determines whether participants obtain one of the vouchers. Participants only learn whether they won on the final page of the survey.[14]

The timing of the steps we use to elicit whether participants actually signed up for the vaccination is crucial: when participants are offered the opportunity to sign up for the vaccination, they do not yet know that we will ask them to provide proof of their registration. Participants learned about the confirmation and the corresponding remuneration only once they reported that they successfully completed the registration. Hence, participants do not have an incentive to misreport their registration in order to qualify for one of the vouchers.

7. Further outcomes and demographics Finally, we elicit participants' post-treatment attitudes and beliefs related to the COVID-19 vaccination, e.g., their stated willingness to ever take such a vaccine. We end the survey by collecting further demographic information, revealing payoffs to participants, and providing them with the opportunity to comment on the survey.

[13]The BayIMCO registration website can be accessed at https://impfzentren.bayern/citizen/ (last visited on 07/25/2022). We provide screenshots illustrating how we elicit and verify the registration decision in Appendix 1.D.6.

[14]In Appendix Table 1.C.5, we assess the explanatory power of Senders' predetermined characteristics in predicting their verified registration status. We find that Senders' trust in the safety of COVID-19 vaccines, and their own willingness to ever take such a vaccine constitute the most important predictors.

1.2.3 Experimental Assignment and Sample Balancing

We use a two-stage random procedure to assign participants into experimental conditions: first, we randomly assign teams to one of the three experimental conditions *"not informing partner"*, *"informing partner after"*, or *"informing partner before"*. Second, within teams, we further randomly assign the roles of Sender and Receiver.[15]. We test for systematic predetermined differences across Senders in different experimental conditions in Appendix Table 1.C.4, yielding that Senders' characteristics are well balanced across conditions.

1.3 Empirical Analysis

1.3.1 Regression Specification

To identify treatment effects, we estimate the following regression model:

$$y_i = \beta_0 + \beta_1 \cdot \text{informing partner}_i + \beta_2 \cdot \text{informing partner before}_i + x_i'\gamma + \epsilon_i, \quad (1.1)$$

where y_i is a dummy variable taking value 100 if Sender i verifiably registered for a COVID-19 vaccination and zero otherwise;[16] *informing partner*$_i$ is a dummy variable taking value 1 if Sender i learns that her registration decision is reported to her partner in the experiment; and *informing partner before*$_i$ is a dummy variable taking value 1 if Sender i's partner is informed *before* instead of *after* his own registration decision. Coefficient β_1 thus captures potential effects related to social pressure, while β_2 measures the additional effect of influence motives. Finally, in some specifications we control for Senders' predetermined characteristics (x_i).[17]

1.3.2 Average Treatment Effects

We begin our analysis by documenting average treatment effects of observability on Senders' behavior which we then decompose to isolate influence motives from social pressure. In Table 1.1, columns 1 and 2, we estimate how Senders' registration likelihood responds when they learn that their Receiver will be informed about their registration decision. When pooling Senders in the *"informing partner before"* and *"informing partner after""* conditions in this way, we obtain small and statistically not significant treatment effects. Thus, the observability of a Sender's decision per se does not alter

[15] We report the resulting assignment into experimental conditions in Appendix Table 1.C.2

[16] Throughout all results, we multiply binary outcomes by 100. As long as estimates are not further scaled, they can be interpreted as percentage-point changes in the outcome, e.g. in the probability of verifiably registering for a COVID-19 vaccination.

[17] We use all characteristics reported in Appendix Table 1.C.1 as controls, except social proximity which we exclude since some participants skipped the corresponding survey items. Our estimates remain unchanged when including it.

Table 1.1: Average treatment effects on Senders' likelihood to register for a COVID-19 vaccination

	Verified registration			
	(1)	(2)	(3)	(4)
Informing partner	1.21	1.16	−0.78	−0.84
	(1.36)	(1.32)	(1.41)	(1.39)
Informing partner before			4.11***	4.12***
			(1.44)	(1.37)
Controls	-	✓	-	✓
Mean, 'Not informing partner'	4.57	4.57	4.57	4.57
Mean, 'Informing partner after'	3.79	3.79	3.79	3.79
Observations	1,401	1,401	1,401	1,401
R^2	0.00	0.09	0.01	0.09

Notes: Results derived from regressions as laid out in Equation 1.1. We employ Senders' verified registration status as the dependent variable, i.e., a dummy variable taking value 100 if a Sender verifiably registered for a vaccination and 0 otherwise. Coefficient estimates can thus be interpreted as percentage-point changes in the probability that a Sender registered for a vaccination. Controls include the full set of variables reported in Appendix Table 1.C.1 with the exception of social proximity. Robust standard errors reported in parentheses. Significance levels: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Senders' registration behavior in our experiment.

In contrast, columns 3 and 4 present the substantial and statistically significant impact of influence motives on Senders' registration decision. Here, we show results based on Equation 1.1 allowing us to separate the impact of influence motives from social pressure under observability: our estimates for β_2 in column 4 document that influence motives increase registration shares among Senders by 4.12 percentage points. Relative to a baseline of 3.79 percent in the *"informing partner after"* condition, this implies a substantial and statistically significant increase in registration likelihood of 109 percent. This result demonstrates that individuals, when choosing their privately optimal action, also account for the externalities they may have on others' decisions. Our estimates for β_1, however, are close to zero and not statistically significant, implying that social pressure effects do not alter Senders' likelihood to sign up for a COVID-19 vaccination in our setting.[18]

The dominance of influence motives over social pressure resonates well with a recent paper by Foerster and van der Weele (2021) who show theoretically that image effects and influence motives can function as substitutes in social signaling: in settings

[18]The statistical significance of our estimates is confirmed by results from Fisher permutation tests summarized in Appendix Figure 1.B.2.

where Senders tend to care less about their (social) image, influence motives may nevertheless motivate Senders to pursue a costly action entailing a positive externality on observing individuals.[19]

1.3.3 Heterogeneous Treatment Effects

One might expect that high- and low-vaccine-trust individuals have differing prosocial motives to influence others' behavior. We therefore examine whether treatment effects vary by Senders' trust in the quality of COVID-19 vaccines, which we measure as the mean of two survey items elicited pre-treatment capturing Senders' beliefs about the safety and efficacy of COVID-19 vaccines.[20] To investigate heterogeneous treatment effects, we split our sample into tertiles according to Senders' trust in vaccines and present results in Table 1.2.

In Panel A, columns 1 to 4, we confirm our main result for the subsample of Senders exhibiting the highest trust in the vaccines: estimates for the "influence motive" parameter β_2 in this group imply a 11.53 percentage point (137 percent) increase in registration shares relative to a baseline of 8.44 percent. As we document in Panel B, treatment effects for Senders with medium levels of trust remain positive but lose significance (4.36 percentage points). In stark contrast, Panel C shows that treatment effects for Senders with the lowest trust in vaccines vanish completely and thus differ substantially from those in the high-trust tertile. The corresponding estimate for the influence motive parameter β_2 is precisely zero, which is due to the fact that not a single Sender in the bottom-trust tertile decided to register for the vaccination within our experiment.

Hence, average treatment effects seem to be driven by Senders exhibiting high levels of trust in COVID-19 vaccines, a demographic with a high likelihood of registering for and receiving the vaccination eventually. Arguably, the key trade-off for many of these Senders was between registering during our survey and doing it at some later point after the experiment. Vaccine supply shortages coupled with strict prioritization of access at the time of our experiment implied that Senders could reasonably expect that their vaccination would only take place in several weeks. This, in turn, left substantial room for delaying the registration, even for strong vaccination supporters. Thus, our

[19] In our experiment, social pressure effects are muted by three factors: first, Senders interact with strangers on the internet instead of neighbors (Bursztyn, González, and Yanagizawa-Drott 2020) or classmates (Bursztyn and Jensen 2015). Second, the number of observers Senders are facing is smaller than in many existing studies (e.g. Perez-Truglia and Cruces 2017). Third, by upholding anonymity throughout the entire experiment, we shut down most instrumental motives underlying social pressure effects arising from potential future encounters with the Receiver. However, while being muted, a recent study by Goette and Tripodi (2021) shows that social pressure effects can arise even under anonymity; hence, we allowed for the possibility of social pressure effects in our pre-analysis plan.

[20] We plot the distribution of Senders' beliefs about the safety and efficacy of COVID-19 vaccines in Appendix Figure 1.B.3.

Table 1.2: Treatment effects on Senders' likelihood to register for a COVID-19 vaccination and their perceived likelihood of influencing partner's registration decision conditional on trust in vaccines

	Verified registration				Perceived likelihood that partner can be influenced (%)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: High trust in vaccines (top tertile)								
Informing partner	2.09	2.44	−3.56	−3.23	9.17***	10.53***	3.95	5.55
	(3.83)	(3.82)	(3.96)	(4.02)	(3.32)	(3.33)	(3.65)	(3.73)
Informing partner before			11.70***	11.53***			10.80***	10.12***
			(4.04)	(3.88)			(3.36)	(3.34)
Mean, 'Not informing partner'	12.00	12.00	12.00	12.00	29.74	29.74	29.74	29.74
Mean, 'Informing partner after'	8.44	8.44	8.44	8.44	33.69	33.69	33.69	33.69
Observations	398	398	398	398	398	398	398	398
R^2	0.00	0.11	0.02	0.13	0.02	0.11	0.04	0.13
Panel B: Medium trust in vaccines (mid tertile)								
Informing partner	0.83	0.32	−1.13	−1.78	7.60***	8.40***	6.88**	7.11**
	(2.36)	(2.30)	(2.40)	(2.38)	(2.75)	(2.78)	(3.07)	(3.05)
Informing partner before			4.01*	4.36*			1.47	2.69
			(2.34)	(2.41)			(2.77)	(2.74)
Mean, 'Not informing partner'	3.61	3.61	3.61	3.61	22.78	22.78	22.78	22.78
Mean, 'Informing partner after'	2.48	2.48	2.48	2.48	29.66	29.66	29.66	29.66
Observations	398	398	398	398	398	398	398	398
R^2	0.00	0.08	0.01	0.09	0.02	0.13	0.02	0.13
Panel C: Low trust in vaccinces (bottom tertile)								
Informing partner	0.00	0.00	0.00	0.00	2.30	1.41	0.23	−0.74
	(0.00)	(0.00)	(0.00)	(0.00)	(2.88)	(2.94)	(3.12)	(3.22)
Informing partner before			0.00	0.00			4.37	4.46
			(0.00)	(0.00)			(2.85)	(2.92)
Mean, 'Not informing partner'	0.00	0.00	0.00	0.00	20.28	20.28	20.28	20.28
Mean, 'Informing partner after'	0.00	0.00	0.00	0.00	20.51	20.51	20.51	20.51
Observations	398	398	398	398	398	398	398	398
R^2	0.00	0.00	0.00	0.00	0.00	0.07	0.01	0.07
Controls	-	✓	-	✓	-	✓	-	✓

Notes: Results derived from regressions as laid out in Equation 1.1 where we restrict the sample to Senders for which we observe both their verified registration status and their perceived likelihood of influencing their partner (N=1194). We employ the following dependent variables: in columns 1-4, we use a dummy variable taking value 100 if a Sender verifiably registered for a vaccination and 0 otherwise; in columns 5-8, we employ Senders' beliefs about the likelihood of influencing their partner (ranging from 0 to 100). We measure Senders' trust in COVID-19 vaccines using the average of Senders' beliefs about the safety and the efficacy of these vaccines. We then run regressions on three subsamples of Senders: in Panel A, we employ Senders with high trust in COVID-19 vaccines (top tertile); in Panel B, Senders with medium levels of trust (mid tertile); and in Panel, C we restrict our sample to Senders with low levels of trust (bottom tertile). Controls include the full set of variables reported in Appendix Table 1.C.1 with the exception of social proximity. Robust standard errors reported in parentheses. Significance levels: $*p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$.

experimental manipulation needed to counteract a preference for delaying the necessary but bothersome task of registering for the COVID-19 vaccination into the future, e.g., due to present bias, but did not need to fundamentally change Senders' views on the vaccination (registration). Indeed, as we document in Appendix Table 1.C.8, Senders' attitudes and beliefs on COVID-19 vaccines remain unaffected.

In Table 1.2, columns 5 to 8, we present treatment effects on Senders' beliefs about their influence on their partner's registration decision, again separately for the three levels of trust in the vaccine.[21] One can see in Panel A that the perceived probability of influencing their partner was significantly shifted up by 10.12 percentage points (30 percent) among high-trust Senders. We expect the strength of influence motives to be a function of Senders' perceived probability of influencing their partner. The substantial upward shift in this belief among responsive high-trust Senders resonates well with the idea that for complying Senders influence motives triggered by our treatment were strong enough to push the benefits of registering during our survey above its costs.[22]

1.3.4 Ruling out Alternative Explanations

We continue by demonstrating that experimenter demand is unlikely to explain our main findings. Due to its positive externalities on others, taking a COVID-19 vaccine, and by extension also registering for a vaccination, should be perceived as a socially desirable action by Senders trusting in their quality. Thus, we expect a certain baseline level of experimenter demand effects to be present in all experimental conditions. Such demand effects would challenge our findings only if Senders in the *"informing partner before"* condition were more likely to infer from our instructions that we expected them to sign up for a COVID-19 vaccination than Senders in the *"informing partner after"* condition. To assess this alternative explanation, we compare treatment effects between a revealed-preference measure (verified registrations) and several self-reported alternatives. If experimenter demand effects explained why Senders are more likely to register in the *"informing partner before"* condition, we would expect to see even larger treatment effects for self-reported outcomes not implying any significant economic costs (Haaland, Roth, and Wohlfart 2023).

In Table 1.3, we contrast our positive and significant estimates for the "influence mo-

[21] Appendix Table 1.C.10 reports average treatment effects on first stage beliefs for our full sample of Senders, corresponding to our main results Table 1.1.

[22] Moreover, coefficient estimates for the change in Senders' perceived probability of influencing their Receiver are likely attenuated by participant inattention, to be expected in a general population online sample. As Appendix Figure 1.B.5 shows, the distribution of Senders' beliefs about the likelihood of influencing their partner reveals a significant amount of bunching. Importantly, however, a substantially larger share of Senders in the *"informing partner before"* condition anticipate very high likelihoods of influencing their partner (> 70 percent), suggesting that among attentive Senders, the perceived probability of influencing their partner – and hence the potential for influence utility – was shifted up quite substantially.

Table 1.3: Comparing average treatment effects on Senders between revealed preference and self-reported outcomes

	Verified registration	Self-reported registration status	Clicked registration link	Self-reported intent to register	Self-reported willingness to take vaccine (post-treatment)
	(1)	(2)	(3)	(4)	(5)
Informing partner	−0.84	1.42	−0.05	1.18	−0.29
	(1.39)	(2.00)	(1.91)	(2.53)	(0.94)
Informing partner before	4.12***	3.16*	2.59	2.29	−0.40
	(1.37)	(1.88)	(1.71)	(2.30)	(0.91)
Controls	✓	✓	✓	✓	✓
Mean, 'Not informing partner'	4.57	9.45	8.84	19.82	49.37
Mean, 'Informing partner after'	3.79	11.01	8.84	21.66	50.46
Observations	1,401	1,401	1,401	1,401	1,401
R^2	0.09	0.13	0.12	0.21	0.85

Notes: Results derived from regressions as laid out in Equation 1.1. We employ the following dependent variables: (column 1) a dummy variable taking value 100 if a Sender verifiably registered for a vaccination; (column 2) a dummy variable taking value 100 if a Sender's reported that she had registered (elicited before verification); (column 3) a dummy variable taking value 100 if a Sender clicked on the registration link forwarding her to BayIMCO; (column 4) a dummy variable taking value 100 if a Sender reported that she will sign up for a vaccination; (column 5) Sender's self-reported willingness to take the vaccine elicited post treatment (ranging from 0 to 100). Controls include the full set of variables reported in Appendix Table 1.C.1 with the exception of social proximity. Robust standard errors reported in parentheses. Significance levels: $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$.

tive" parameter β_2 obtained from regressions employing Senders' verified registration status as the dependent variable (column 1) with much smaller and statistically not significant estimates derived from regressions with self-reported alternatives: when using Senders' self-reported registration status, β_2 amounts to a 29 percent increase over the mean (column 2) – compared to a 109 percent increase for the verified status. Estimates for β_2 are even smaller when employing a dummy measuring whether a Sender clicked on the link forwarding her to the BayIMCO website as the dependent variable (column 3). The same applies to estimates based on Senders' self-reported intent to register (column 4) which we elicit after Senders saw the treatment messages, yet before we offered them the opportunity to sign up. Finally, even when using Senders' self-reported likelihood of ever taking the vaccine as the dependent variable, we obtain precisely estimated zero effects for β_2 (column 5).

Hence, *self-reported* registration willingness does not differ between the *"informing partner before"* and the *"informing partner after"* conditions, while *verified* registration willingness clearly does.[23] Together these findings render it unlikely that an *"informing partner before"*-specific experimenter demand effect biases our average treatment effects on Senders' verified registration status; instead, moving from *self-reported* to *verified* willingness seems to weed out a general experimenter demand effect regard-

[23]We further corroborate this finding in Appendix Table 1.C.8 where we show that Senders' beliefs and attitudes regarding the vaccine, e.g., Senders' perceptions of its social desirability, remain unaffected by our experimental manipulations.

ing registration willingness present in all experimental conditions.

In the Appendix, we provide evidence speaking against additional alternative explanations: in Appendix Table 1.C.9, we show that Senders in the *"informing partner before"* condition are not just signaling to their partner that they are willing to register for a vaccination; they actually follow through and are thus unlikely to lie strategically. Finally, in Appendix Figure 1.B.4, we demonstrate that Senders in the *"informing partner before"* condition do not differentially try to cheat and sidestep our registration verification process.

1.3.5 Peer Effects on Receivers

In a final exercise, we show that Senders are not successful in actually influencing Receivers' behavior within the experiment. We use the following regression model:[24]

$$y_i = \phi_0 + \phi_1 \cdot \text{informed before}_i + x_i'\zeta + \epsilon_i, \tag{1.2}$$

where y_i is a dummy variable taking value 100 if the Receiver decides in the same way as his partner and zero otherwise. Our main explanatory variable *informed before*$_i$ is a dummy which equals 1 if Receiver i was informed about his partner's decision before we offered the Receiver the opportunity to sign up for a vaccination and 0 otherwise.

In Table 1.4, Panel A, column 1, we base the same decision dummy on Receivers' verified registration status and obtain a small negative and statistically not significant coefficient estimate for ϕ_1. This implies that Receivers who had been informed about their Sender's registration decision before they could decide themselves were no more likely to make the same choice as their partner than those who learned about their Sender's choice only after their own decision. Estimates remain small and statistically not significant throughout columns 2 to 4, where we employ self-reported alternative measures of Senders' willingness to sign up for a vaccination. As we document in Table 1.4, Panel B, Receivers' attitudes and beliefs about COVID-19 vaccinations remained similarly unaffected by Senders' decisions. Only Receivers' self-reported likelihood to freeride on others' vaccination decisions is negatively affected by being informed about the Sender's decision beforehand; the effect is, however, only marginally significant. Together, these results suggest that Senders' decisions do not exhibit meaningful peer effects on Receivers' behavior.

The absence of peer effects can be explained by the specific setting we study: Receivers learned about Senders' decisions, yet not vice versa, and were thus not subject to social pressure, a potential channel for conformity. Moreover, Receivers' behavior may only

[24] As we document in Appendix Table 1.C.11, Receivers' predetermined characteristics are well balanced across experimental conditions.

Table 1.4: Average treatment effects on Receivers

	(1)	(2)	(3)	(4)
Panel A: Effects on registration decision				
	$\mathbb{I}$[Receiver decides like Sender]			
	Verified registration	Self-reported registration status	Clicked registration link	Self-reported intent to register
Informed before	−0.16	2.63	3.71	3.80
	(2.34)	(3.31)	(2.97)	(4.16)
Mean, 'Informed after'	92.80	83.47	86.02	71.19
Observations	456	456	456	456
R^2	0.04	0.03	0.06	0.07
Panel B: Effects on attitudes and beliefs				
	Δ Pre/post treatment			
	Self-reported willingness to take vaccine	Safety & efficacy of vaccine	Likelihood to freeride on others' vaccination decision	Social desirability of vaccine
Informed before	−0.39	−0.09	−0.17*	−0.11
	(1.01)	(0.09)	(0.09)	(0.08)
Mean, 'Informed after'	0.17	0.06	0.08	0.07
SD, 'Informed after'	9.72	1.04	1.12	0.98
Observations	456	456	456	456
R^2	0.03	0.04	0.05	0.39
Controls	✓	✓	✓	✓

Notes: Results derived from regressions as laid out in Equation 1.2. In Panel A, outcomes are "same decision dummies" taking value 100 if the Receiver decided in the same way as the Sender he was matched with and 0 otherwise. Coefficient estimates can thus be interpreted as percentage-point changes in the probability that the Receiver decides the same way as the Sender. We construct "same decision dummies" for the following dependent variables: (Panel A, column 1) verified registrations; (Panel A, column 2) self-reported registration status (elicited before verification); (Panel A, column 3) participants' click on the link forwarding participants to BayIMCO; and (Panel A, column 4) self-reported intent to register. In Panel B, outcomes capture changes in Receivers attitudes or beliefs pre- versus post-treatment. We consider such changes for the following dependent variables: (Panel B, column 1) willingness to take the vaccine; (Panel B, column 2) beliefs about safety and efficacy of the vaccine; (Panel B, column 3) likelihood to freeride on others' vaccination decision; and (Panel B, column 4) beliefs about the social desirability of the vaccine. Controls in Panel A include the full set of variables reported in Appendix Table 1.C.1 with the exception of social proximity; each column of Panel B further excludes the pre-treatment level of the respective outcome measure. Robust standard errors reported in parentheses. Significance levels: $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$.

be affected with a delay, i.e., after we can observe them in the experiment, or only after receiving several signals from others. However, the absence of peer effects could also suggest that in certain decision environments people perceive themselves as more pivotal than they actually are, such that influence motives alter people's behavior even in absence of actual peer effects. In line with this interpretation, the evidence compiled in Table 1.2, columns 5 to 8, suggests that high-vaccine-trust Senders, who strongly responded to our treatment, also substantially shift their perceived probability of influencing their partner upwards; an anticipation that did not translate into an actual impact on Receivers.

1.4 Conclusion

Our work provides evidence for influence motives in social signaling in the context of COVID-19 vaccinations. We show that individuals, when choosing their privately optimal behavior, also take into account the externalities they may have on others' decision-making. Our evidence thus helps explain why individuals behave differently when being observed by others, even in absence of social pressure effects.

Social influence motives were likely of particular relevance during the early roll-out of COVID-19 vaccines in Germany: in this phase of the pandemic, the externalities of individuals' vaccination decisions were highly salient in public discourse and widespread uncertainty about potential benefits and costs of taking the vaccine offered substantial scope for vaccine supporters (and opponents) to signal their views to others. This holds especially for individuals with strong views on the vaccination who may perceive themselves as better informed than others. Cultural norms may have also reinforced social influence motives in the German setting, where individuals are frequently called upon to consider themselves as role models for others, e.g., at traffic lights where they are reminded that jaywalking might teach children to replicate dangerous behavior. The anonymity of the Sender-Receiver interaction in our experiment likely dampened social pressure effects and may, in turn, have increased the role of influence motives, as suggested by recent theory (Foerster and van der Weele 2021); however, one might also expect influence motives to be even larger among groups of friends or colleagues, especially for individuals who consider themselves influential for their peers and are not worried about their social standing in the group.

Future work could further explore the role of influence motives in other settings and investigate the underlying mechanisms. In particular, we believe that developing a better understanding of the formation of individuals' – perhaps inaccurate – beliefs about their impact on others constitutes an interesting avenue for future research.

Appendix to Chapter 1

1.A Experimental Details

In this section, we provide more details on our experimental design allowing us to disentangle influence motives from social pressure effects. We pre-registered all features of our experimental design at the AEA RCT registry under ID AEARCTR-0007437 (Esguerra, Vollmer, and Wimmer 2021) before the experiment commenced.[25] The experiment was approved by the Ethics Committee of the Department of Economics at LMU Munich, protocol 2021-01. We begin by discussing our sample in more detail and continue by illustrating the experimental design at greater length. In the final subsection, we discuss the assignment of Senders and Receivers to our experimental conditions and show that Senders' predetermined characteristics are balanced across these conditions.

1.A.1 Sample

We recruited participants for our survey from CINT, an online panel provider. During our experiment's field time in April 2021, approximately 15 percent of the Bavarian population had already received at least one vaccination and a further 30 percent had registered in the central system. We exclude both of these groups from our experiment by screening them out at the start of our survey. In total, 1,857 participants completed our experiment, for which we report summary statistics on their characteristics in Appendix Table 1.C.1: 51 percent of our participants reported to be willing to get vaccinated against COVID-19 at some point (elicited pre-treatment), which is – due to our exclusion of already vaccinated and registered individuals – somewhat lower than the vaccination willingness of 65 percent elicited in a nationally representative study at the same point in time (Betsch, Wieler, and Habersaat 2020; COSMO 2021). Roughly half of our sample is female; mean age and monthly net income are 40.9 years and €2,907, respectively, compared to the official state averages of 43.7 years in 2017 (Bayerisches Landesamt für Statistik 2019) and €2,549 in 2018 (GESIS – Leibniz-Institut für Sozialwissenschaften 2019). Hence, our sample seems suitably representative of the Bavarian population as a whole.

[25] In our pre-analysis plan, we referred to the behavioral motive of interest as "anticipated peer effects"; we have since changed our terminology to "influence motives".

1.A.2 Implementation

For the technical implementation of our online experiment, we used the open-source software oTree which allows us to let participants interact during our survey in real time (Chen, Schonger, and Wickens 2016). Our survey was hosted on a HEROKU server in Germany.

1.A.3 Survey Design and Treatments

1. Introduction Prior to starting the survey, participants invited by the online panel provider are informed that they are about to take part in a scientific survey studying attitudes towards COVID-19 vaccinations. They then receive information on data protection and are asked to consent to these terms and their participation in the survey. We begin by screening out all participants who indicate to have already been vaccinated or registered for a COVID-19 vaccination. From all remaining participants we collect basic demographic information as well as a rich set of attitudes, beliefs, and preferences related to the vaccination. For example, we elicit participants beliefs about the safety and efficacy of the vaccine as well as their beliefs about the social desirability of taking such vaccine. To elicit attitudes and beliefs related to COVID-19, we rely on pre-tested survey items adopted from the COVID-19 Snapshot Monitoring project (COSMO 2021).

2. Joint problem solving task Before entering the main stage of the experiment, we build teams consisting of two randomly paired participants. Within teams, participants are randomly assigned either to the role of Sender (she) or Receiver (he). Teams work on a joint problem solving task adopted from Goette and Tripodi (2021) which we use to induce social proximity between the partners and to allow participants to verify that they are interacting with a human subject and not a chatbot. The task consists of four consecutive questions, in which teams are presented with historical paintings and are asked to select the corresponding artist from a list. Each correct answer increases participants' probability of winning an Amazon voucher, but only if their partner selects the correct artist as well. To allow for coordination between partners, we provide them with the option to exchange text messages.[26] Participants are informed as to whether they won any of the vouchers on the final page of the survey.

3. Social proximity After the joint task, we again follow Goette and Tripodi (2021) and elicit perceived social proximity between partners using the *"oneness"* scale (Cialdini et al. 1997), as a fast and simple way of measuring relationship closeness.[27] Gächter,

[26] We present a screenshot of the joint problem solving task showing a sample painting and the chat window in Figure 1.D.1 in Appendix 1.D.

[27] The oneness scale is computed as the mean of the Inclusion of Other in the Self (IOS) scale (Aron, Aron, and Smollan 1992) and the (ii) WE scale (Cialdini et al. 1997). We provide screenshots of how we

Starmer, and Tufano (2015) confirm the original results underlying the *"oneness"* scale in a large general-population sample and conclude that it is a useful tool to meaningfully measure social proximity without the need to draw on more elaborate relationship inventories. In Appendix Table 1.C.6, we show that the *"oneness"* scale correlates as expected with plausible predictors of social proximity: Senders report higher levels of social proximity to their partner if their partner shares the same gender or educational level. At the same time, reported social proximity declines in larger age and income differences. Partner (dis-)similarity in age and in gender – the two key characteristics participants were told about their partner – are the strongest predictors of our social proximity measure.

4. Treatment Next, teams enter the experiment's treatment stage, where we use two experimental manipulations to identify influence motives in Senders' decision to sign up for a COVID-19 vaccination: we vary (1) whether the Sender's decision to register for a COVID-19 vaccination is reported to the Receiver and (2) whether this happens before or after the Receiver's decision.

The main intuition of our design is illustrated in Appendix Figure 1.B.1. For each experimental condition, we report the treatment instructions shown to the Sender and the corresponding decision sequence as implemented in the experiment. Irrespective of the condition to which we assigned teams, Senders were always offered the opportunity to sign up for the vaccination before Receivers and were explicitly told that they would not learn about the decision of their partner afterwards.

In the *"not informing partner"* condition, we inform Senders that their decision on whether to register for a vaccination will not be reported to their partner. As a result, neither influence motives nor social pressure effects should affects Senders' registration decisions.

In the *"informing partner after"* condition, Senders learn that their decision will be shared with their partner. However, we highlight to Senders that their partner will only be informed about their registration decision once he (the partner) has already taken his own registration decision. Therefore, while social pressure effects may arise, Senders cannot influence their partner's decision within the experiment and, consequently, influence motives should play no role in this condition.

In the third and final condition, *"informing partner before"*, we inform Senders that their partner will learn about their registration decision *before* he takes his own registration decision. As in the previous condition, Senders in this condition are subject to social pressure effects. On top, Senders should infer that they can now influence their partner's registration decision within the experiment and update their beliefs

elicited the oneness scale in Appendix Figure 1.D.2.

about their potential impact on their partner's decision accordingly. Hence, by comparing Senders' willingness to sign up for a vaccination between Senders who can (*"informing partner before"*) and those who cannot (*"informing partner after"*) influence their partner's registration decision, we can separate influence motives from social pressure effects in social signaling.

Irrespective of the experimental condition, all Receivers are equally informed that they will decide after their Sender but do not learn ex ante whether and when they will be informed about their partner's decision. This abstracts from any differential anticipated behavior among Receivers and allows for the estimation of a treatment effect on Receivers only caused by differences in Sender behavior. While Senders take their decision, all Receivers are directed to a waiting page where they remain for a maximum of 60 seconds before they can continue with the next question. Only then are Receivers in the *"informing partner before"* condition informed about the Sender's registration decision and can register themselves. In contrast, Receivers in the *"informing partner after"* condition directly move on to their own registration decision and are only informed about their partners' decision thereafter.

5. First stage Subsequently, we ask Senders how likely they think it is that they can influence their partner's registration decision using a slider ranging from 0 to 100. We use this metric to obtain an estimate of the first-stage effect of our manipulations on Senders' beliefs about their impact on their partner's registration decision.

6. Main outcome Next, we elicit our main outcome by asking participants whether they wished to sign up for a COVID-19 vaccination right away. If participants answered "yes", they were forwarded to the official registration website (BayIMCO) outside of our survey.[28] Participants who responded "no" were forwarded to the next module of our survey. On average, it took participants in our experiment five to six minutes to complete the online registration form. Once participants completed the form, they obtained an email from BayIMCO officially confirming their registration. We use this confirmation email to verify whether participants indeed registered for a vaccination by asking them to enter the sending address and the subject line in a survey form. For this task, we incentivized participants by informing them that by reporting both pieces of information correctly they would qualify for one of 30 additional €20 Amazon vouchers. Once participants had entered their information, their responses were checked by our system. If both answers were correct, a lottery determined whether participants obtained one of the Amazon vouchers. Participants only learned whether they had won any of the Amazon vouchers after they had answered all questions, i.e., on the final page of the survey.

[28] In Appendix 1.D.6 we provide screenshots illustrating how we elicited and confirmed whether participants signed up for a COVID-19 vaccination via BayIMCO.

The timing of the steps we used to elicit whether participants actually signed up for the vaccination is crucial in this context: when we offered participants the opportunity to sign up for the vaccination, participants did not know that we would ask them to provide proof of their registration. We informed participants about the confirmation and the corresponding remuneration only after they had reported to us that they successfully completed the registration. Hence, participants did not have an incentive to misreport their registration in order to qualify for one of the vouchers. One may still worry that participants misreporting their registration status tried to find out the address and the subject line of the confirmation email to nevertheless qualify for one of the vouchers. It is, however, very unlikely that participants successfully managed to cheat, since the address from which the confirmation email was sent changed over time. Even if participants found a screenshot of the confirmation email by searching the internet, the screenshot had to be fairly recent to keep up with the changes of the confirmation email over time.

7. Further outcomes Finally, we collect post-treatment attitudes and beliefs related to the COVID-19 vaccination, including participants' stated willingness to ever take a COVID-19 vaccine alongside with their beliefs regarding the safety and efficacy of the vaccine, its social desirability, and associated free-riding problems.[29] In addition, we collect further demographic information including income, education, county, and zip code of residence. On the final page of the survey, we reveal payoffs to participants and provide them with the opportunity to comment on the survey.

1.A.4 Additional steps taken to identify influence motives

In order to identify influence motives in social signaling, our design aims to maximize the difference in Senders' beliefs about their influence on their partner's registration decision between the "*informing partner before*" and "*informing partner after*" conditions. To achieve this, we designed both the decision Senders take as well as the interaction with their partner to be *"one-shot"*. To ensure that the interaction is one-shot in nature, we paired individuals who had likely not met before and upheld anonymity throughout the experiment. Anonymity facilitates identifying influence motives as it limits Senders' chances of influencing their partner to that particular encounter: Senders in the *"informing partner before"* condition should realize that their opportunity to influence their partner's decision is either now, by sending a signal in the experiment, or never. Of course, Senders' decisions within the experiment may influence Receivers' behavior after the experiment has ended, as Receivers may contemplate their partner's decision in the experiment for a while and register for a vaccination at some later point in time. In principle, Senders in the *"informing partner after"* condition may realize as

[29] We also collect the same set of beliefs before the treatment to analyze within-individual changes arising from the treatment.

well that their actions during the experiment might influence Receivers' behavior *after* the experiment. If that was the case, influence motives would also encourage Senders in this condition, potentially attenuating behavioral differences relative to the *"informing partner before"* condition.

Moreover, the fact that the decision itself – and thus its potential externality on the Receiver – is one-shot, may render it more salient from the perspective of the Sender. Combined, the one-shot decision and the one-shot interaction help us identify influence motives. The role of these design features also suggests a reason why Karlan and McConnell (2014) – who used a similar set of experimental manipulations – did not find evidence for influence motives: to conduct their experiment, they recruited participants from the same peer group (college students from the same university). As a result, Senders might have already known Receivers and anticipated to meet them again in the future, reducing the relative importance of the signal sent within the experiment. A similar logic applies to the decision they studied: they asked Senders to decide about a donation to a university institution, a decision which Senders could take multiple times in the future.

We opted for a setting with limited scope for social pressure to arise, which is likely conducive to identifying influence motives – in line with Foerster and van der Weele (2021) who suggest that influence motives and social pressure effects may be substitutes. In our experiment, social pressure effects are potentially muted by three factors: first, Senders interact with strangers on the internet instead of neighbors (Bursztyn, González, and Yanagizawa-Drott 2020) or classmates (Bursztyn and Jensen 2015). Second, the number of observing Receivers a Sender is facing is smaller than in many existing studies (e.g. Perez-Truglia and Cruces 2017). Third, by upholding anonymity throughout the entire experiment, we shut down most instrumental motives underlying social pressure effects arising from potential future encounters with the Receiver.[30] However, while being muted, a recent study by Goette and Tripodi (2021) shows that social pressure effects can still play a role even in quasi-anonymous online settings like ours. Due to this we chose a design holding social pressure constant between the key experimental conditions and allowed for the possibility that social pressure effects might arise in formulating our pre-analysis plan.

1.A.5 Experimental assignment and sample balancing

We used a two-stage random procedure to assign participants into experimental conditions: first, we randomly assigned teams to one of the three experimental conditions *"not informing partner"*, *"informing partner after"*, or *"informing partner before"*. Second, within those teams, we further randomized who was assigned the role of Sender and

[30] For a discussion of the distinction between instrumental and hedonic motives underlying social-image effects see Bursztyn and Jensen (2017).

Receiver, respectively. We report the resulting assignment into experimental conditions in Appendix Table 1.C.2. The discrepancy between the number of participants in the *"informing partner after"* and *"informing partner before"* conditions is an artefact of the specific randomization procedure used. We used "on the fly" randomization to assign participants into experimental conditions as they entered the survey. Due to the random nature of the assignment process, the effective share in each condition slightly deviates from the target shares we specified in our pre-analysis plan.

Since we are primarily interested in Senders' decisions, we opted for an implementation using fewer Receivers than Senders. Hence, some pairs were formed of two Senders rather than a Sender and a Receiver. To employ only factually true experimental instructions, Senders in the *"informing partner after"* and the *"informing partner before"* conditions were informed that their registration decision *may* be shared with their partner.[31] To further reduce the number of Receivers in our experiment, pairs in the *"not informing partner"* condition always consisted of two Senders. Since Senders' decisions in this condition were not shared with their partner anyway, these Senders' partners could also be other Senders while relying on factually true information throughout.

To assess whether Senders' predetermined characteristics are balanced across experimental conditions, we conducted pairwise comparisons of 21 predetermined characteristics across all three conditions using the following regression model:

$$characteristic_i = \alpha + \beta \cdot treat_i + \epsilon_i,$$

where $treat_i$ is a dummy variable corresponding to either the *"informing partner after"* or the *"informing partner before"* condition, and where we omit one condition from our sample for every pair-wise comparison. In Table 1.C.4, we report the group means separately for each condition alongside the p-values obtained from these regressions. Out of the 63 estimates reported in Table 1.C.4, only one is significant at the 5-percent level, suggesting that Senders' predetermined characteristics are well balanced across experimental conditions. This finding is further supported by the p-values obtained from tests for joint significance of all predetermined characteristics reported at the bottom of Table 1.C.4.

[31] To be precise, we informed Senders that their partner would learn about their decision only "with high probability".

1.B Additional Figures

Figure 1.B.1: Treatment messages and corresponding implementation in the survey

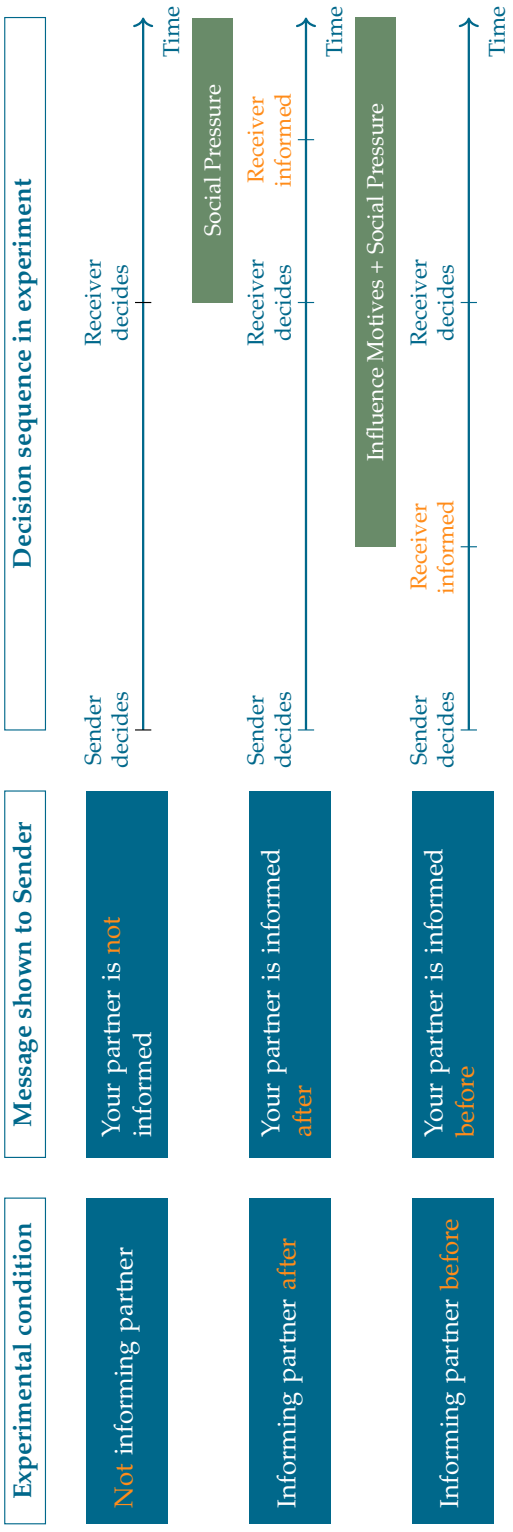
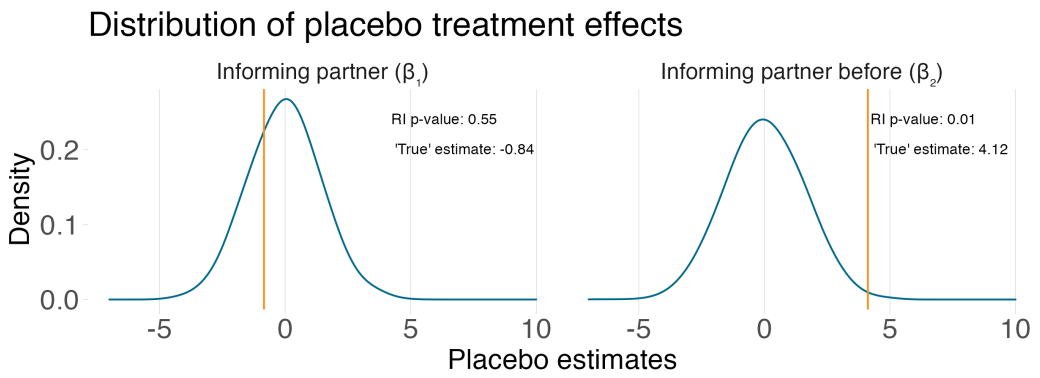
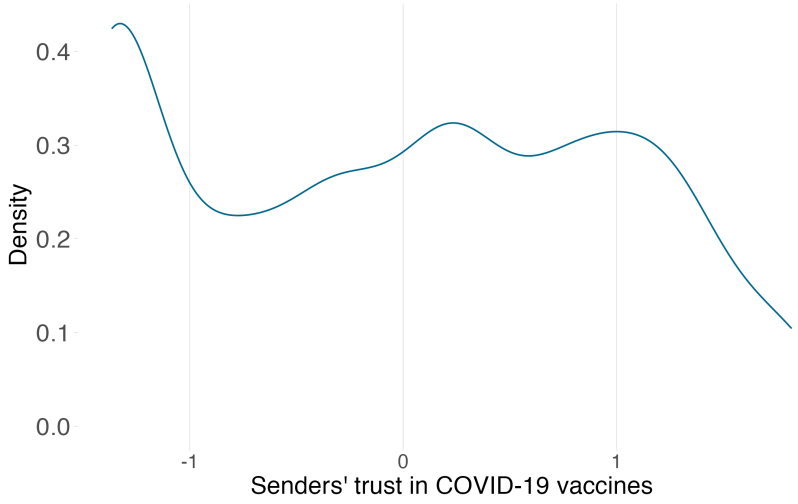


Figure 1.B.2: Results from randomization inference



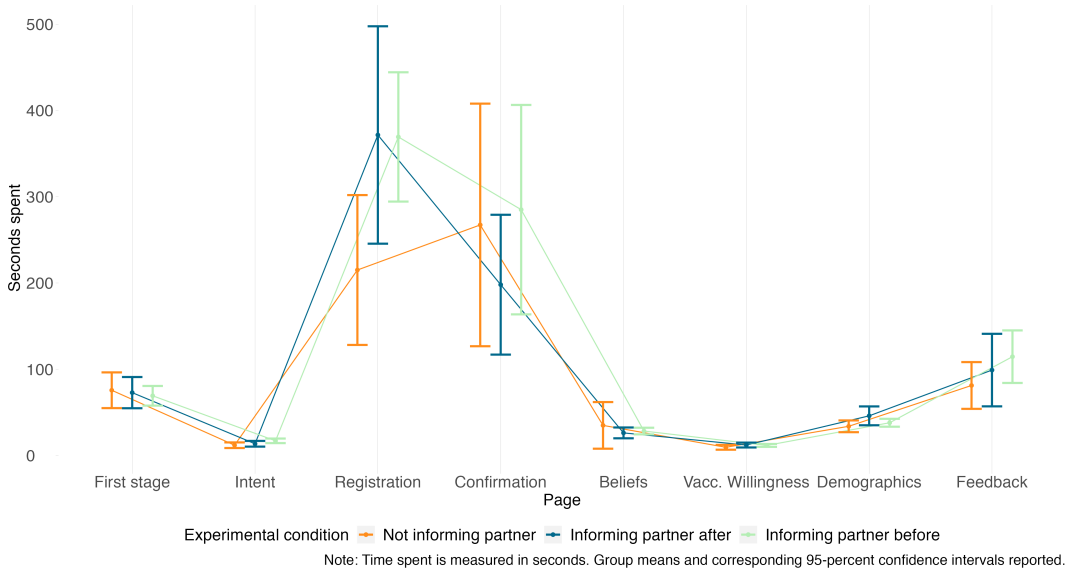
Notes: Distribution of placebo estimates derived from randomly re-assigning Senders to placebo treatment groups over 5,000 iterations and calculating the share of "placebo treatment effects" that exceed the "true treatment effect" in (absolute) magnitude. Panel (a) reports the resulting distribution and Fisher exact p-value for coefficient β_1 based on Equation 1 and Panel (b) for coefficient β_2, respectively. The outcome in both panels is Senders' verified registration status.

Figure 1.B.3: Distribution of Senders' trust in COVID-19 vaccines



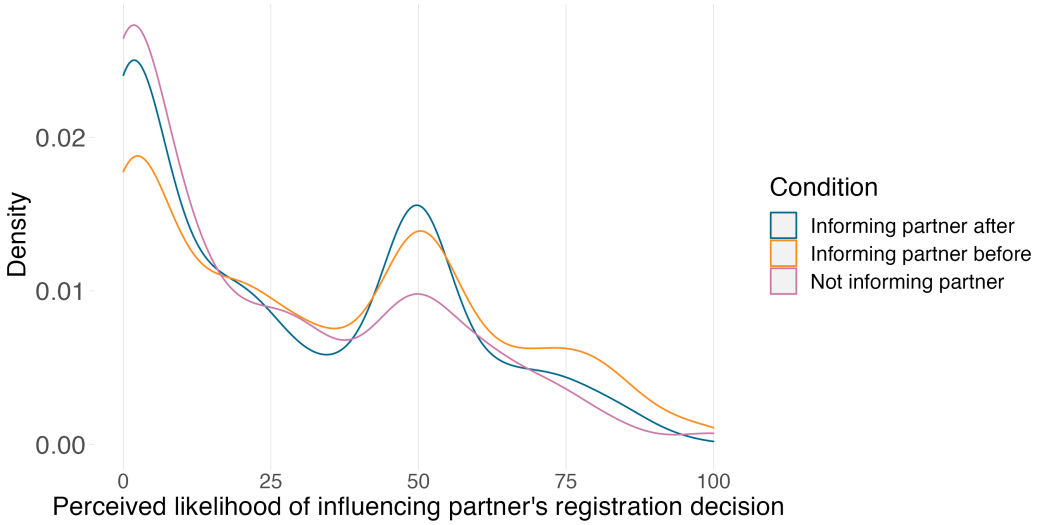
Notes: Distribution of Senders' trust in COVID-19 vaccines (scaled to mean = 0 and sd = 1). We measure trust in vaccine quality as the standardized mean of two survey items elicited pre-treatment capturing Senders' beliefs about the safety and efficacy of COVID-19 vaccines, respectively.

Figure 1.B.4: Time spent on each page post treatment by experimental condition



Notes: Senders' mean time spent on all survey pages after the treatment module alongside 95-percent confidence intervals by experimental condition. Time spent on each page is measured in seconds. The sample of Senders is limited to those who could provide proof of their registration.

Figure 1.B.5: Distribution of Senders' beliefs about their influence on their partner



Notes: Density plot of Senders' beliefs about the likelihood of influencing their partners' registration decision, reported separately by experimental condition.

1.C Additional Tables

Table 1.C.1: Summary statistics for full sample (Senders and Receivers)

Statistic	Mean	St. Dev.	Min	Max	N
Demographics					
Age	40.90	14.35	18.00	79.00	1,857
Female (%)	54.55	49.81	0.00	100.00	1,857
Monthly net household income (in 1,000 €)	2.91	1.60	1.10	7.50	1,857
Upper secondary degree (%)	38.50	48.67	0.00	100.00	1,857
Local characteristics					
Mean incidence rate (2nd wave)	138.73	40.67	65.64	301.07	1,857
Population in zip (in 1,000)	14.81	9.85	0.60	48.05	1,857
Lives in urban area (¿=100,000 inhab.)	29.46	45.60	0.00	100.00	1,857
Turnout in 2017	77.54	4.30	59.90	90.20	1,857
AfD vote share in 2017	12.22	3.06	5.49	26.42	1,857
Unemployment rate (%)	2.35	0.93	0.05	5.50	1,857
Attitudes and beliefs about COVID-19 vaccines					
Safety of vaccines, pre	0.00	1.00	−1.23	1.83	1,857
Efficacy of vaccines, pre	0.00	1.00	−1.42	1.65	1,857
Social desirability of vaccination, pre	0.00	1.00	−1.17	1.50	1,857
Likelihood to freeride on others' vaccination decision, pre	0.00	1.00	−1.09	1.81	1,857
Estimated willingness to take vaccine in state (%)	59.11	20.16	0.00	100.00	1,857
Preferences					
Own willingness to take vaccine (%)	51.31	37.09	0.00	100.00	1,857
Altruism	0.00	1.00	−2.36	2.70	1,857
Desire to influence others	0.00	1.00	−3.03	1.67	1,857
Social image concerns	0.00	1.00	−1.84	2.31	1,857
Social proximity					
Oneness	0.00	1.00	−1.06	2.25	1,526

Notes: All variables classified as "local characteristic" do not vary on the individual but on the zip code or municipality ("Gemeinde") of residence level.

Table 1.C.2: Number of Senders and Receivers assigned to each condition

Condition	Role	Observations
Not informing partner	Sender	328
Informing partner after	Sender	554
Informing partner before	Sender	519
Informed after	Receiver	236
Informed before	Receiver	220

Table 1.C.3: Description of all variables

Variable	Type	Answer Options	Min	Max	Source	Survey Item	Notes
Demographics							
Age	Numeric				Survey	How old are you?	
Female	Categorical	male, female	0	1	Survey	Are you male or female?	Dummy == 1 if participant is female
Monthly net household income (in 1,000 €)	Categorical	seven income brackets			Survey	What was your household's monthly net income last year?	
Upper-secondary degree	Categorical	primary, secondary, upper-secondary etc.	0	1	Survey	What is your highest educational degree (general or vocational)?	Dummy == 1 if participant has an upper secondary degree (Abitur)
Local characteristics							
Mean incidence rate (2nd wave)	Numeric				Administrative		Administrative COVID-19 incidence data from Robert Koch Institut (2022)
Population in zip (in 1,000)	Numeric				Administrative		Population data from Suche-Postleitzahl.org (2022)
Lives in urban area ($\geq$ 100,000 inhab.)	Categorical				Administrative		Dummy == 1 if municipality has more than 100,000 inhabitants
Turnout in 2017	Numeric				Administrative		Election data from Statistische Ämter des Bundes und der Länder (2022)
AfD vote share in 2017	Numeric				Administrative		
Unemployment rate (in %)	Numeric				Administrative		Unemployment data from Bayerisches Landesamt für Statistik (2022)
Attitudes and beliefs about COVID-19 vaccines							
Safety of vaccines	Categorical		1	7	Survey	To what extent do you agree with the following statement? I have full confidence that the vaccination against COVID-19 is safe.	Attitudes and beliefs items based on COSMO (2021)
Efficacy of vaccines	Categorical		1	7	Survey	To what extent do you agree with the following statement? I have full confidence that the vaccination against COVID-19 is effective.	
Social desirability of vaccines	Categorical		1	7	Survey	To what extent do you agree with the following statement? I view vaccinations as a collective effort against the spread of COVID-19.	
Likelihood to freeride on others' vaccination decision	Categorical		1	7	Survey	To what extent do you agree with the following statement? If everyone is vaccinated against COVID-19, I don't need to get vaccinated too.	
Estimated willingness to take vaccine in state	Numeric		0	100	Survey	What do you think? What proportion of people in Bavaria are willing to get vaccinated against COVID-19?	
Preferences							
Own willingness to take vaccine	Numeric		0	100	Survey	How likely is it that you will get vaccinated against COVID-19?	
Altruism	Numeric		0	10	Survey	How much would you be willing to give to a good cause without expecting anything in return?	Combined both survey items as described in Falk et al. (2018)

Notes: All variables classified as "local characteristic" do not vary on the individual but on the zip code or municipality ("Gemeinde") of residence level.

(continued on next page)

Variable	Type	Answer Options	Min	Max	Source	Survey Item	Notes
Altruism	Numeric		0	1000	Survey	Today you unexpectedly received 1,000 EUR. How much of the money would you donate to a good cause?	
Desire to influence	Categorical		1	7	Survey	How well does the following statement apply to you as a person? I like it when people accept my suggestions.	Scaled, unweighted average of all three survey items from that category; items based on Bennett (1988)
Desire to influence	Categorical		1	7	Survey	How well does the following statement apply to you as a person? I like it when my ideas and opinions have an impact on other people.	
Desire to influence	Categorical		1	7	Survey	How well does the following statement apply to you as a person? I would like the feeling of having influenced other people's lives.	
Social image concerns	Categorical		1	7	Survey	How well does the following statement apply to you as a person? It is important to me to impress others.	Scaled, unweighted average of all three survey items from that category; items based on Alba et al. (2014)
Social image concerns	Categorical		1	7	Survey	How well does the following statement apply to you as a person? I think a lot about whether I am good enough compared to others	
Social image concerns	Categorical		1	7	Survey	How well does the following statement apply to you as a person? It is important to me how I am perceived by others	
Social proximity							
Oneness	Categorical		1	7	Survey	Which of the following figures best reflects how connected you feel with your partner?	Scaled, unweighted average of all survey items from that category; oneness item sources described in Appendix 1.A
Oneness	Categorical		1	7	Survey	To what extent would you refer to yourself and your partner [name] as "we"?	
Registration decision							
Verified registration	Categorical		0	1	BayIMCO		Dummy == 1 if registration could be verified
Self-reported registration	Categorical	yes, no	0	1	Survey	Have you successfully registered?	Dummy == 1 if replied "yes"
Clicked link forwarding to BayIMCO	Categorical		0	1	Survey		Dummy == 1 if clicked on link forwarding participant to BayIMCO.
Self-reported intent to register	Categorical	yes, no	0	1	Survey	Would you like to register for a COVID-19 vaccination?	Dummy == 1 if replied "yes"
First stage belief							
Perceived likelihood of influencing partner's registration decision	Numeric		0	100	Survey	What do you think? How likely is it that your decision to register or not to register will influence your partner's decision?	

Notes: All variables classified as "local characteristic" do not vary on the individual but on the zip code or municipality ("Gemeinde") of residence level.

Table 1.C.4: Senders' predetermined characteristics compared across experimental conditions

	Group means			Test for equal means: p-values			N
	Before	After	Not	Before vs. After	Before vs. Not	After vs. Not	
Attrition							
Completed survey (in %)	73.20	73.57	76.28	0.87	0.24	0.30	1,892
Demographics							
Age	40.67	41.36	40.43	0.43	0.82	0.36	1,401
Female (in %)	56.07	53.61	51.83	0.42	0.23	0.61	1,401
Monthly net household income (in 1,000 €)	2.85	2.85	2.99	0.97	0.21	0.21	1,401
Upper secondary degree (in %)	37.38	39.71	40.24	0.43	0.41	0.88	1,401
Local characteristics							
Mean incidence rate (2nd wave)	138.48	140.53	137.02	0.41	0.60	0.22	1,401
Population in zip (in 1,000 inhabitants)	14.21	15.17	14.91	0.11	0.29	0.71	1,401
Lives in urban area ($\geq$ 100,000 inhabitants)	28.90	31.77	31.40	0.31	0.44	0.91	1,401
Turnout (%)	77.54	77.60	77.50	0.81	0.91	0.75	1,401
AfD vote share (%)	12.23	12.18	12.17	0.78	0.78	0.98	1,401
Unemployment rate (%)	2.34	2.40	2.36	0.33	0.81	0.53	1,401
Attitudes and beliefs about COVID-19 vaccines							
Safety of vaccines	-0.02	-0.01	0.01	0.80	0.69	0.86	1,401
Efficacy of vaccines	-0.01	-0.02	-0.01	0.84	0.99	0.87	1,401
Social desirability of vaccination	-0.02	0.01	0.00	0.69	0.81	0.92	1,401
Likelihood to freeride on others' vaccination decision	0.02	-0.08	0.07	0.12	0.46	0.03**	1,401
Estimated willingness to take vaccine in state (%)	58.37	58.41	59.83	0.97	0.30	0.30	1,401
Preferences							
Own willingness to take vaccine (%)	50.78	51.40	49.57	0.78	0.65	0.48	1,401
Altruism	-0.02	0.04	-0.02	0.34	0.97	0.39	1,401
Desire to influence others	0.02	-0.05	0.02	0.24	0.99	0.30	1,401
Social image concerns	0.01	0.00	-0.01	0.77	0.80	0.99	1,401
Social proximity							
Oneness	-0.05	0.05	-0.06	0.13	0.92	0.14	1,140
Test for joint significance				0.71	0.92	0.42	

Notes: Group means of Senders' predetermined characteristics alongside p-values testing for equal means reported. p-values are derived from the following regressions comparing predetermined characteristics between pairs of conditions: $characteristic_i = \alpha + \beta \cdot treat_i + \epsilon_i$, where $treat_i$ is a dummy variable corresponding to either the "*informing partner after*" or the "*informing partner before*" condition, and where we omit one condition from our sample for every pair-wise comparison. *Not* refers to the *not informing partner* condition. All variables classified as "local characteristic" do not vary on the individual but on the zip code or municipality ("Gemeinde") of residence level. Significance levels: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 1.C.5: What predicts Senders' registration status?

	Verified registration																			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Age	0.31																			0.71
	(0.58)																			(0.61)
Female		−0.24																		0.60
		(0.61)																		(0.63)
Monthly net household income (in 1,000 €)			1.01																	0.52
			(0.61)																	(0.62)
Upper secondary degree (in %)				0.88																0.07
				(0.63)																(0.68)
Mean incidence rate (2nd wave)					−1.14**															0.04
					(0.48)															(0.60)
Population in zip (in 1,000)						−0.58														0.09
						(0.62)														(0.75)
Lives in urban area ($\geq$ 100,000 inhab.)							−1.33**													−1.49*
							(0.55)													(0.84)
Turnout in 2017								1.82***												1.37
								(0.64)												(1.03)
AfD vote share in 2017									−0.44											−0.14
									(0.62)											(0.93)
Unemployment rate (in %)										−1.59***										0.34
										(0.56)										(1.12)
Safety of COVID-19 vaccines											5.62***									3.31***
											(0.68)									(1.14)
Efficacy of COVID-19 vaccines												4.95***								−1.32
												(0.63)								(1.04)
Social desirability of COVID-19 vaccines													5.18***							0.02
													(0.69)							(1.23)
Likelihood to freeride on others' vacc. decision														−3.34***						−0.15
														(0.56)						(0.53)
Estimated willingness to take vaccine in state (%)															3.29***					0.61
															(0.60)					(0.59)
Own willingness to take vaccine (%)																5.98***				3.87***
																(0.67)				(1.14)
Altruism																	0.17			−0.71
																	(0.54)			(0.57)
Desire to influence others																		2.41***		1.72**
																		(0.61)		(0.68)
Social image concerns																			1.02*	−0.83
																			(0.54)	(0.63)
Observations	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401	1,401
R^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.06	0.05	0.05	0.02	0.02	0.07	0.00	0.01	0.00	0.09

Notes: Results from regressions of the following type reported in columns 1 to 19: $y_i = \alpha + \beta \cdot x_i + \epsilon_i$, where y_i denotes Sender i's verified registration status and x_i is one of Sender i's predetermined characteristics. In column 20, we assess the predictive power of all predetermined characteristics jointly in one regression. Significance levels: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 1.C.6: What predicts perceived social proximity between partners?

	Social proximity				
	(1)	(2)	(3)	(4)	(5)
Same gender	0.09				0.10
	(0.06)				(0.07)
Same educational level		0.01			0.01
		(0.07)			(0.07)
Absolute age difference			-0.06^{**}		-0.05
			(0.03)		(0.03)
Absolute income difference				0.03	0.03
				(0.03)	(0.03)
Mean, social proximity	0.00	0.00	0.00	0.00	0.00
SD, social proximity	1.00	1.00	1.00	1.00	1.00
Observations	1,140	959	1,140	959	959
R^2	0.00	0.00	0.00	0.00	0.01

Notes: Results from regressions of the following type reported in columns 1 to 4: $y_i = \alpha + \beta \cdot x_i + \epsilon_i$, where y_i denotes Sender i's perceived social proximity between herself and her partner in the experiment. x_i is either a dummy taking value 1 if Sender i shares this predetermined characteristics with her partner and zero otherwise or the absolute difference between Sender i's response and her partner's response. In column 5, we employ all characteristics jointly in the same regression. Significance levels: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 1.C.7: All outcomes by experimental conditions

	Observations			Group means			Test for equal means: p-values		
	Before	After	Not	Before	After	Not	Before vs. After	Before vs. Not	After vs. Not
Registration decision									
Verified registration status (in %)	519	554	328	7.90	3.79	4.57	0.00***	0.04**	0.58
Self-reported registration status (in %)	519	554	328	13.87	11.01	9.45	0.16	0.05*	0.46
Clicked registration link (in %)	519	554	328	11.18	8.84	8.84	0.20	0.26	1.00
Self-reported intent to register (in %)	519	554	328	23.31	21.66	19.82	0.52	0.23	0.51
Attitudes towards COVID-19 vaccination									
Self-reported willingness to take vaccine (in %)	519	554	328	49.69	50.46	49.37	0.73	0.90	0.67
Changes in beliefs pre-post treatment									
Safety of vaccine (in sd)	519	554	328	0.03	0.00	-0.05	0.73	0.28	0.44
Efficacy of vaccine (in sd)	519	554	328	0.00	0.01	-0.01	0.92	0.87	0.80
Social desirability of vaccination (in sd)	519	554	328	0.02	0.00	-0.04	0.80	0.42	0.56
Likelihood to freeride of others' vaccination decision (in sd)	519	554	328	-0.02	0.01	0.01	0.59	0.63	1.00
First stage belief									
Perceived likelihood of influencing partner's registration decision (in %)	519	554	328	33.51	27.93	24.36	0.00***	0.00***	0.07*

Notes: Raw means conditional on treatment status reported for full sample of Senders. All changes in beliefs (pre/post treatment) were standardized to exhibit mean = 0 and sd = 1 in the full sample of Senders. Significance levels: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 1.C.8: Treatment effects on changes in Senders' beliefs and attitudes

	Δ Beliefs safety of vaccine	Δ Beliefs efficacy of vaccine	Δ Beliefs social desirability of vaccination	Δ Likelihood to freeride on others' vaccination decision
	(1)	(2)	(3)	(4)
Informing partner	0.05	−0.01	0.08	−0.00
	(0.07)	(0.07)	(0.06)	(0.07)
Informing partner before	0.02	0.00	−0.00	−0.02
	(0.06)	(0.06)	(0.05)	(0.06)
Controls	✓	✓	✓	✓
Mean, 'Not informing partner'	-0.05	-0.01	-0.04	0.01
Mean, 'Informing partner after'	0.00	0.01	0.00	0.01
Observations	1,401	1,401	1,401	1,401
R^2	0.05	0.08	0.38	0.02

Notes: Results derived from regressions as laid out in Equation 1. We employ the following changes in attitudes (pre/post treatment) as dependent variables: (column 1) beliefs about the safety of the vaccine; (column 2) beliefs about the efficacy of the vaccine; (column 3) beliefs about the social desirability of taking the vaccine; and (Column 4) self-reported tendency to freeride on others' commitment to take the vaccine. All changes in beliefs were standardized to obtain mean = 0 and sd = 1 in the full sample of Senders. Controls include the full set of variables reported in Appendix Table 1.C.1 with the exception of social proximity and the pre-treatment levels of these attitudes and beliefs. Robust standard errors reported in parentheses. Significance levels: $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$.

Table 1.C.9: Strategic lying

	Self-reported intent to register	Verified registration	Self-reported intent NOT verified
	(1)	(2)	(3)
Informing partner	1.18	−0.84	2.02
	(2.53)	(1.39)	(2.43)
Informing partner before	2.29	4.12***	−1.83
	(2.30)	(1.37)	(2.16)
Controls	✓	✓	✓
Mean, 'Not informing partner'	19.82	4.57	15.24
Mean, 'Informing partner after	21.66	3.79	17.87
Observations	1,401	1,401	1,401
R^2	0.21	0.09	0.13

Notes: Results derived from regressions as laid out in Equation 1. We employ the following dependent variables: (column 1) dummy variable taking value 1 if a Sender reported to be willing to register (elicited before verification); (column 2) a dummy variable taking value 1 if a Sender reported that she registered for a vaccination and could provide proof of her registration; (column 3) a dummy variable taking value 1 if a Sender reported that she had signed up but failed to provide proof of her registration. Controls include the full set of variables reported in Appendix Table 1.C.1 with the exception of social proximity. Robust standard errors reported in parentheses. Significance levels: $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$.

Table 1.C.10: Treatment effects on Senders' first-stage beliefs

	Perceived likelihood that partner can be influenced (%)			
	(1)	(2)	(3)	(4)
Informing partner	6.26***	5.90***	3.58*	3.31*
	(1.77)	(1.73)	(1.94)	(1.90)
Informing partner before			5.57***	5.35***
			(1.79)	(1.69)
Controls	-	✓	-	✓
Mean, 'Not informing partner'	24.36	24.36	24.36	24.36
Mean, 'Informing partner after'	27.93	27.93	27.93	27.93
Observations	1,194	1,194	1,194	1,194
R^2	0.01	0.11	0.02	0.12

Notes: Results derived from regressions as laid out in Equation 1 where we employ Senders' beliefs about how likely they can influence their partner's registration decision as the dependent variable. Controls include the full set of variables reported in Appendix Table 1.C.1 with the exception of social proximity. Robust standard errors reported in parentheses.

Table 1.C.11: Receivers' predetermined characteristics compared across experimental conditions

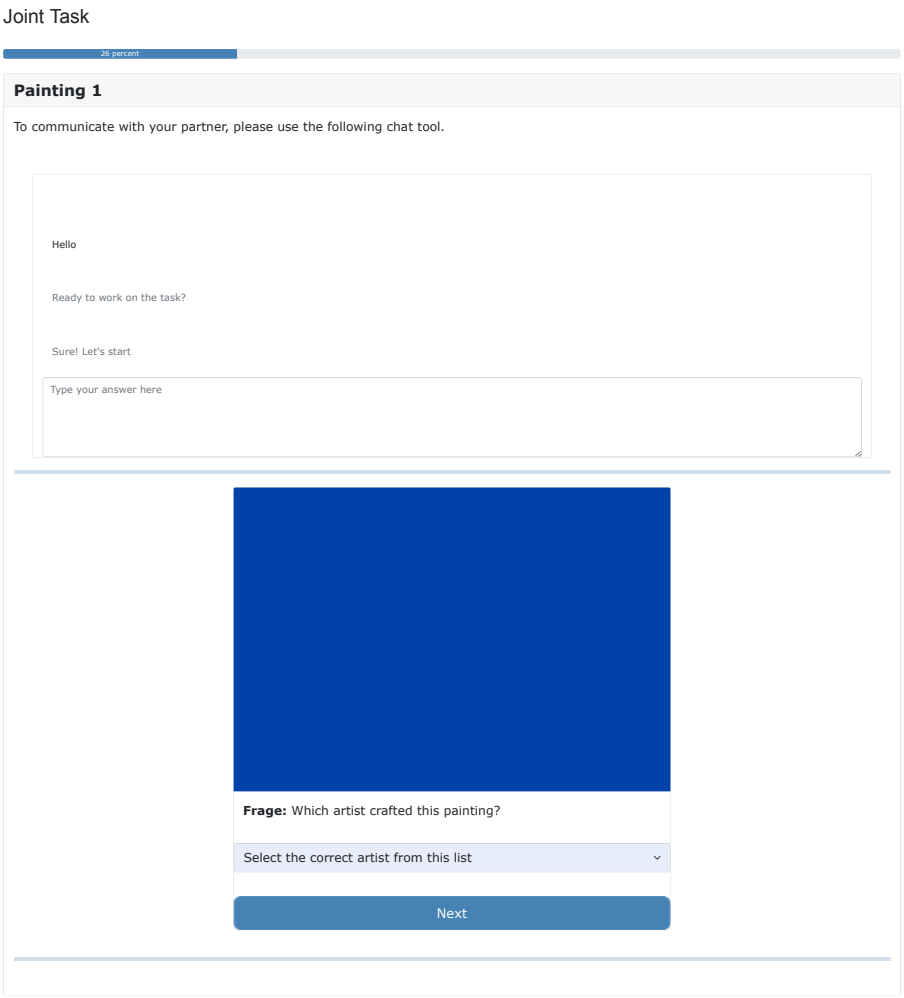
	Group means		p-value	
	Before	After	Before vs. After	N
Attrition				
Completed survey	73.09	70.66	0.50	635
Demographics				
Age	39.76	42.03	0.10	456
Female (%)	56.36	55.51	0.86	456
Monthly net household income (in 1,000 €)	2.86	3.10	0.12	456
Upper secondary degree (%)	36.82	37.29	0.92	456
Local characteristics				
Avg. incidence rate (2nd wave)	139.21	137.01	0.56	456
Population in zip (in 1,000 inhabitants)	15.30	14.70	0.52	456
Lives in urban area (¿=100,000 inhaitants.)	28.64	23.31	0.20	456
Turnout (%)	77.35	77.62	0.51	456
AfD vote share (%)	11.99	12.55	0.05*	456
Unemployment rate (%)	2.40	2.25	0.08*	456
Attitudes and beliefs about COVID-19 vaccines				
Safety of vaccines	0.05	0.01	0.68	456
Efficacy of vaccines	0.08	0.01	0.48	456
Social desirability of vaccination	0.05	-0.02	0.48	456
Likelihood to freeride on others' vaccination decision	0.05	0.00	0.59	456
Estimated willingness to take vaccine in state (%)	60.74	59.84	0.64	456
Preferences				
Own willingness to take vaccine (%)	53.22	52.91	0.93	456
Altruism	-0.02	-0.01	0.93	456
Desire to influence others	-0.03	0.07	0.31	456
Social image concerns	-0.03	0.02	0.60	456
Social proximity				
Oneness	0.13	-0.07	0.04**	386
Test for joint significance			0.51	

Notes: Group means of Receivers' predetermined characteristics alongside p-values testing for equality of means. P-values are derived from the following regressions: $\text{characteristic}_i = \alpha + \beta \cdot \text{informed before}_i + \epsilon_i$, where informed before_i is a dummy taking value 1 for all Receivers in the *informed before* condition. All variables classified as "local characteristic" do not vary on the individual but on the zip code or municipality ("Gemeinde") of residence level. Significance levels: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

1.D Screenshots

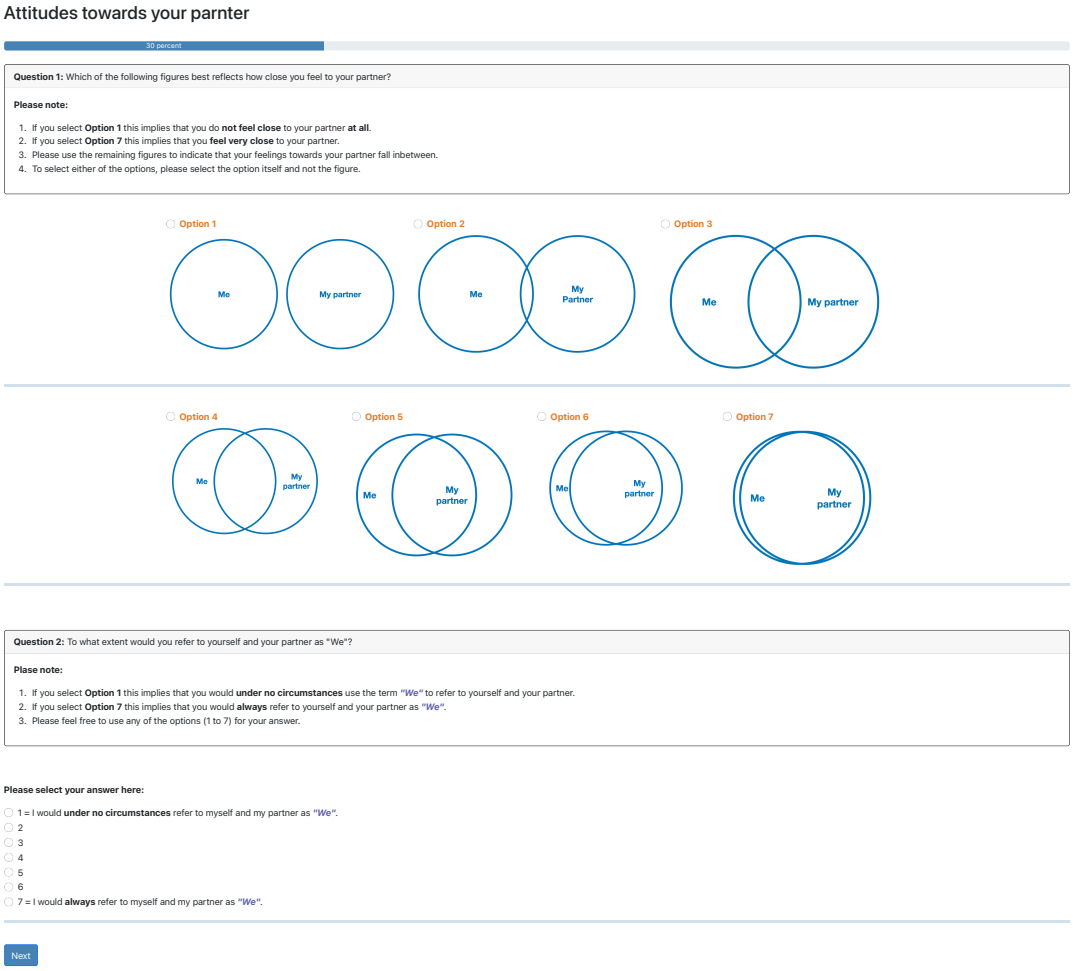
1.D.1 Joint problem solving task

Figure 1.D.1: Survey page showing chat window and historical painting (placeholder)



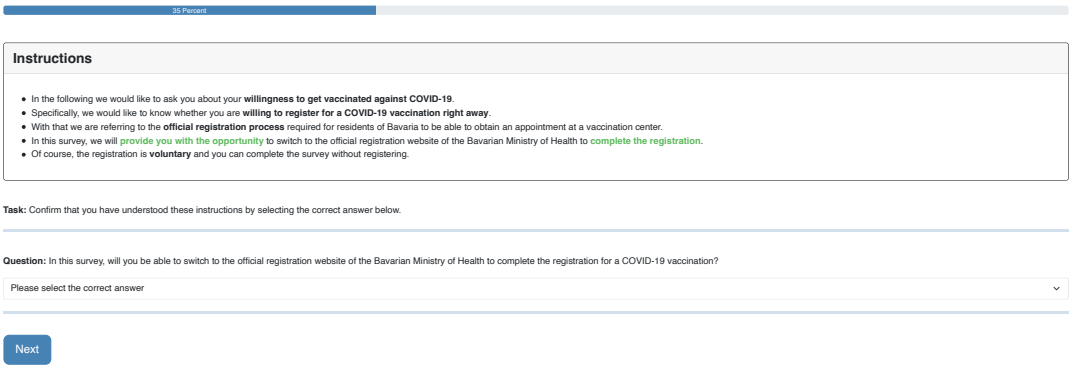
1.D.2 Oneness elicitation

Figure 1.D.2: Survey page documenting elicitation of social proximity



1.D.3 General instructions

Figure 1.D.3: Survey page providing general instructions on the opportunity to register for a COVID-19 vaccination



1.D.4 Treatment instructions – Senders

Figure 1.D.4: Survey page providing treatment instructions for Senders in the "not informing partner" condition

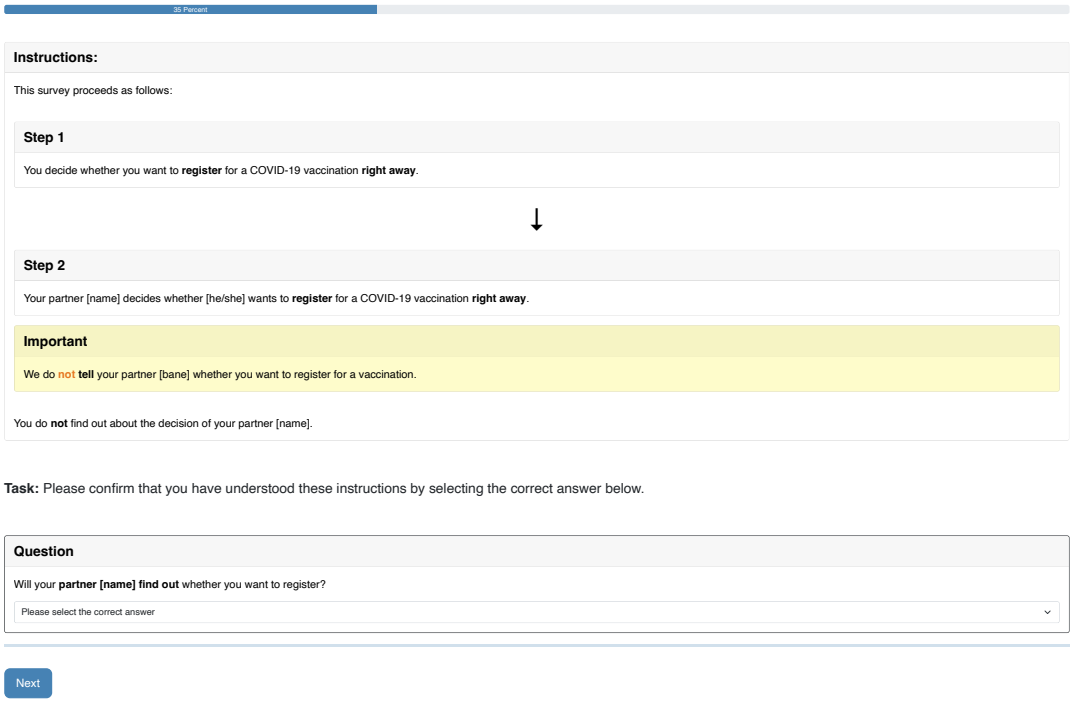


Figure 1.D.5: Survey page providing treatment instructions for Senders in the "informing partner after" condition

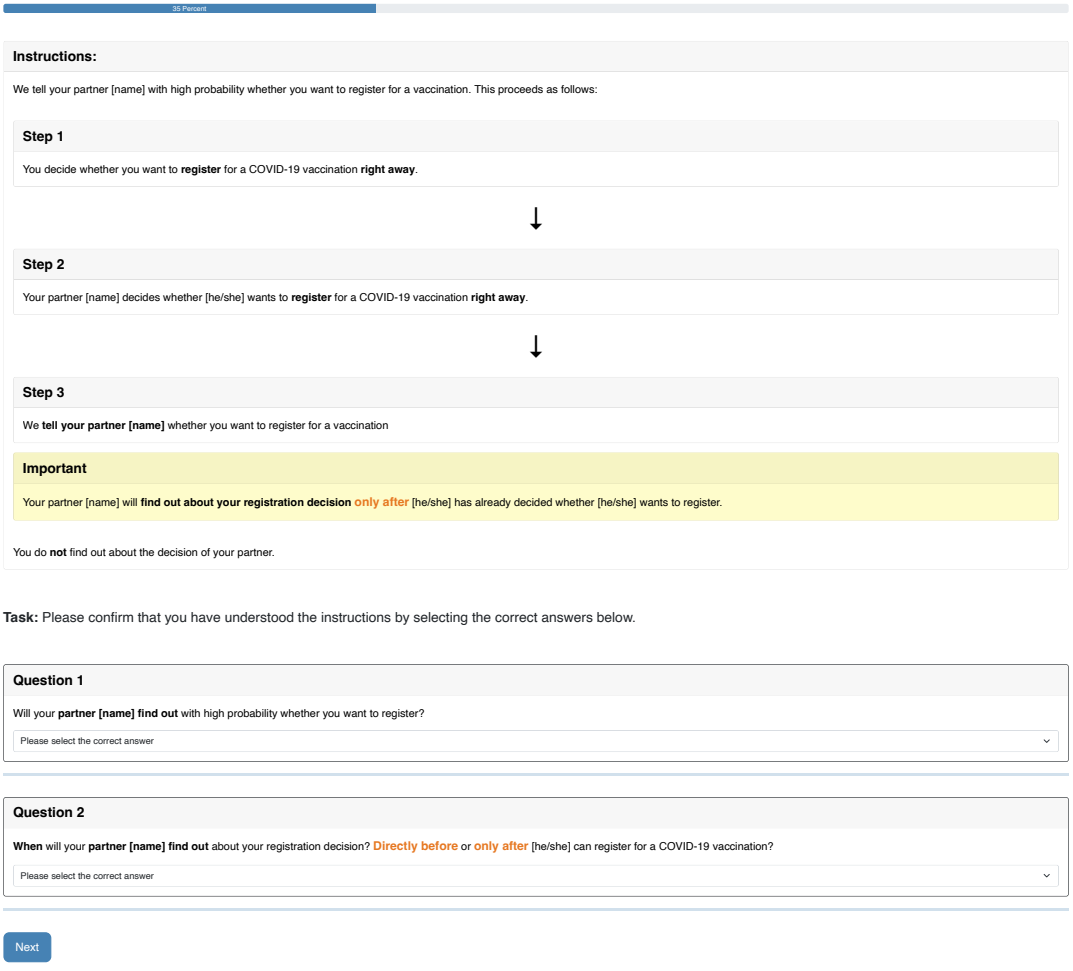
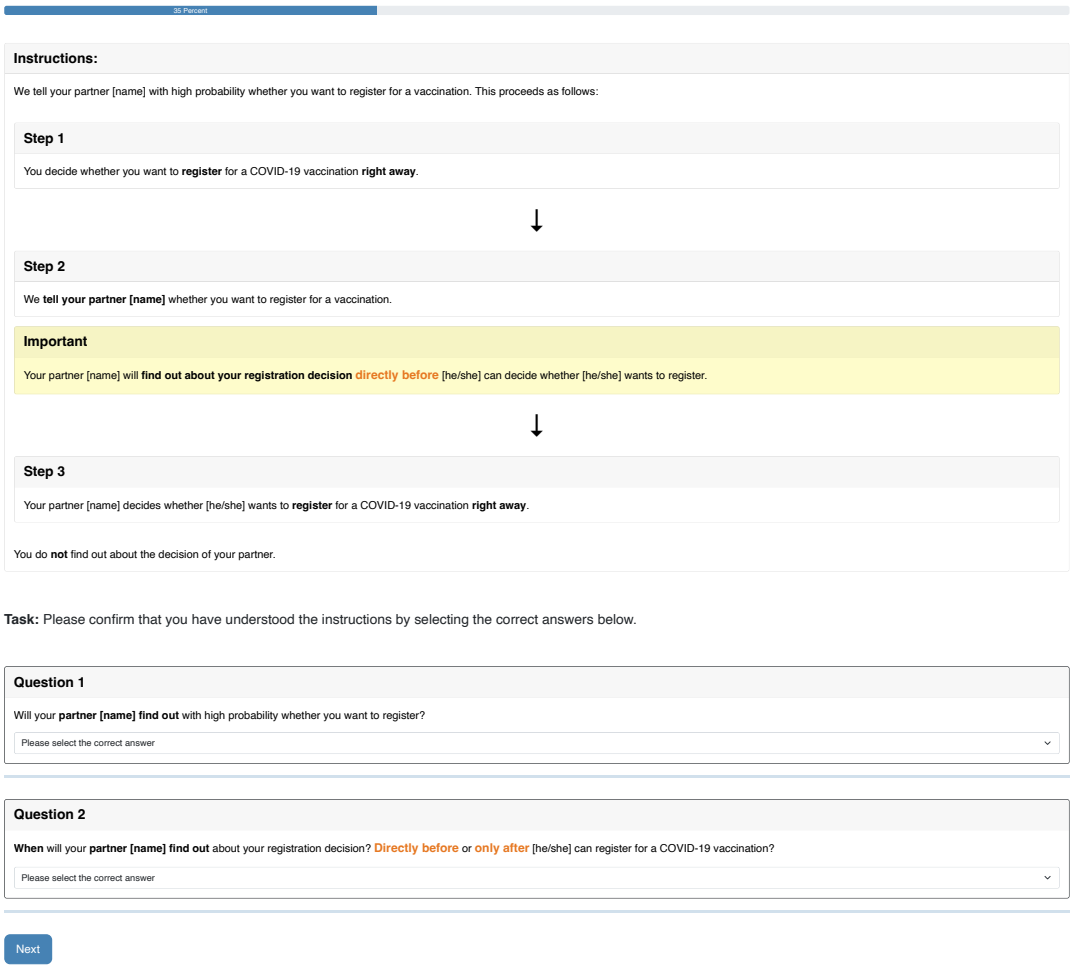
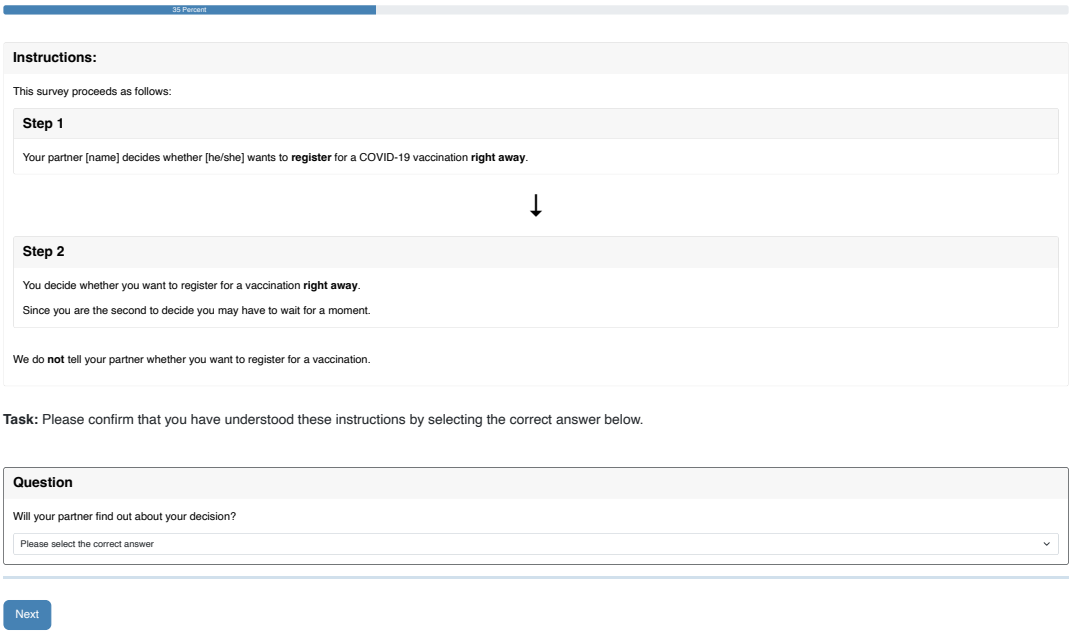


Figure 1.D.6: Survey page providing treatment instructions for Senders in the "informing partner before" condition



1.D.5 Treatment instructions – Receivers

Figure 1.D.7: Survey page providing treatment instructions for all Receivers



1.D.6 Registration for COVID-19 vaccination

Figure 1.D.8: Survey page eliciting intended willingness to register and providing link to official registration website (BayIMCO)

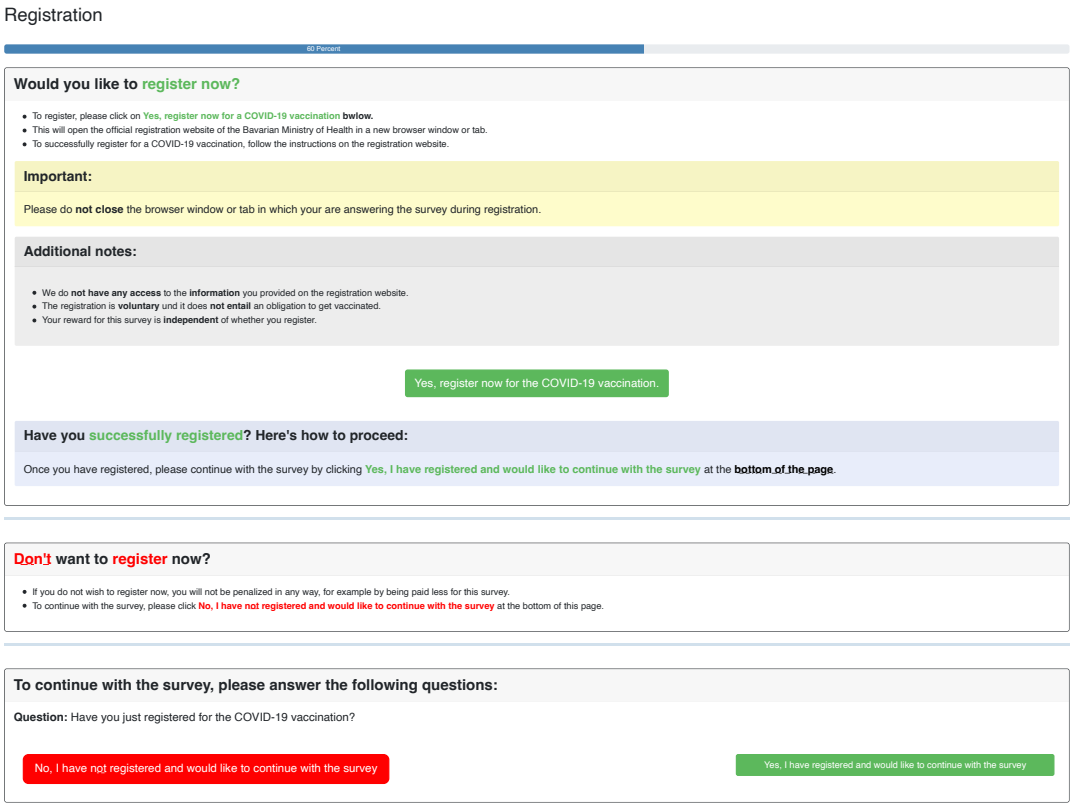


Figure 1.D.9: Landing page of the official registration website (BayIMCO)

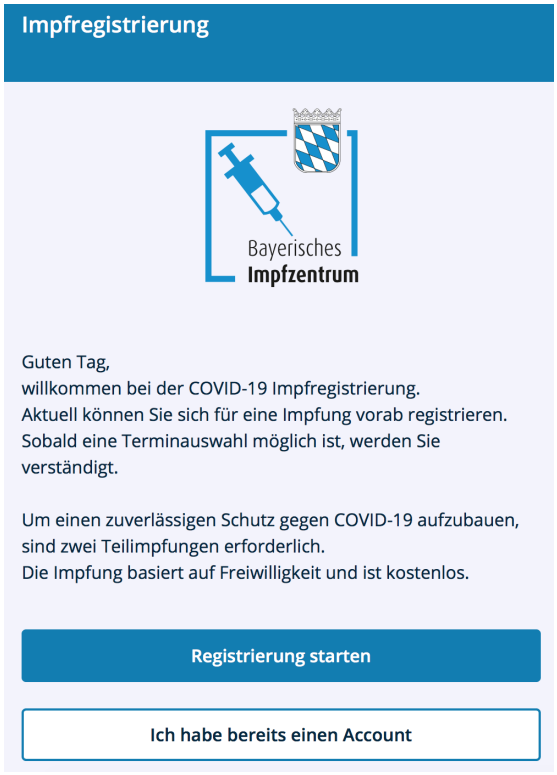
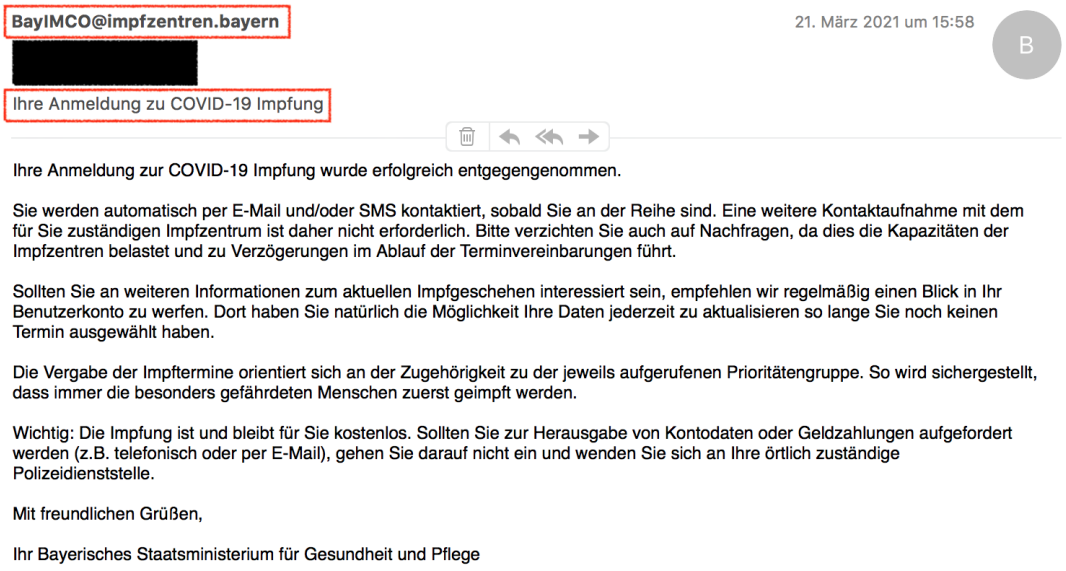


Figure 1.D.10: Confirmation email highlighting sending address and subject line



BayIMCO@impfzentren.bayern

[redacted]

21. März 2021 um 15:58

B

Ihre Anmeldung zu COVID-19 Impfung

Ihre Anmeldung zur COVID-19 Impfung wurde erfolgreich entgegengenommen.

Sie werden automatisch per E-Mail und/oder SMS kontaktiert, sobald Sie an der Reihe sind. Eine weitere Kontaktaufnahme mit dem für Sie zuständigen Impfzentrum ist daher nicht erforderlich. Bitte verzichten Sie auch auf Nachfragen, da dies die Kapazitäten der Impfzentren belastet und zu Verzögerungen im Ablauf der Terminvereinbarungen führt.

Sollten Sie an weiteren Informationen zum aktuellen Impfgeschehen interessiert sein, empfehlen wir regelmäßig einen Blick in Ihr Benutzerkonto zu werfen. Dort haben Sie natürlich die Möglichkeit Ihre Daten jederzeit zu aktualisieren so lange Sie noch keinen Termin ausgewählt haben.

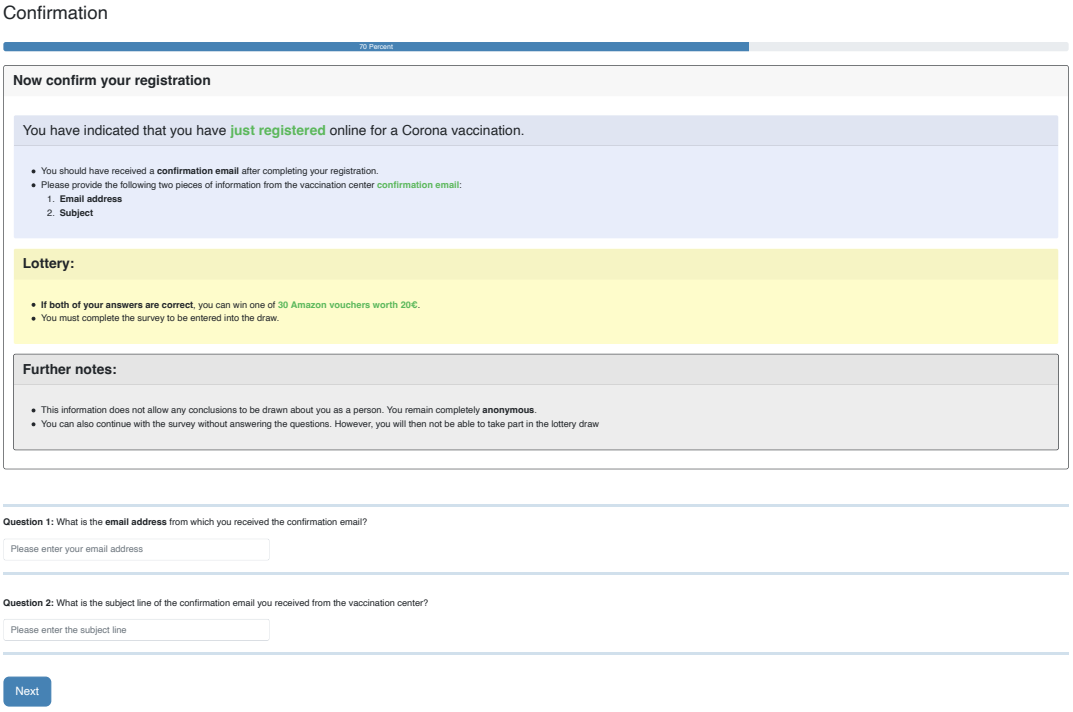
Die Vergabe der Impftermine orientiert sich an der Zugehörigkeit zu der jeweils aufgerufenen Prioritätengruppe. So wird sichergestellt, dass immer die besonders gefährdeten Menschen zuerst geimpft werden.

Wichtig: Die Impfung ist und bleibt für Sie kostenlos. Sollten Sie zur Herausgabe von Kontodaten oder Geldzahlungen aufgefordert werden (z.B. telefonisch oder per E-Mail), gehen Sie darauf nicht ein und wenden Sie sich an Ihre örtlich zuständige Polizeidienststelle.

Mit freundlichen Grüßen,

Ihr Bayerisches Staatsministerium für Gesundheit und Pflege

Figure 1.D.11: Survey page explaining verification of registration



1.E Survey Instrument[32]

I Basic demographic information

Question 1: Are you male or female?
Question 2: How old are you?
Question 3: In which federal state do you live?

new page

Since the end of last year (December 2020), *vaccinations against the coronavirus* (COVID-19 vaccinations) have been administered in Germany.

Question: Have you already received a COVID-19 vaccination? Reply options: *Yes or No*

new page

Did you know that?
In Bavaria, it is possible to register for a COVID-19 vaccination already, even though the actual vaccination may not take place for a few months. Registration takes place either online or by telephone at the Bavarian vaccination centres.

Question: Have you already registered for a COVID-19 vaccination? Reply options: *Yes or No*

new page

II Attitudes towards the COVID-19 vaccination

We would like to start by asking you a few basic questions regarding how you feel about the COVID-19 vaccination.
There are now several vaccines against the coronavirus on the German market. Vaccination is officially recommended for adults of all ages (exception: not during pregnancy and breastfeeding for the time being, as no data on safety and efficacy are yet available).

To what extent do you agree with the following statements?

- **Statement 1:** I have full confidence that the vaccination against COVID-19 is

[32]This section provides a translation of the original German-language survey instrument. The full original survey instrument was attached to our pre-registration at the AEA RCT Registry with ID AEARCTR-0007437 (Esguerra, Vollmer, and Wimmer 2021).

safe.
Reply options: *Likert scale (1-7) with 1: do not agree at all, 7: agree completely.*

- **Statement 2:** I have full confidence that the vaccination against COVID-19 is effective.
  Reply options: *Likert scale (1-7) with 1: do not agree at all, 7: agree completely.*
- **Statement 3:** I view vaccinations as a collective effort against the spread of COVID-19.
  Reply options: *Likert scale (1-7) with 1: do not agree at all, 7: agree completely.*
- **Statement 4:** If everyone is vaccinated against COVID-19, I don't need to get vaccinated too.
  Reply options: *Likert scale (1-7) with 1: do not agree at all, 7: agree completely.*

Question 1: How likely is it that you will get vaccinated against COVID-19?
Instruction: Please use the bar/slider for your answer. Click on the bar at the bottom to reveal the slider. Then move the slider to give your answer. 0 percent means "definitely not willing to get vaccinated". 100 percent means "definitely willing to get vaccinated".

Question 2: What do you think? What proportion of people in Bavaria are willing to get vaccinated against COVID-19?
Instruction: Please use the bar/slider for your answer. Click on the bar at the bottom to reveal the slider. Then move the slider to give your answer. 0 percent means "no one is willing to get vaccinated". 100 percent means "everybody is willing to get vaccinated".

new page

III Broader set of attitudes

How well do the following statements apply to you as a person?

- **Statement 1:** I like it when people accept my suggestions.
  Reply options: *Likert scale (1-7) with 1: do not agree at all and 7: agree completely.*
- **Statement 2:** I like it when my ideas and opinions have an impact on other people.
  Reply options: *Likert scale (1-7) with 1: do not agree at all and 7: agree completely.*
- **Statement 3:** I would like the feeling of having influenced other people's lives.
  Reply options: *Likert scale (1-7) with 1: do not agree at all and 7: agree completely.*

new page

How well do the following statements apply to you as a person?

- **Statement 1:** It is important to me to impress others.
  Reply options: *Likert scale (1-7) with 1: do not agree at all and 7: agree completely*.
- **Statement 2:** I think a lot about whether I am good enough compared to others.
  Reply options: *Likert scale (1-7) with 1: do not agree at all and 7: agree completely.*
- **Statement 3:** It is important to me how I am perceived by others.
  Reply options: *Likert scale (1-7) with 1: do not agree at all and 7: agree completely.*

new page

We now ask you about your behavior in certain situations.

Question: How much would you be willing to give to a good cause without expecting anything in return?
Reply options: *0: Not at all willing, 10: Extremely willing*

Imagine the following situation: Today you unexpectedly received 1,000 EUR.
Question: How much of the money would you donate to a good cause? *Note: You can enter whole numerical values from 0 to 1,000 here.*

new page

IV Joint task

Please read the following instructions carefully before proceeding with the survey.

- In the next section of our survey, we ask you to solve a short task together with another participant of this survey.
- Your task is to match famous pieces of art to the respective artist together with your partner.
- In this task, you can win one of 30 Amazon vouchers worth €10.
- You can communicate with your fellow player by means of a chat.
- To facilitate communication, please enter your first name or a nickname below.

Question: What is your first name or nickname?

Hint:

- In order to remain anonymous, please make sure to enter only your first name.

- You can also choose another name here. However, the name should correspond to your gender.

new page

We ask you to solve the upcoming task together with your partner.
Your partner is: [*name*]

[He/she] is [*xx*] years old. [He/she] lives in Bavaria.

Task: Together with your partner, match the following four pieces of art with the correct artist.

Hints:

1. You and your partner have 60 seconds for each piece of art.
2. If you and your partner correctly match at least three pieces of art, you can win one of 30 Amazon vouchers worth €10.
3. You must complete the full survey to qualify for one of the vouchers.
4. To increase your chances of winning, it is important that you and your partner work together.
5. You will receive points only if you both give the correct answer.
6. Use the chat window to communicate with your partner via text messages and coordinate your answers. The chat window is available for the entire task.
7. Its a good idea to introduce yourself to your partner with a short message right away.

[***Chat window***]

Final hints before the tasks begins: You may have to wait for a moment until your partner [*name*] has read the instructions and responds to you.
Reminder: You can win one of 30 Amazon vouchers worth €10.

new page

[*Painting is shown for 1 Minute.*]

Question: Which artist painted this piece of art?
Reply options: *Participants can choose one artist from a drop-down menu.*

[*This process is repeated four times. During this time the participants have the option to use the chat window to communicate.*]

new page

Question: Which of the following figures best reflects how connected you feel with your partner [*name*]?

Hints:

1. Option 1 means that you do not feel connected to your partner [*name*] at all.
2. Option 7 means that you feel very close to your partner [*name*].
3. Use the remaining options (2-6) to grade your answer.
4. To select one, click on the option in the header and not on the image.

new page

Please still think of your partner [*name*].
Question: To what extent would you refer to yourself and your partner [*name*] as "we".

Hints:

1. Option 1 means that you would definitely not refer to the two of you as "we".
2. Option 7 means that you would definitely speak refer to the two of you as "we".
3. Use the remaining options (2-6) to grade your answer.

new page

V Explanations on the survey

Instructions: In the following, we would like to ask you about your willingness to get vaccinated against COVID-19. Specifically, we would like to know whether you are willing to register for a COVID-19 vaccination right away. With that we are referring to the official registration process required for residents of Bavaria to be able to obtain an appointment at a vaccination center. In this survey, we will provide you with the opportunity to switch to the official registration website of the Bavarian Ministry of Health to complete the registration. Of course, the registration is voluntary and you can also complete the survey without registering.

Task: Confirm that you have understood these instructions by selecting the correct answer below.

Question: During this survey, will you be able to switch to the official registration website of the Bavarian Ministry of Health to complete the registration for a COVID-19 vaccination?
Reply options: *Yes or No*

new page

V.A Instructions Senders *"not informing partner"*

Instructions:
The survey proceeds as follows:

Step 1: You decide whether you want to register for a COVID-19 vaccination right away.
Step 2: Your partner [*name*] decides whether [he/she] wants to register for a COVID-19 vaccination right away.

Important: We do not tell your partner [*name*] whether you want to register for a vaccination.

You do not find out about the decision of your partner [*name*].

Task: Confirm that you have understood the instructions by selecting the correct answer below.
Question: Will your partner [*name*] find out whether you want to register?
Reply options: *Yes/No*

V.B Instructions Senders *"informing partner after"*

Instructions:
We will tell your partner [*name*] with a high probability whether you want to register for a vaccination. This proceeds as follows:

Step 1: You decide whether you want to register for a COVID-19 vaccination right away.
Step 2: Your partner [*name*] decides whether [he/she] wants to register for a COVID-19 vaccination right away.
Step 3: We tell your partner [*name*] whether you want to register for a vaccination.
Important: Your partner [*name*] will find out about your registration decision **only**

after [he/she] has already decided whether [he/she] wants to register.

You do not find out about the decision of your partner [*name*].

Task: Confirm that you have understood the instructions by selecting the correct answers below.

Question 1: Will your partner [*name*] find out with a high probability whether you want to register?
Reply options: *Yes/No*
Question 2: When will your partner [*name*] find out about your registration decision?
Directly before or **only after** [he/she] can register for a COVID-19 vaccination?
Reply options: *Directly before/Only after*

V.C Instructions Senders *"informing partner before"*

Instructions:
We will tell your partner [*name*] with a high probability whether you want to register for a vaccination. This proceeds as follows:

Step 1: You decide whether you want to register for a COVID-19 vaccination right away.
Step 2: We tell your partner [*name*] whether you want to register for a vaccination.
Important: Your partner [*name*] will find out about your registration decision **directly before** [he/she] can decide whether [he/she] wants to register.
Step 3: Your partner [*name*] decides whether [he/she] wants to register for a COVID-19 vaccination right away.

You do not find out about the decision of your partner [*name*].

Task: Confirm that you have understood the instructions by selecting the correct answers below.

Question 1: Will your partner [name] find out with a high probability whether you want to register?
Reply options: *Yes/No*
Question 2: When will your partner [*name*] find out about your registration decision?
Directly before or **only after** [he/she] can register for a COVID-19 vaccination?
Reply options: *Directly before/Only after*

V.D Instructions Receivers *"informed before"* and *"informed after"*

Instructions: The survey proceeds as follows:

Step 1: Your partner [*name*] decides whether [he/she] wants to register for a COVID-19 vaccination right away.
Step 2: You decide whether you want to register for a vaccination now. Since you are the second to decide you may have to wait for a moment.

We do not tell your partner [*name*] whether you want to register for a vaccination.

Task: Please confirm that you have understood these instructions by selecting the correct answer below.
Question: Will your partner find out about your decision?
Reply options: *Yes/No*

new page

VI Vaccination willingness

VI.1.A First stage Senders *"not informing partner"*

Reminder: Below we will provide you and your partner [*name*] with the opportunity to go to the official registration website of the Bavarian Ministry of Health to complete the registration process.

Your partner [*name*] will not know whether you wish to register for a COVID-19 vaccination.

Remember: Your partner [*name*] **will not** learn about your registration decision.
Question 1: What do you think? How likely is it that your decision to register or not to register will influence your partner's decision?

Hints:

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to give your answer.
- 0 percent means "there is no way I can influence my partner with my decision".
- 100 percent means "I can definitely influence my partner with my decision".

Remember: Your partner [*name*] **will not** learn about your registration decision.
Question 2: What do you think? How likely is it that your partner will make the same decision as you?

Hints:

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to give your answer.
- 0 percent means "my partner will definitely not decide the same way I do".
- 100 percent means "my partner will definitely decide like me" .

VI.1.B First stage Senders *"informing partner after"*

Reminder: Below we will provide you and your partner [*name*] with the opportunity to go to the official registration website of the Bavarian Ministry of Health to complete the registration process.

Your partner [*name*] will learn with a high probability whether you wish to register for a COVID-19 vaccination.

Remember: Your partner [*name*] will learn about your registration decision only after [he/she] has already decided whether to register for COVID-19 vaccination now.
Question 1: What do you think? How likely is it that your decision to register or not to register will influence your partner's decision?

Hints:

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to give your answer.
- 0 percent means "there is no way I can influence my partner with my decision".
- 100 percent means "I can definitely influence my partner with my decision".

Remember: Your partner [*name*] will learn about your registration decision only after [he/she] has already decided whether to register for COVID-19 vaccination now.
Question 2: What do you think? How likely is it that your partner will make the same decision as you?

Hints:

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to give your answer.
- 0 percent means "my partner will definitely not decide the same way I do".
- 100 percent means "my partner will definitely decide like me" .

VI.1.C First stage Senders *'informing partner before'*

Reminder: Below we will provide you and your partner [*name*] with the opportunity to go to the official registration website of the Bavarian Ministry of Health to complete the registration process.

Your partner [*name*] will learn with a high probability whether you wish to register for a COVID-19 vaccination.

Remember: Your partner [*name*] will learn about your registration decision right before [he/she] decides whether to register for a COVID-19 vaccination.
Question 1: What do you think? How likely is it that your decision to register or not to register will influence your partner's decision?

Hints:

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to give your answer.
- 0 percent means "there is no way I can influence my partner with my decision".
- 100 percent means "I can definitely influence my partner with my decision".

Remember: Your partner [*name*] will learn about your registration decision right before [he/she] decides whether to register for a COVID-19 vaccination.
Question 2: What do you think? How likely is it that your partner will make the same decision as you?

Hints:

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to give your answer.
- 0 percent means "my partner will definitely not decide the same way I do".
- 100 percent means "my partner will definitely decide like me" .

new page

VI.2.A Registration intent Senders *"not informing partner"*

Reminder: if you live in Bavaria and want to get vaccinated, this registration is required to get a vaccination appointment at a Bavarian vaccination center.
Your partner [*name*] **will not** learn if you want to register for a COVID-19 vaccination.

Question: Would you like to register for a COVID-19 vaccination?
Reply options: *Yes/No*

VI.2.B Registration intent Senders *'informing partner after'*

Reminder: if you live in Bavaria and want to get vaccinated, this registration is required to get a vaccination appointment at a Bavarian vaccination center.
Your partner [*name*] will learn with a high probability if you wish to register for a COVID-19 vaccination.

Important: Your partner [*name*] will learn about your registration decision only after [he/she] has already decided whether to register for a COVID-19 vaccination.

Question: Would you like to register for a COVID-19 vaccination?
Reply options: *Yes/No*

VI.2.C Registration intent Senders *'informing partner before'*

Reminder: if you live in Bavaria and want to get vaccinated, this registration is required to get a vaccination appointment at a Bavarian vaccination center.
Your partner [*name*] will learn with a high probability if you wish to register for a COVID-19 vaccination.

Important: Your partner [*name*] will learn about your registration decision directly before [he/she] decides whether to register for a COVID-19 vaccination.

Question: Would you like to register for a COVID-19 vaccination?
Reply options: *Yes/No*

VI.2.D Registration intent Receivers *'informed after'*

Reminder: if you live in Bavaria and want to get vaccinated, this registration is required to get a vaccination appointment at a Bavarian vaccination center.
Your partner will not know if you want to register.

Question: Would you like to register for a COVID-19 vaccination?
Reply options: *Yes/No*

VI.2.E Registration intent Receivers *'informed before'*

Reminder: if you live in Bavaria and want to get vaccinated, this registration is required to get a vaccination appointment at a Bavarian vaccination center.
Your partner will not know if you wish to register.

Important: Your partner [*name*] [would like/would not like] to register for a COVID-19 vaccination.

Question: Would you like to register for a COVID-19 vaccination?
Reply options: *Yes/No*

new page

VI.3 Registration for COVID-19 vaccine

Would you like to register now?
To register, please click on **Yes, register now for a COVID-19 vaccination** below.
This will open the official registration website of the Bavarian Ministry of Health in a new browser window or tab. To successfully register for a COVID-19 vaccination, follow the instructions on the registration website.

Important: Please do not close the browser window or tab in which you are answering the survey during registration.

Additional Notes: We do not have any access to the information you provide on the registration website. Registration is voluntary and it does not entail an obligation to get vaccinated. Your reward for this survey is independent of whether you register.

Button: *Yes, register for the COVID-19 vaccination right away.*

[*Opens the link to the official registration website.*]

Have you successfully registered?

Here's how to proceed: once you have registered, please continue with the survey by clicking **Yes, I have registered and would like to continue with the survey** at the bottom of this page.

Don't want to register now?
If you do not wish to register now, you will not be penalized in any way, for example by being paid less for this survey. To continue with the survey, please click **No, I have not registered and would like to continue with the survey** at the bottom of this page.

To continue with the survey, please answer the following question:
Question: have you just register for the COVID-19 vaccination?

Reply options:

- *No, I have not registered and would like to continue with the survey*
- *Yes, I have registered and would like to continue with the survey*

new page

VI.4 Confirmation of registration for COVID-19 vaccination

Now confirm your registration: You have indicated that you have just registered online for a Corona vaccination.
You should have received a confirmation email after completing your registration.
Please provide the following two pieces of information from the confirmation email sent out by the vaccination center:

1. Email Address
2. Subject

Lottery: If both of your answers are correct, you can win one of 30 Amazon vouchers worth 20€.

You must complete the survey to qualify for the lottery.

Further notes: Providing this information does not allow us to infer anything about you as a person. You remain completely anonymous. You can also continue with the survey without answering the questions. However, you will then not be able to participate in the lottery draw.

Question 1: What is the email address from which you received the confirmation email?
Question 2: What is the subject of the confirmation email you received from the vaccination center?

new page

VI.5 What do you think about the COVID-19 vaccine?

Question 1: What do you think? How safe is the COVID-19 vaccination?
Reply option: *Likert scale (1-7) with 1: not at all safe, 7: extremely safe.*

Question 2: What do you think? How effective is the COVID-19 vaccination?
Reply option: *Likert scale (1-7) with 1: not at all effective, 7: extremely effective.*

Question 3: What do you think? To what extent is it socially desirable to get vaccinated against COVID-19?
Reply option: *Likert scale (1-7) with 1: not at all socially desirable, 7: extremely socially desirable*

Question 4: To what extent do you agree with the following statement? Statement: if everyone is vaccinated against COVID-19, I don't need to get vaccinated too.
Reply option: *1: do not agree at all, 7: agree completely*

new page

Question: How likely are you to get vaccinated against COVID-19?
Please use the bar/slider for your answer.

- Click on the bar at the bottom to reveal the slider.
- Then move the slider to make your selection.
- 0 percent means "definitely not willing to get vaccinated."
- 100 percent means "definitely willing to get vaccinated."

new page

VII Further demographic information

To conclude this survey, please provide some general information.

Question 1: What county do you live in?
Question 2: What is your zip code?
Question 3: What was your household's monthly net income last year?
Note: We mean the sum that results from wages, salaries, income from self-employment, pensions, income from public aid, income from letting, housing allowances, child benefits and all other incomes, after the deduction of taxes and social security contributions.
Reply options:

- Less than 1,100 EUR
- 1.100 - 1.500 EUR
- 1,501 - 2,000 EUR
- 2,001 - 2,600 EUR
- 2,601 - 4,000 EUR
- 4,001 - 7,500 EUR
- More than 7,500 EUR

Question 4: What is your highest educational degree (general or vocational)?

new page

VIII End of survey

Thank you for participating in our survey!

In the following, we list your performance in the task in which you had to assign artworks to artists together with your partner and inform you whether you have won one of the Amazon vouchers. Afterwards, we ask you to answer two more questions about this survey yourself and give you the opportunity to give us feedback on the survey.

- Unfortunately, you have not won one of the raffled Amazon vouchers./Congratulations, you have won one of the raffled Amazon vouchers.

- If you would like to know how you and your partner did on your shared task, please click here. [*Upon clicking the button, participants' answers and the corresponding solutions open in the same window.*]
- *For Receivers 'informed after':* Finally, we would like to inform you that your partner [*name*] [registered/did not register] for a COVID-19 vaccination.
- Thank you again for participating in our survey.

Please answer the following questions to complete the survey:

Question 1: What do you think? What was the purpose of this survey?
Question 2: Where on the political spectrum would you place this survey?
Hints: Please use the slider to tell us the extent to which you felt this survey was leaning more toward the political right or toward the political left.
Click on the bar below to reveal the slider. Then move the slider to make your selection.

Feedback If you would like to give us any feedback on the survey, please feel free to do so here.

Would you like to close the survey now?

Click on ***Close survey***

CHAPTER 2

Is Right-Wing Populist Rhetoric Contagious? Evidence from Parliamentary Speeches in Germany

2.1 Introduction

For a long time, far-right political rhetoric and ideas had been ostracized in Western democracies. Yet, the recent rise of right-wing populism across many countries has been accompanied by an increasing normalization and acceptability of such language in the political discourse (Guriev and Papaioannou 2022). Spreading extreme ideas by saying the previously intolerable has been part of the successful playbook of right-wing populists such as Viktor Orbán in Hungary, Jair Bolsonaro in Brazil, or Donald Trump in the United States. Recent research has highlighted that such language has consequences and can have detrimental effects on political attitudes, social norms, and even violent behavior (Bursztyn, Egorov, and Fiorin 2020; Djourelova 2023; Müller and Schwarz 2020, 2021). While a growing body of research has documented the consequences of such changes in acceptability, less work exists that investigates the mechanisms leading to the spread of right-wing populist ideas. We argue that day-to-day exposure plays a key role in this normalization process and show how contact with right-wing populism makes politicians from mainstream parties adopt and converge to the language employed by the extreme right.[1]

In this paper, we study how the first-time entry of a right-wing populist party, the *Alternative für Deutschland* (AfD), to the federal parliament of Germany, the *Bundestag*, affected the political rhetoric of incumbent politicians. Using techniques from natural language processing on several thousand parliamentary speeches, we construct different measures of rhetorical similarity to the language employed by right-wing AfD politicians. To induce variation in politicians' exposure to right-wing populists, we exploit a quasi-exogenous component in the allocation of parties to parliamentary committees. This allows us to analyze the causal effect of individual-level contact with AfD politicians on rhetorical similarity to right-wing political speech.

We find that politicians who are relatively more exposed to right-wing politicians in

[1] Convergence in our framework refers to the habitualization of right-wing rhetoric in the political discourse by increasing usage of distinctively right-wing vocabulary but does not necessarily imply convergence in ideology.

committees use language more similar to right-wing rhetoric. More precisely, comparing a politician with the highest to one with the lowest relative AfD exposure increases the cosine similarity to right-wing speech by 0.1 of a standard deviation, an effect size comparable to the average distance between the main center-left SPD and center-right CDU/CSU parties. Importantly, our difference-in-differences approach allows us to estimate this effect *within* individual speakers, highlighting how politicians converge to AfD rhetoric in response to higher exposure. Our findings imply that direct contact and confrontation with right-wing populism might exert a contagion effect on political language, even in a democracy that places a high social stigma on far-right ideology and rhetoric.

We corroborate this main result with two alternative measures of rhetorical similarity to right-wing populism: relatively higher AfD exposure also makes politicians use language more similar to extra-parliamentary speeches by the far-right AfD politician Björn Höcke, who is known to employ an extreme right-wing rhetoric. Furthermore, we find that speakers are more likely to use populist-specific phrases in their speeches as identified in the German-language populist dictionary by Gründl (2022). Placebo tests suggest that the effect is specific to exposure to right-wing populism and does not extend to any interaction with other politicians of a different political ideology.

Finally, we explore why politicians might adopt right-wing language in their publicly displayed speeches. Building on insights from theories in social psychology on communication accommodation (Giles and Ogay 2007), we hypothesize that such language use follows strategic motives with respect to electoral support. Indeed, our results show that the contagious effect of AfD exposure on political rhetoric increases with the intensity of local competition with the AfD in a politician's electoral district.

This study contributes to a number of active research agendas in economics and political science. First, our article adds to the rapidly growing literature on populism and political change, as recently reviewed by Guriev and Papaioannou (2022). Specifically, it aims to contribute to a better understanding of how populist politicians can influence political and social norms and, ultimately, affect behavior. A number of existing studies have shown how the electoral success of populism can increase the acceptability of extreme political rhetoric and social norms up to the point of fanning hate crimes (Albornoz, Bradley, and Sonderegger 2020; Bursztyn, Egorov, and Fiorin 2020; Hagemeister 2022; Müller and Schwarz 2020, 2021; Romarri 2022; Schilter 2018). The strong connection between language and norms has been emphasized by Gentzkow, Shapiro, and Taddy (2019) who argue that changes in political rhetoric might contribute to differences in animus in the broader public. Consistent with this argument, Djourelova (2023) documents how even small differences in language alone can have wide-ranging impacts on political attitudes. Newman et al. (2021) find how the use of

explicitly inflammatory speech by political elites can have an emboldening effect on expressing prejudiced opinions among the general public. In our setting, we study the spread of right-wing language *within* the political elite, potentially setting a precedent for the subsequent normalization and further dissemination to a wider audience.

Second, this study is also embedded in the literature on strategic policy responses of mainstream parties to rising populism (Meguid 2005). Using text data from party manifestos, Abou-Chadi (2014) shows that parties' strategic reactions differ vis-à-vis radical right and green contenders. When radical right parties gain electoral support, convergence to anti-immigration positions follows suit, while in contrast parties de-emphasize ecological issues in response to green competitors. Similarly, work by van Spanje (2010) and Abou-Chadi and Krause (2020) provides evidence for a contagious effect on anti-immigration stances of mainstream parties across Europe in response to radical right parties' appearance. While the study by Hjorth and Larsen (2020) on Denmark demonstrates how accommodating strategies can be beneficial in terms of electoral success for left-wing parties, other studies find inconclusive or conflicting results on the effectiveness of such accommodation to radical right parties (Bale et al. 2009; Dahlström and Sundell 2012; Krause, Cohen, and Abou-Chadi 2023; Spoon and Klüver 2020). We extend existing research on party-level accommodation by studying rhetorical changes of individual politicians in the face of newly emerging right-wing populism.

Third, we advance the existing literature on the effects of polarization and populism on parliamentary speech.[2] Previous studies of political speech have, among others, studied plenary debates in Sweden (Magnusson et al. 2018), the UK (Gurciullo et al. 2015), Norway (Fiva, Nedregård, and Øien 2021), or the European Parliament (Greene and Cross 2015). For the case of Germany, Lewandowsky et al. (2022) and Atzpodien (2022) explore how the entry of the AfD to the Bundestag and state parliaments, respectively, affects issue-specific polarization in plenary debates, with only the latter finding evidence for an increase in polarization over immigration. Similarly, Breyer (2022) analyzes parliamentary speeches in Austria and Germany and finds that both mainstream and populist parties use more populist rhetoric when in opposition than when in government. Whereas most of these studies only provide correlational evidence, a notable exception is the work by Valentim and Widmann (2021) that exploits variation in the timing of elections when AfD politicians enter German state parliaments. They find that politicians of other parties respond by using more positive,

[2]This also relates to a body of research studying the effects of populism on party manifestos. Rooduijn, Lange, and van der Brug (2012) analyze whether populism has contagious effects on the party manifestos of non-populist established parties in Western democracies, finding that manifestos of mainstream parties have not become more populist in recent years. Similarly, Han (2014) analyzes the potential impact of radical right-wing parties on policy positions of mainstream parties regarding multiculturalism and immigration.

rather than negative, emotional rhetoric in their speeches. Our study goes beyond existing approaches by exploiting a novel source of variation in individual-level exposure to right-wing politicians in parliament. This allows us to study within-speaker changes in political rhetoric in the same parliament and to shed light on the important role of day-to-day work interactions with right-wing colleagues. Furthermore, we go beyond sentiment analysis and party positions by employing both similarity and dictionary measures of distance to right-wing rhetoric.

Finally, our empirical approach adds to a rapidly growing literature in economics and political science that studies large-scale text data combining methods from natural language processing with the toolkit for causal inference of applied econometrics (Gentzkow, Kelly, and Taddy 2019; Gentzkow, Shapiro, and Taddy 2019; Hager and Hilbig 2020; Kelly et al. 2021; Widmer, Galletta, and Ash 2022; Wilkerson and Casas 2017). In particular, the addition of a novel source of variation due to a quasi-exogenous committee allocation rule may offer new research opportunities to study the effects of individual-level exposure to other politicians.

The remainder of this paper is structured as follows: Section 2.2 provides background on right-wing populism in Germany and discusses existing research on political rhetoric and strategic accommodation. Section 2.3 describes the data and construction of our measures of similarity to right-wing rhetoric. Section 2.4 introduces our identification strategy and explains how allocation rules to parliamentary committees in the Bundestag lead to quasi-exogenous variation in exposure to right-wing populists. Section 2.5 presents our main results as well as a number of robustness checks and discusses evidence on strategic reasons for rhetorical change. Finally, Section 2.6 concludes.

2.2 Background

2.2.1 Right-Wing Populism in Germany

Since the re-establishment of parliamentary democracy in 1949 after the end of the Nazi dictatorship, far-right parties had for a long time only played a minor role in (West) German politics. At the federal level, no far-right or right-wing populist party had managed to cross the 5% electoral threshold for parliamentary representation in the German Bundestag.[3] In the federal election of September 2013, a newly established right-wing party called *Alternative für Deutschland* (AfD, Alternative for Germany) fell

[3] At the state- and municipal-level, a number of radical right-wing parties such as the *Sozialistische Reichspartei* (SRP, Socialist Reich Party) - which was banned by the German Federal Constitutional Court in 1952 - the *Deutsche Volksunion* (DVU, German People's Union), the *Republikaner* (REP, Republicans), and the *Nationaldemokratische Partei Deutschlands* (NPD, National Democratic Party of Germany) enjoyed geographically and temporarily limited electoral success that never proved to be sustainable in the long-run.

just short of overcoming this threshold when it won 4.7% of the votes. Subsequently, the AfD continued to gain electoral support and established itself in several state parliaments, albeit undergoing an increasing radicalization and a strong shift to the right in the context of the 2015 European migration crisis. In the next federal election in September 2017, the AfD scored 12.6% of the votes and entered the Bundestag for the first time as the third largest parliamentary group and the strongest opposition party. The AfD's continued electoral success appears to be sustainable since it re-entered the Bundestag with a 10.3% of the vote share in the 2021 federal election. Furthermore, the AfD is currently (as of January 2023) represented in 15 of 16 German state parliaments, as well as in the European Parliament.

While having been founded in early 2013 in the context of the European debt crisis as a socially conservative party with soft eurosceptic views (Arzheimer 2015), the AfD veered increasingly to the right of the political spectrum and evolved into a populist radical right-wing party with a distinctively anti-immigration, anti-refugee, and anti-Islam platform (Arzheimer and Berning 2019). This ideological shift to the far right also manifested itself in a significant change in the language used by the AfD in speeches, party manifestos, and social media posts with an increasing usage of words related to Islam, migration, and the nation/Germany (Cantoni, Hagemeister, and Westcott 2020).[4] Parts of the AfD have also cooperated with the xenophobic PEGIDA ("Patriotic Europeans Against the Islamization of the Occident") movement that organizes anti-immigrant rallies mostly in East Germany. Prominent members of the AfD have held speeches at PEGIDA rallies, such as Björn Höcke, the de facto leader of the far-right faction within the AfD *"Der Flügel"* ("The Wing"). This intra-party group had been put under surveillance by the *Federal Office for the Protection of the Constitution* for being considered a "secured extreme right-wing threat against the free democratic constitutional order" (Bundesamt für Verfassungsschutz 2020). Although *"Der Flügel"* was officially dissolved in 2020, both the main federal party itself, several state-level associations of the AfD as well as the AfD's youth organization *Junge Alternative* (JA, Young Alternative) continue to be classified by domestic intelligence agencies as a "subject of extended investigation to verify a suspicion" for suspected right-wing extremism (Bundesamt für Verfassungsschutz 2023). Furthermore, following classifications by political scientists (Arzheimer and Berning 2019; Hansen and Olsen 2018), we argue that the AfD can be considered as a populist radical right party in the spirit of Mudde (2007). According to this definition, populism among Western far-right parties can be understood as politicizing the "pure people" against the "corrupt elite", reflecting a dichotomous understanding of society.

[4]This increasing radicalization of the AfD is furthermore exemplified by the fact that two of its three initiators (Bernd Lucke and Konrad Adam), two former party leaders (Frauke Petry and Jörg Meuthen) as well as multiple members in the Bundestag and state parliaments left the party claiming that it had become too radical.

2.2.2 Accommodation

As the success of the AfD in consecutive elections at various legislative levels appears to be enduring, the question arises of how existing "traditional" parties and their politicians react to and deal with this new populist competitor on their right. Initially, after the entrance of the AfD to the Bundestag and the different state parliaments, all mainstream parties tried to emphasize the formation of a *cordon sanitaire* against the AfD with the exclusion of any formal cooperation.[5] As documented by Heinze (2022), increasing signs of minor cooperation between established parties and the AfD as well as a turn toward *ad hoc toleration* could be observed at the municipal and the state level: while there has been no formation of official coalitions so far, mainstream parties have elected AfD candidates to parliamentary offices and debated motions by the AfD on a case-by-case basis. The arguably biggest violation of this non-cooperation policy happened in the federal state of Thuringia in February 2020, when Thomas Kemmerich from the liberal FDP was elected minister-president with the votes of the AfD and the conservative CDU. Kemmerich quickly had to step down amongst massive public outcry and resistance from the FDP and CDU federal leaderships. The case exemplifies the increasing difficulties parties and individual politicians are facing in response to the sustained electoral success of the AfD. Especially in some states in East Germany, where the AfD has managed to repeatedly score close to or more than 25% of the vote share, the formation of government coalitions as well as the functioning of parliamentary routines become increasingly difficult.[6] This raises the question whether both parties and individual politicians might resort to an *accommodation* strategy towards the AfD. In the following, we will discuss a number of existing theoretical frameworks and empirical findings for potential accommodating reactions to new – in particular radical right-wing and populist – parties.

Accommodation by Parties Since the AfD has shown to be able to repeatedly gain considerable shares of votes at different electoral levels, it is essential to examine the reactions of established parties to such an electoral threat. In particular, existing studies from political science have studied whether and how parties adopt their policy platforms in response to the rise of (populist right-wing) contenders. Using text data from party manifestos in Western European countries, Abou-Chadi (2014) shows that

[5] For example, the AfD has so far been denied by the other parties the election of a Bundestag vice-president from their ranks of which traditionally every parliamentary group received at least one position. While all of the six candidates presented by the AfD since 2017 failed to receive the required simple majority, they have increasingly scored more votes than the AfD itself has seats, hinting at an increased questioning of this formal exclusion practice among some MPs from other parties.

[6] In the 2017 federal elections, the AfD received the second-largest vote share with 21.9% in East Germany (vs. 10.7% in West Germany), even coming out as the strongest party in the state of Saxony (27.0%). Furthermore, the AfD received more than a fifth of the vote share in the state elections of Brandenburg 2019 (23.5%), Mecklenburg-Western Pomerania 2016 (20.8%), Saxony 2019 (27.5%), Saxony-Anhalt 2016 (24.3%) and 2021 (20.8%), and Thuringia 2019 (23.4%).

parties' reactions to radical right and green contenders differ: when radical right-wing parties are able to gain substantial electoral support, convergence to anti-immigration positions follows suit. If green parties gather stronger support, however, existing parties de-emphasize ecological issues. In a similar vein, the empirical findings by van Spanje (2010) point towards a contagion impact on entire party systems with respect to immigration policy positions in Western European countries following electoral success of the extreme right. Relatedly, Abou-Chadi and Krause (2020) show that mainstream parties in European democracies change their immigration policies if radical right parties enter parliament. With respect to the effectiveness of such strategies, a survey experiment in Denmark by Hjorth and Larsen (2020) highlights how accommodation by left mainstream parties can attract anti-immigration voters at the expense of pro-immigration voters. As former voters of left mainstream parties switch to other left parties without anti-immigration stances, this can in turn lead to an increased overall support for left parties. Accommodation towards right-wing positions might in this way foster the political prospects of the mainstream left in governing coalitions. However, other studies find conflicting or inconclusive results on the effectiveness of strategic accommodation to radical right parties (Bale et al. 2009; Dahlström and Sundell 2012; Krause, Cohen, and Abou-Chadi 2023; Spoon and Klüver 2020). Given that the AfD received substantial and continued support in elections at different levels in Germany, we might expect some form of reaction to this electoral threat among existing parties, in particular, as it has been shown that the AfD was successful in politicizing issues that were previously less controversial and, respectively, less politicized (Engler et al. 2022; Gessler and Hunger 2021; Hansen and Olsen 2022).

Accommodation by Individuals While much attention has been paid to strategic accommodation decisions by entire parties, the accommodating behavior of *individual* politicians in the face of newly emerging (populist right-wing) parties has not been thoroughly examined. One reason for this might be that due to the traditionally strong party discipline – especially in parliamentary systems across Europe – it might seem that individual MPs have less room for potentially accommodating decisions in terms of voting behavior or the choice of policy platforms.

Therefore, in this paper we study changes in the political rhetoric of individual politicians in publicly held parliamentary speeches. This has a number of advantages with respect to alternative sources available for text analysis: While party manifestos and policy papers are often the product of widespread cooperation among party members and the party leadership, parliamentary speeches are more directly attributable to individual politicians. Furthermore, party manifestos are typically only drafted for election campaigns, whereas parliamentary speeches are given on a regular basis, allowing us to more directly capture reactions to exposure to right-wing populists as

well as to take care of time-specific trends.[7] Parliamentary speech also differentiates itself from legislative text, since the latter is a very formal type of language with multiple individuals involved in the writing process, whereas speeches leave more room for individual rhetorical accentuation.

For our analysis of accommodation in parliamentary speech, we draw on the framework of *Communication Accommodation Theory* developed in social psychology (see Giles and Ogay 2007). This framework aims to predict and explain individual language adjustments as a function of creating, maintaining, or decreasing the social distance in personal interactions. In particular, communication accommodation theory consists of four main components: first, communication is context-specific and contingent on the receiver. For example, individuals communicate differently when talking to their friends than when talking to people they do not know. Second, language use is the result of habit formation and is subject to gradual change. Communication experience and social interactions shape the way language is used. Third, communication is used in part to indicate and signal their attitudes toward each other and can therefore be seen as a "barometer of the level of social distance" (Giles and Ogay 2007, p. 294). In this sense, *accommodation* is a movement toward and away from others by changing communicative behavior. Among the different possible accommodative strategies speakers can use, the most frequent ones are *convergence* – adapting one's own communication to become more similar to others – and *divergence*, i.e., accentuating the differences between self and others. Fourth, accommodation entails benefits and costs. The benefit of accommodation is that greater similarity to the conversational partner might lead to greater approval, respect, or even direct social rewards from the accommodated speaker.

Taken together, in our context of parliamentary speeches in the German Bundestag, this framework implies that politicians face a trade-off: with increasing support for right-wing populism, they could choose *converging* accommodation toward right-wing rhetoric in order to win support from both the right-wing populist electoral base as well as the right-wing politicians themselves. The cost of this strategy could be an alienation from in-group politicians as well as the own electoral base, which might sanction right-wing populist accommodation with lower support. Alternatively, politicians might opt for *divergence* in accommodation towards right-wing political speech and choose a language that is clearly distinct from right-wing populist rhetoric. A benefit of this strategy might be increasing support from in-group politicians and the non-populist voter base, at the cost of losing voters attracted by right-wing populism, as well as lower potential of cooperation with right-wing populist politicians.

[7] An advantageous feature of our setting is that plenary speeches are often given in the afternoon right after meetings of parliamentary committees, where politicians have been in direct contact with AfD colleagues as will be explained with more detail in Section 2.4.

2.3 Data

2.3.1 Parliamentary Speech Data

Our empirical analysis is based on the *Open Discourse* dataset by Richter et al. (2020), a corpus of (plenary) parliamentary speeches in the German Bundestag. The dataset consists of all plenary protocols with the texts and metadata of speeches since the first session of the Bundestag in 1949, as well as demographic information on the speakers, such as their age, gender, occupation, and place of residence. For our analysis, we choose a time window around the first-time entry of the AfD in the German Bundestag after the federal elections in 2017: our dataset contains all speeches of the 18th Bundestag between October 2013 and September 2017 as well as all speeches of the 19th Bundestag between October 2017 and December 2019.[8]

To render the data more suitable for our analysis, we perform a number of pre-processing steps in the following order: first, we exclude speeches by the President and Vice-Presidents of the Bundestag, the respective chairperson of the plenary sessions, or other speeches related to special functions, as they are likely to merely reflect administrative content. Second, we only keep speeches by speakers who are members of the Bundestag and were a member in at least one parliamentary committee during the analyzed period. This ensures a comparable setting for all analyzed speeches, since members of the government, members of the parliament in special functions, and external speakers might systematically differ in how and about what they speak. Third, we correct a number of corpus-specific text issues: we remove punctuation including characters specific to the German language and the context (e.g., –, used to denote speech breaks), as well as digits, other numerical characters, and stopwords. Fourth, and as the final pre-processing step, we lemmatize the remaining tokens. A more detailed description of all steps of data preparation and pre-processing, including the software packages employed, is provided in Appendix Section 2.C.1. Our final dataset consists of 39,310 speeches held by 931 different speakers over the course of 57 months between October 2013 and December 2019.[9]

[8] We decided to not use speeches after January 2020 until the end of the 19th Bundestag in September 2021, as this period was heavily influenced by the COVID-19 pandemic. In particular, as discussed in Section 2.4, our empirical strategy critically hinges on direct and repeated personal contact between MPs in parliamentary committees. However, with the outbreak of the pandemic, the Bundestag changed its rules of procedure to allow for the participation in committee sessions via electronic means of communication and reduced the necessary quorum of attending members to one quarter instead of the usual 50% majority (Deutscher Bundestag 2020). Therefore, we cannot directly compare the level and quality of personal interaction with AfD members in committees during this time period with the period prior to the COVID-19 pandemic.

[9] Figure 2.A.1 in the Appendix shows the distribution of the speeches in our dataset over time and by party.

2.3.2 Committee Data

We gather data on Bundestag committees (*Bundestagsausschüsse*) from multiple sources: committee names and lists of committee members for the 18th Bundestag (2013-2017) and 19th Bundestag (2017-2021) were retrieved from the website of the Bundestag (Deutscher Bundestag 2022a). Since the names of committees and their responsibilities for different policy areas might slightly change over legislative periods, we manually harmonized committees based on the committee names in the 19th Bundestag (2017-2021). Throughout all of our analyses, we only evaluate full membership in committees and disregard if MPs are deputy or stand-by members in committees as they do not regularly attend committee sessions.[10]

We merge the information on committee membership – that is constant within a legislative period – to the main speech-level dataset via the name and party affiliation of a speaker. In addition to the information on age, gender, residency, and occupation of MPs contained directly in the *Open Discourse* dataset of parliamentary speeches, we furthermore add constituency-level data on results in federal elections as well as on which MPs stood as candidates in which electoral district obtained from Bundeswahlleiter (2022).

2.3.3 Measuring Similarity to Right-Wing Populist Rhetoric

Cosine Similarity Our preferred measure of the similarity of a speech to right-wing populist language is the standardized average cosine similarity to AfD speeches. More specifically, we construct the AfD cosine similarity score for speech i as the average over all pairwise cosine similarities of speech i with all AfD speeches $j \in J$

$$\text{AfD Cosine Similarity}_i = \frac{1}{J} \sum_{j=1}^{J} \frac{\sum_{k=1}^{K} a_k b_k}{\sqrt{\sum_{k=1}^{K} a_k^2 \sum_{k=1}^{K} b_k^2}} \tag{2.1}$$

where a_k and b_k are *term-frequency inverse-document-frequency* (tf-idf) weighted counts of word k in speeches i and j. We use tf-idf weighting and calculate tf-idf scores for each speech because words with particularly high frequencies or extremely low occurrence are usually not informative.[11] These scores take into account both the frequency of words within a given speech as well as the relative frequencies of words with respect to the overall corpus of speeches. The tf-idf weighted count of word k in speech i is given by

$$a_k = tf(i,k) \cdot idf(k) = \frac{f_{k,i}}{\sum_{k \in i} f_{k,i}} \cdot ln \frac{I}{|\{i \in I : k \in i\}|} \tag{2.2}$$

[10] Several committees confirmed to us in writing that stand-by members attending committee sessions is the exception rather than the rule and that personal attendance is usually only observed in case of full members being sick or otherwise incapacitated.

[11] A more detailed description on the implementation can be found in Appendix Section 2.C.2.

where $f_{k,i}$ is the frequency of word k in speech i and I is the total number of speeches. For ease of interpretation and comparison, we standardize the cosine similarity measure with mean zero and standard deviation one. As speeches differ in length, we also calculate cosine similarities to AfD speeches using different sample restrictions on the minimum number of terms of a speech.

Speeches at Far-Right Rallies As a second outcome measure, we compute the average cosine similarity to speeches given by Björn Höcke at far-right rallies in 2015 and 2016.[12] Björn Höcke is the chairman of the AfD in the East German state of Thuringia and is the de facto leader of the increasingly influential hard-line nationalist faction within the AfD.[13] Höcke has repeatedly made headlines with a number of highly controversial statements which have been considered to exhibit racist and xenophobic views as well as elements of historical revisionism and fascism.[14]

The speeches held by Höcke in 2015 focused on asylum policy and the contemporaneous large influx of refugees and how, according to Höcke, the government was actively trying to harm the German population. In the January 2016 speeches, Höcke additionally exploits for political purposes the events of the 2015 New Year's Eve sexual assaults in Cologne. In his speeches, Höcke uses clearly identifiable patterns and elements of populism and nativism (Mudde and Rovira Kaltwasser 2018). Many statements allude that there is allegedly too much immigration to Germany that poses a threat to the security and culture of native Germans. For example, Höcke claims that "we have hundreds of thousands of illegal immigrants in hiding, we have millions of Muslims living in non-integrated parallel societies" (January 13, 2016) or that "the millions of young men who are now being let in will also be legalized by the Germany abolitionists of the *Altparteien* ["old parties", derogatory term for established parties]" (September 30, 2015). In his speech on January 27, 2016, Höcke proclaims that "we want to live according to our values and customs and norms, we want to preserve our culture, we do not want to go back to the Middle Ages, we want to keep our country!". Another important topic of his speeches is the purported antagonism between the established political elites and the German people. In his speech on September 30, 2015, Höcke says about a local politician: "[...] because he stands up for the rule of the people, he can no longer stand the fact that the media-political pseudo-elite in this country tramples on the will of the people!". Some passages even contain barely veiled warn-

[12] The four speeches were held in Erfurt, Thuringia, on September 30, 2015, October 28, 2015, and January 13, 2016, as well as in Magdeburg, Saxony-Anhalt, on January 27, 2016, and have a length of 1,574, 2,432, 1,653, and 1,686 terms, respectively. We have retrieved the texts from the transcripts of the speeches provided by Enderstam (2020).

[13] As described in Section 2.2.1, Höcke was the de facto leader of the far-right faction *"Der Flügel"* within the AfD that was put under surveillance by domestic intelligence services and later dissolved.

[14] For example, Höcke has criticized the Memorial for the Murdered Jews of Europe in Berlin as a "monument of shame" and called for "a 180 degree turnover" in Germany's remembrance of the Nazi era (Bennhold and Eddy 2019).

ings about upheaval or revolt: "Sometimes one could think that our country is being deliberately plunged into chaos in order to establish an authoritarian order." (January 13, 2016). Overall, Höcke employs a radical and extremist language that constitutes a sharp departure from the established consensus on German political rhetoric.[15]

For each speech in our dataset, we calculate a measure of cosine similarity to the corpus of Höcke speeches using the same approach as for the similarity to AfD speeches described above in Equation (2.1). This measure is intended to approximate similarity to a clearly far-right and arguably more extreme populist rhetoric outside of the specific form and norms surrounding parliamentary speeches.

Populism Dictionary As our third measure of similarity to right-wing rhetoric, we construct a populism score from the German-language populism dictionary provided by Gründl (2022). This dictionary is based on distinctively populist rhetoric in German-speaking social media posts by politicians and parties in Austria, Germany, and Switzerland. It scans the speeches on a sentence level and counts the sentences in which it identifies words or phrases which are identified as populist or point to populist rhetoric.[16] Again, we standardize the resulting outcome measure with mean zero and standard deviation one such that a higher relative number of sentences with populist phrases in a speech indicates a higher degree of populism. Of the 238 words and phrases contained in the dictionary, 98 appear in the analyzed corpus of parliamentary speeches. The majority of the phrases are, according to the classification of populist ideology from Gründl (2022), associated with anti-elitism (77), 16 are about sovereignty and five are attributed to people-centrism. Among the most frequent phrases are for example *"sogenannt"* ("so-called", 4,696 appearances), *"Bürokrat"* ("bureaucrat", 513), or *"manipuliert"*("manipulated", 141), but also more distinct words like *"undemokratisch* ("undemocratic", 82), *"Elite/Eliten"*("elite/elites", 35) or *"Volksverräter"* ("traitor to the nation/people", 2) appear in our measure. A full list of terms found in our corpus is provided in Table 2.C.1 in the Appendix; for the full list of dictionary items refer to Gründl (2022).

2.3.4 Validation

Before moving to our empirical analysis, we want to verify whether our similarity measures are able to accurately capture patterns of right-wing populist rhetoric. Figure 2.1 displays the average score by party for our three different measures of similarity to right-wing rhetoric. The upper panel of Figure 2.1 provides the party averages of

[15]The excessive use of words such as *"Volk"*, oftentimes linked to Nazi ideology and rhetoric, or derogatory terms such as *"Altparteien"* (old parties) or *"Asylorkan"* (asylum hurricane) provide other striking examples.

[16]For more details on the construction of the populist dictionary measure, see Appendix Section 2.C.2 and Gründl (2022).

the standardized average cosine similarity to the whole corpus of AfD speeches in the dataset. As expected, speeches by members of the AfD themselves have the highest cosine similarity compared to all other AfD speeches.[17] With respect to the other parties, we can roughly differentiate two groups: first, speeches by conservative (CDU/CSU), social-democratic (SPD), and liberal (FDP) members of the Bundestag are less similar to the AfD than AfD speeches themselves, with the CDU/CSU being the closest in rhetorical terms. The remaining parties, the Left party and the Greens, are furthest away in terms of rhetorical similarity to the AfD. This emerging pattern is reassuring since it mirrors the ideological distribution from right to left in the Bundestag fairly well. In particular, the fact that the CDU/CSU is closest in terms of shared rhetoric and the Greens are farthest from the AfD is in line with how close these parties associate or distance themselves from the far right.

The middle panel of Figure 2.1 is analogous to the upper panel, now showing the average cosine similarity by party of MPs' speeches to the speeches by Björn Höcke. The results are very similar, except that liberals and social democrats switch their positions. However, again, speakers from the Greens and the Left party are significantly farther away in terms of rhetorical similarity to these rather extreme speeches than MPs of other parties.

Finally, the lower panel of Figure 2.1 displays the similarity to populist rhetoric in terms of the German-language populism dictionary by Gründl (2022). More specifically, it displays the frequency of the usage of distinctively populist words after standardization. The figure shows that, as expected, MPs from the AfD are by far most likely to use such populist words in their speeches.[18] With respect to the other parties, the emerging pattern differs from the previous figures: speeches from the Left party are significantly less likely to use populist words but more so than the remaining other parties. Given that the Left party has been categorized by political scientists as a populist far-left party itself, this finding is not surprising (Rooduijn et al. 2019). Overall, the observed pattern is in line with theoretical expectations and shows that also the populist dictionary approach does well at identifying populist right-wing rhetoric.[19] At the same time, this shows that the populism dictionary approach deviates from our other cosine similarity measures of right-wing populist speech and seems to capture

[17] When calculating the cosine similarity of one individual AfD speech, we leave out that speech from the sample of AfD speeches they are compared to in order to avoid mechanically higher cosine similarities.

[18] In the non-standardized scale, the AfD scores a mean populist dictionary measure of 0.99 (sd = 1.32), indicating that on average one sentence per speech contains a populist phrase. The values for the other parties are as follows: Left (mean = 0.55, sd = 0.94), SPD (mean = 0.39, sd = 0.75), CDU/CSU (mean = 0.39, sd = 1.32), FDP (mean = 0.36, sd = 0.70) and Greens (mean =0.32, sd = 0.69).

[19] In the original paper that analyzes texts from social media posts on Facebook and Twitter, Gründl (2022) finds that the AfD, followed by the Left party, has the highest score in terms of the populist dictionary. It is reassuring that we can reproduce this ranking for our different corpus of parliamentary speeches in the German Bundestag.

another aspect of AfD rhetoric.

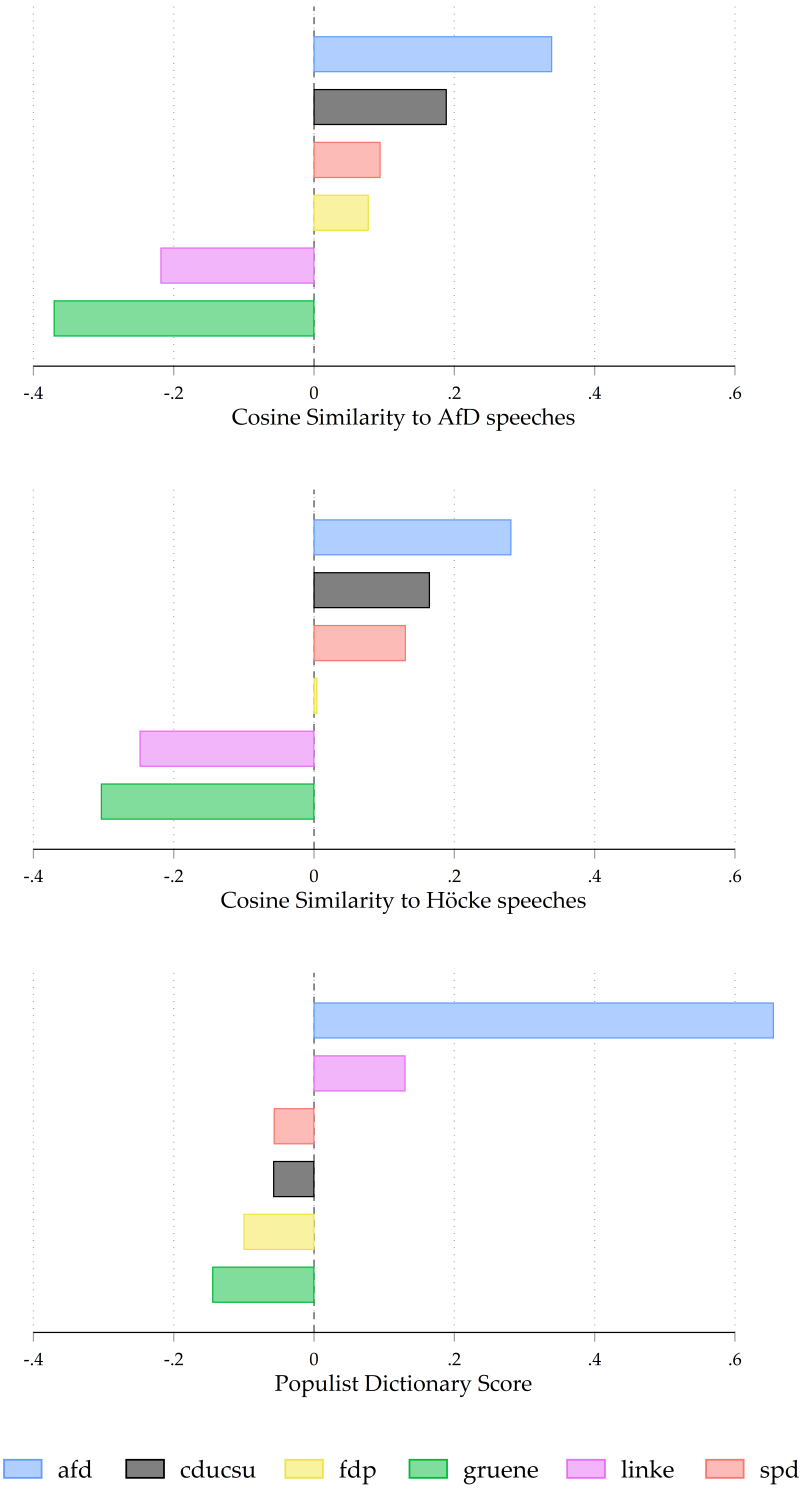
In Figure 2.2, we furthermore explore the time dimension in the usage of right-wing rhetoric in our dataset and plot the development over time of our three similarity measures. As before, the upper panel displays the average standardized cosine similarity to AfD speeches in the Bundestag, the middle panel shows the average cosine similarity to Höcke speeches, while the lower panel shows the standardized scores from the populism dictionary provided by Gründl (2022). We display the party averages by month in which the speech was recorded, with the vertical dashed lines indicating the entry of the AfD in the Bundestag after the federal election in September 2017. Importantly, there seems to be a large variation over time in terms of how similar speeches are to right-wing rhetoric across all of our three measures, and most parties seem to move together in this aspect. This indicates important time-specific aspects in Bundestag speeches, e.g., due to which topics are more frequently discussed in a month or how polarized the debate at a certain time is. This underlines the need to account for such time-specific variation in our empirical analysis which we will address with the inclusion of month fixed effects and controls generated by a Latent Dirichlet Allocation (LDA) topic model.[20]

In a last step, we also formally study the correlation between our preferred measure of AfD cosine similarity and the other measures of similarity to right-wing rhetoric in Appendix Table 2.B.1. We find that both a higher cosine similarity to Höcke speeches as well as a higher number of words from the populist dictionary significantly predicts a higher cosine similarity to AfD speeches. Importantly, both correlations remain highly significant when including speaker fixed effects, i.e., only comparing speech similarity measures within one speaker, as well as adding topic controls, month fixed effects and excluding speeches by AfD and FDP members, in line with our main empirical specification presented in Section 2.5.1.[21] Overall, the strong correlation between these three very differently constructed measures gives us confidence that we can validly identify similarity to right-wing or populist rhetoric.

[20] Details on the implementation of the topic modeling are provided in Appendix Section 2.C.3.

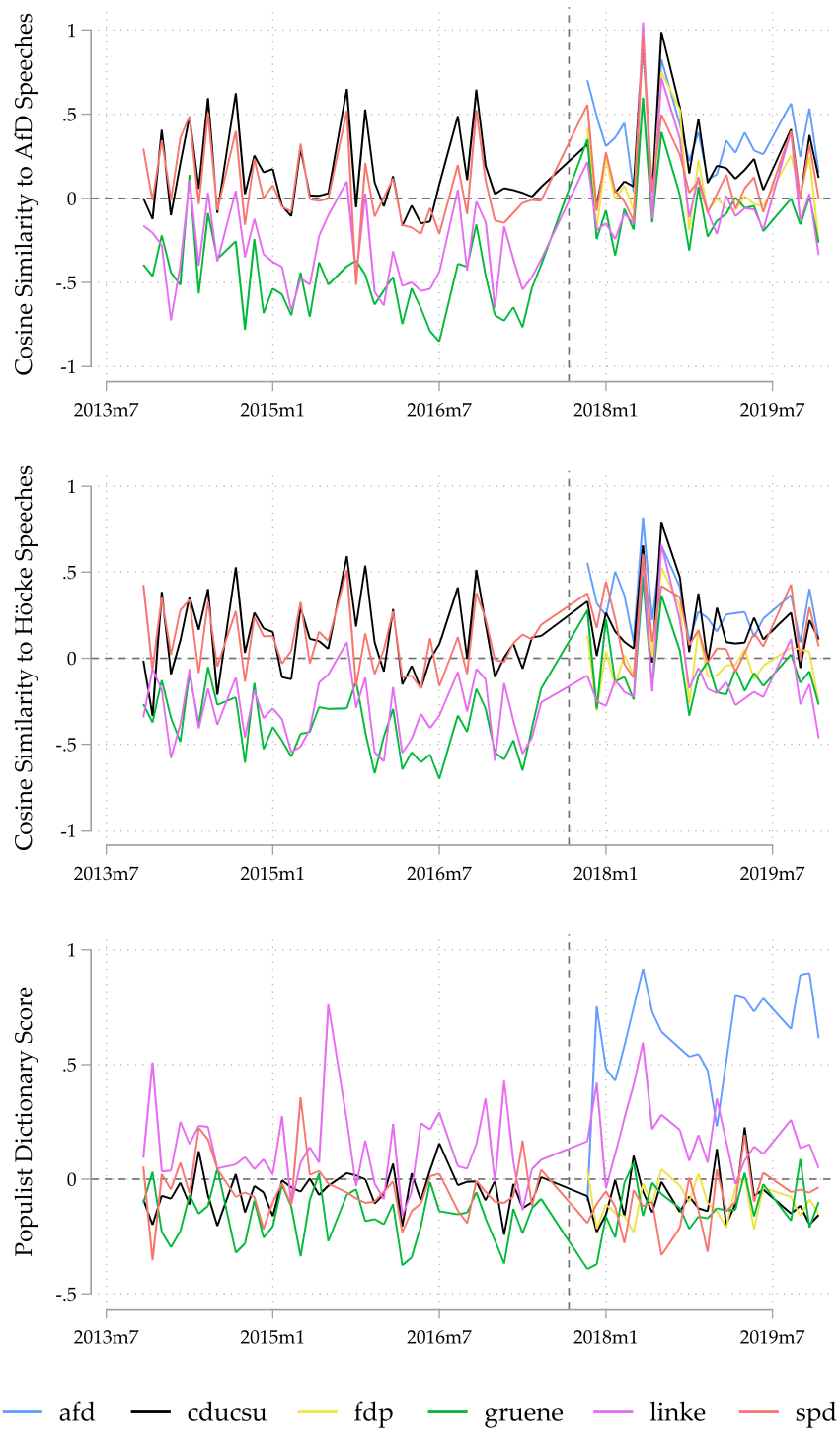
[21] We exclude speeches by the FDP as it was also not represented in the Bundestag in our pre-treatment electoral period from 2013 to 2017.

Figure 2.1: Similarity measures by party



Notes: All outcome variables have been standardized with mean zero and standard deviation one. For the construction of each outcome, the sample was restricted to speeches with a minimum length of 100 terms.

Figure 2.2: Similarity measures over time



Notes: Excludes months with few speeches (n<30) due to summer breaks and around change of legislative period (August 2015, September 2017, October 2017, July 2019). All outcome variables have been standardized with mean zero and standard deviation one. For the construction of each outcome, the sample was restricted to speeches with a minimum length of 100 terms.

2.4 Identification Strategy

According to our main hypothesis, the new presence of the AfD in the Bundestag and the active participation of right-wing populist politicians in parliamentary work might influence the language and rhetoric used by other politicians. However, simple comparisons of rhetorical similarity to the AfD across time or between parties are unlikely to identify the *causal* effect of exposure to the AfD due to a number of potentially serious endogeneity concerns. For example, the salience of different topics, especially those related to right-wing electoral success such as migration, might change over time and could simultaneously drive rhetoric similarity measures which would invalidate a simple comparison of speeches before and after the entry of the AfD.

We try to overcome such concerns by exploiting variation in *individual-level* exposure to the AfD within the parliamentary committees of the Bundestag (*Ausschüsse*). Using this novel source of variation has a number of advantages: conceptually, we study personal exposure to right-wing colleagues in repeated encounters in the context of the day-to-day work routine of members of parliament. In line with the ideas of communication accommodation theory discussed in Section 2.2, we thereby focus on the impact of human interactions on language and political rhetoric. In terms of the empirical analysis, using this source of variation across different politicians *within* the same parliament allows us to hold a number of potential confounders constant. This feature is a particular advancement compared to other studies that have exploited variation *between* different parliaments in terms of exposure to right-wing populists (Atzpodien 2022; Valentim and Widmann 2021).

In the following, we first provide a brief institutional description of the central role of committees for the functioning of the German Bundestag. Second, we provide a description of the mechanism that is used for the allocation of committee seats to different parties in parliament. We show that this allocation mechanism yields arguably exogenous variation in party-level exposure to the AfD that we can exploit to study a potential individual-level contagion effect of right-wing populism on political rhetoric.

2.4.1 Committees in the Bundestag

Political scientists have classified the Bundestag as predominantly exhibiting characteristics of a so-called *working parliament* (*Arbeitsparlament*) in which most work is done in topic-specific specialized parliamentary committees that prepare legislative proposals that are then submitted to the plenary for approval (Ismayr 2001, p. 167; Schmidt 2021, p. 148).[22] Therefore, the time spent on debating, working, and voting inside com-

[22]The opposing type of parliamentary work is the so-called *speech parliament* (*Redeparlament*) where the plenary session is the main arena in which debates are held and legislative decisions are made. The

mittees typically largely outweighs the time spent on debating and giving speeches in the plenary sessions.[23] Committees are hence the central place for policy-making and inter-party political discussions and exchanges in the Bundestag.

While the German constitution prescribes that Bundestag committees on foreign affairs, defense, petitions, and European Union affairs must be formed, the exact number and specializations of the committees are not determined and decided upon by the members of the Bundestag for each legislative period. Typically, the topical specialization of committees mirrors those of the federal ministries and their competences. As a result, the number and specialization of committees varies from one legislative period to the other, reflecting changes in the relevance and overlapping of different policy areas.[24]

Usually, however, committee meetings are not public and, therefore, speech transcripts are not available.[25] In contrast, the plenary sessions of the Bundestag are the most visible arena of parliamentary work where members of parliament hold speeches that are livestreamed and transcribed. Hence, MPs are well aware that their speeches will be visible to other members, their own party and its leadership as well as the media and, ultimately, voters. Both plenary sessions and committee meetings are typically held in the same week when the Bundestag is officially *"in session"* which occurs at least in 20 weeks per year and for which MPs usually travel to and work in their Berlin offices located in the *Reichstag* and surrounding buildings. Usually, committee meetings are scheduled for Wednesday morning and plenary sessions are held on Wednesday afternoon (Deutscher Bundestag 2022b). This scheduling sequence gives us confidence that plenary speeches might at least to some degree be given in reaction to debates in the preceding committee meetings and, hence, might give room for exerting an influence on the rhetoric used by speakers.

canonical example for such a *speech parliament* is the House of Commons of the United Kingdom, while the United States Congress is seen as the prototype of a *working parliament*.

[23]This priority of committee work can be quantified by the fact that there have been almost ten times as many committee meetings (38,731) than plenary sessions (4,106) from 1949 to 2017 as recorded in the Bundestag statistics compiled by Feldkamp (2018, pp. 214–216).

[24]For instance, the War Victims and Repatriates committee (*Ausschuss für Kriegsopfer- und Heimkehrerfragen*) played an important role in immediate post-WWII politics but the issue is no longer relevant today and the committee does not exist anymore. On the other hand, the Digital Agenda committee (*Auschuss für Digitale Agenda*), that was for the first time established after the 2013 election, represents the emergence of a new policy area.

[25]There are some exceptions when committee meetings are public, often due to a hearing that deviates from standard committee procedure. For example, committees may wish to gather information from external experts on certain legislative proposals, with the session focusing on speeches given by invited experts and not on speeches given by MPs who rather ask questions.

2.4.2 Allocation of Committee Seats

The size of committees, i.e., the number of members that have full voting rights, is not fixed but depends on the importance of their respective policy agenda and the amount of legislative work involved. The different parties represented in the Bundestag jointly decide on the size of committees at the beginning of each legislative period when committees are formed. In the main periods of interest in our empirical analysis, there exist 23 main committees in the 18$^{\text{th}}$ Bundestag (2013-17) that have between 14 and 46 members, while in the 19$^{\text{th}}$ Bundestag (2017-21) there are 24 main committees with 14 to 49 members.[26]

Once the absolute size of committees is established, seats are allocated to parties on the premise of ensuring proportional representation, i.e., the share of seats of a party in a given committee should equal the share of seats this party has in the Bundestag. As the number of available seats in a committee is finite and relatively low, a perfect proportional representation is, however, not always attainable and committee shares might deviate from the share of seats in the plenary. In order to ensure a fair representation and, in particular, to avoid discrimination against smaller parties, the Bundestag uses the *Sainte-Laguë/Schepers* rule for the allocation of committee seats to parties.[27] The rule is based on the idea of iteratively calculating an allocation quotient from the following formula: for each party p and its already allocated number of seats s, an allocation quotient Q is calculated based on the share of the party's seats in parliament V:

$$\mathrm{Q}_p = \frac{V_p}{2s_p + 1} \tag{2.3}$$

[26] Table 2.B.3 and Table 2.B.4 in the Appendix display the name and size of the committees in the 18$^{\text{th}}$ and 19$^{\text{th}}$ Bundestag, respectively, as well as the absolute number of seats assigned to each party in a given committee. We exclude a number of non-standard committees from our analysis: the committee on election audit (*Wahlprüfungsausschuss*) is excluded as it has the specific task of auditing whether the elections for the Bundestag and the European Parliament were conducted lawfully and without intervention. The committee meets significantly less often than other committees and consisted of only 9 members in both periods of interest. The mediation committee (*Vermittlungsausschuss*) is the common committee between the Bundestag and the Bundesrat, which is the parliamentary body representing the 16 German states at the federal level. Its main function is to mediate between the interests of the Bundestag and the Bundesrat in case of disagreement in the legislative process. As this committee consists of both members from the Bundestag and Bundesrat, we exclude it from our analysis. We also exclude the joint committee (*Gemeinsamer Ausschuss*) as its only function is to work as an emergency parliament in case of a state of defence and does not regularly meet. Furthermore, we exclude sub-committees (*Unterausschüsse*) that can be formed within the main committees, as well as five investigative committees (*Untersuchungsauschuss*) that are temporarily formed *ad-hoc* to investigate specific cases of potential misconduct by the government. Finally, we also exclude the two temporary main committees (*Hauptausschuss*) that were formed for one month in 2013 and two months in 2017/18 as a stand-in committee until the constitution of the main committees while negotiations for the formation of a coalition government were on-going.

[27] The rule has been applied for the allocation of committee seats in the Bundestag since 1980, and since 2009 it also determines the allocation of plenary seats in the Bundestag as well as the allocation of the German seats in the European Parliament.

An iterative procedure that starts with $s = 0$ for all parties allocates a seat to the party with the highest quotient. If more than one party has the highest quotient, the seat is randomly allocated to one party rather than the other. After the allocation of the seat, the quotient is recalculated. The process ends when all available seats in a committee have been allocated.

Figure 2.3 visualizes how the *Sainte-Laguë/Schepers* rule leads to plausibly exogenous variation in the share of seats assigned to a party, here for the example the AfD. Panel (A) shows the allocated number of seats to all parties by the *Sainte-Laguë/Schepers* rule based on different total committee sizes. Panel (B) visualizes the change in the allocated share of AfD members for differently sized committees. There are a number of distinct jumps in the share of AfD members when the marginal additional seat in a committee is allocated to the AfD. The dashed vertical lines indicate the number of seats in existing committees Hence, we can observe that there were committee sizes where the AfD was "lucky" in the sense of being overrepresented due to having just gained the next additional seat and in areas where the AfD was "unlucky", respectively. Panel (C) shows the actual distribution of speeches held after 2017 in our sample by the respective share of AfD members in a speaker's committees. Reassuringly, we find substantial variation in relative AfD exposure in our sample of speeches.[28]

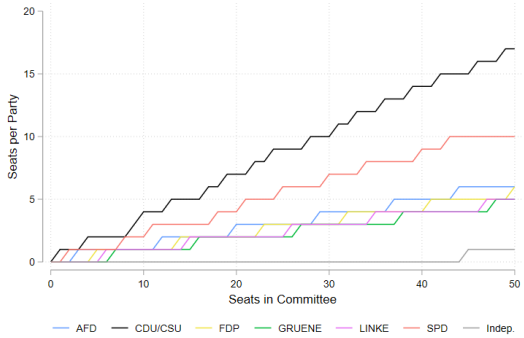
The share of AfD members in a given committee of a certain size therefore arguably features an exogenous component.[29] Two politicians in committees of comparable size might therefore have a different relative exposure to far-right AfD politicians. For example, a politician represented in the *Digital Agenda* committee (with a total of 21 members) has to work on a regular basis with three colleagues from the AfD, implying a relative share of 14.3% AfD members. In contrast, a politician in the committee for *Culture and Media Affairs* (with a total of 18 members) faces only two AfD members in her committee meetings with a relatively lower share of 11.1% AfD members. Table 2.B.4 in the Appendix summarizes the distribution of AfD members across all Bundestag committees, exhibiting variation in the relative share of AfD members in committees of different size. From the perspective of politicians of other parties, this implies variation in the exposure to right-wing populist politicians and their ideology in their day-to-day parliamentary work. In the following empirical analysis, we will exploit this variation to analyze the effect of this exposure on the rhetorical similarity to right-wing populism in parliamentary speeches given by these politicians.

[28] Note that the empirical mean of 12.97% of AfD committee members in our speech sample corresponds exactly to the relative share of the 92 AfD members among the 709 total Bundestag MPs.

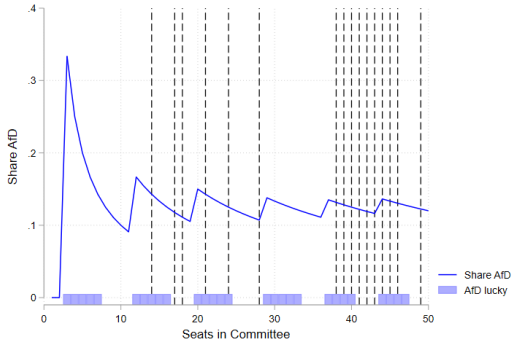
[29] In the spirit of a regression discontinuity design, one could assume that politicians in committees are not able to manipulate the size of committees to be either just to the right or just to the left of a jump in the share of AfD members function. As shown in Appendix Figure 2.A.2, relative committee sizes are quite stable over time and, in particular, there seems to be little movement in relative committee sizes between the 2013-2017 and 2017-2021 legislative periods that are of interest for our empirical analysis.

Figure 2.3: Sainte-Laguë/Schepers Rule

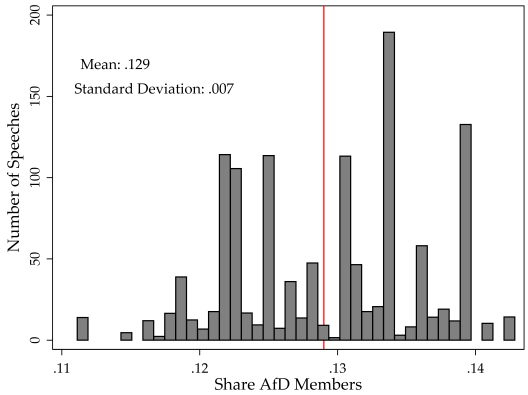
(a) Absolute Number of Members



(b) Share of AfD Members



(c) Distribution Share of AfD Members



Notes: Panel (A) shows the absolute number of members for each party for different sizes of committees according to the *Sainte-Laguë/Schepers* rule for the 2017-2021 legislative period. Panel (B) shows the assigned share of AfD members based on the *Sainte-Laguë/Schepers* rule for different potential sizes of committees. Shaded ranges on x-axis indicate seat numbers for committees that are midpoints between seat numbers where the AfD gains an additional seat according to the *Sainte-Laguë/Schepers* rule. Dashed vertical lines indicate the total number of seats in existing committees. Panel (C) shows the distribution of the associated share of AfD committee members for all speeches in our sample held after September 2017.

2.5 Results

2.5.1 Main Results

Our goal is to estimate the causal effect of individual exposure to radical right-wing AfD politicians on similarity to right-wing populist rhetoric. However, a simple comparison of the relative committee exposure to AfD members on speech similarity might suffer from selection bias. As we have previously argued, the committee allocation procedure leads to variation in the share of seats assigned to a party, and hence individual-level variation in exposure to the AfD; yet, individual assignment of politicians to committees might still be endogenous. For example, parties could strategically select politicians for committees with relatively higher AfD presence due to some individual characteristics such as ideological solidity or distance to right-wing populism.[30]

To address such endogeneity concerns, we exploit that our data comprises speeches before the AfD's entry into the Bundestag. We run a difference-in-differences regression comparing speeches of the *same* politicians before and after being differentially exposed to right-wing politicians. In particular, we estimate the following regression model:

$$Similarity_{ist} = \beta \cdot Share\,AfD\,Members_{s(i)} \times Post_t + X_i'\gamma + \delta_t + \phi_s + \epsilon_{ist} \tag{2.4}$$

where $Similarity_{ist}$ is one of our measures of similarity to right-wing rhetoric for the plenary speech i held by speaker s at time t. Our main explanatory variable $Share\,AfD\,Members_{s(i)}$ measures the share of AfD politicians among all full members of the committee of which politician s is a full member in the 19$^{\text{th}}$ Bundestag.[31] $Post_t$ is a dummy variable equal to 1 if plenary speech i was held after the entry of the AfD in September 2017. As shown in Figure 2.2, there is substantial variation over time in rhetorical similarity, for which we account by including month fixed effects δ_t as well as controls for 20 topics generated by a Latent Dirichlet Allocation (LDA) topic model (X_i).[32] Crucially, we also include speaker fixed effects ϕ_s that control for all time-invariant factors related to an individual speaker. The inclusion of this rel-

[30] In Table 2.B.5 in the Appendix, we empirically investigate such selection and regress the share of AfD committee members on a number of observable individual characteristics of MPs. We find that female and older MPs tend on average to sit in committees with relatively fewer AfD members. Interestingly, results in column (2) suggest that electoral competition with the AfD – as measured by the absolute vote share and relative distance to the AfD in an MP's electoral district in the last federal election – does not seem to predict assignment into committees. Nevertheless, these findings confirm that our empirical strategy needs to account for potential individual-level selection into committees.

[31] If a politician is a full member in multiple committees, we assign her the average share of AfD members across all respective committees.

[32] Details on the implementation of the topic modeling are provided in Appendix Section 2.C.3.

atively demanding set of 437 speaker fixed effects should alleviate concerns relating to unobserved characteristics influencing political speech and selection into committees. Throughout all specifications, we cluster standard errors on the committee times electoral period level. Our main coefficient of interest is given by β: a positive and significant coefficient would indicate that more AfD members in a given committee increase similarity to right-wing rhetoric. However, a negative effect would suggest that direct exposure to AfD politicians might lead members of other parties to rhetorically distinguish themselves more from right-wing speech.[33]

Table 2.1 presents our main results from estimating the regression specification as shown in Equation 2.4. Column (1) shows the effect on our preferred measure of rhetorical similarity, the standardized average cosine similarity to all AfD speeches, by comparing speeches given by the *same* politicians before and after the entry of the AfD into the Bundestag. Furthermore, topic controls and month fixed effects assure that the estimated effect is not confounded by time- or topic-specific trends in plenary speeches. We obtain a positively estimated coefficient for β significant at the 10 percent level, implying an increase in similarity to AfD rhetoric with higher exposure to right-wing politicians. The magnitude of the effect is non-negligible: comparing a politician in a committee with the highest to one in a committee with the lowest relative AfD exposure (corresponding to an increase in the share of AfD members by 0.03 as indicated in Table 2.B.2) increases the AfD cosine similarity by 0.1 (3.356×0.03) of a standard deviation. To put this into perspective, this increase corresponds roughly to the 0.09 difference in the average standardized AfD cosine similarity between speakers of the center-left SPD and the center-right CDU/CSU, as shown in the upper panel of Figure 2.1.

When looking at alternative measures of rhetorical similarity to right-wing speech, we find very comparable results: Column (2) shows a positive and strongly significant effect of higher committee exposure to the AfD on the average cosine similarity to extra-parliamentary speeches held by extreme right-wing AfD politician Björn Höcke. Column (3) reports a likewise positive effect on the number of sentences with populist words as classified in the German-language populism dictionary by Gründl (2022). As all outcomes were standardized to allow for easier comparison of magnitudes, we can further note that the estimated effect sizes are reassuringly similar.

Taken together, our main results reported in Table 2.1 provide evidence for a contagious effect of direct exposure to far right-wing politicians on using similar language

[33]Note that β is not mechanically driven by AfD speeches. AfD speeches will both feature a higher AfD cosine similarity and tend to come from politicians sitting in committees with high shares of AfD members. However, the difference-in-differences design with speaker fixed effects requires that speeches included in our analysis come from politicians who were present in both legislative periods, i.e., before and after the entry of the AfD. Thus, the sample of speeches in our design excludes speeches from AfD politicians (as well as speeches by the FDP who also only re-entered parliament in 2017).

Table 2.1: Main Results

	AfD Similarity	Höcke Similarity	Pop. Dictionary
	(1)	(2)	(3)
Share AfD × Post	3.356*	3.868***	4.194**
	(1.932)	(1.321)	(1.630)
Topic Controls	✓	✓	✓
Month FE	✓	✓	✓
Speaker FE	✓	✓	✓
Observations	17,383	17,383	17,383

Notes: Table reports coefficients and standard errors from linear regressions as laid out in Equation 2.4. The independent variable of interest is the interaction between the (average) share of AfD members of all committees in which a politician is a full member and an indicator whether the speech was recorded in the 19th German Bundestag (2017-2021). The dependent variables are as follows: (Column 1) the standardized average cosine similarity to AfD speeches after pre-processing and tf-idf vectorization; (Column 2) the standardized average cosine similarity to speeches by Björn Höcke after pre-processing and tf-idf vectorization; (Column 3) the standardized number of sentences with words from the German-language populist dictionary by Gründl (2022). Topic controls are derived from a 20-topic LDA model. The sample comprises plenary speeches by members of the German Bundestag held between October 2013 and December 2019 with a minimum length of 100 terms from parties that were represented throughout the whole period (CDU/CSU, SPD, The Left, and Alliance90/The Greens). Standard errors clustered at the committee times electoral period level are reported in parentheses: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

in public speeches. Notably, this effect can even be detected *within* politicians who seem to adapt their rhetoric once they have to deal with more extreme right-wing politicians in their daily committee work after 2017. Furthermore, the change in language is not only detectable in similarity to rhetoric used by the AfD itself in parliamentary speeches, but also extends to arguably more extreme rhetoric, as showcased by the similarity to Höcke speeches, and the usage of distinctively populist vocabulary.

2.5.2 Robustness Checks

FDP Placebo So far, our results have suggested a contagious effect of exposure to far-right politicians on other MPs' rhetoric. Yet, it remains unclear to what extent such contagion is specific to (right-wing) populism or whether politicians generally adopt their language use in reaction to any increased exposure to newly elected colleagues, independent of their ideology. The specific setting of the German Bundestag provides us with a useful placebo exercise to shed more light on this question: at the same time as the AfD was elected into parliament for the first time at the 2017 federal elections,

Table 2.2: FDP Placebo

	AfD Similarity	Höcke Similarity	Pop. Dictionary	FDP Similarity (res.)
	(1)	(2)	(3)	(4)
Share FDP × Post	4.189	1.716	-1.432	-0.583
	(2.732)	(1.481)	(1.625)	(0.577)
Topic Controls	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
Speaker FE	✓	✓	✓	✓
Observations	17,383	17,383	17,383	17,383

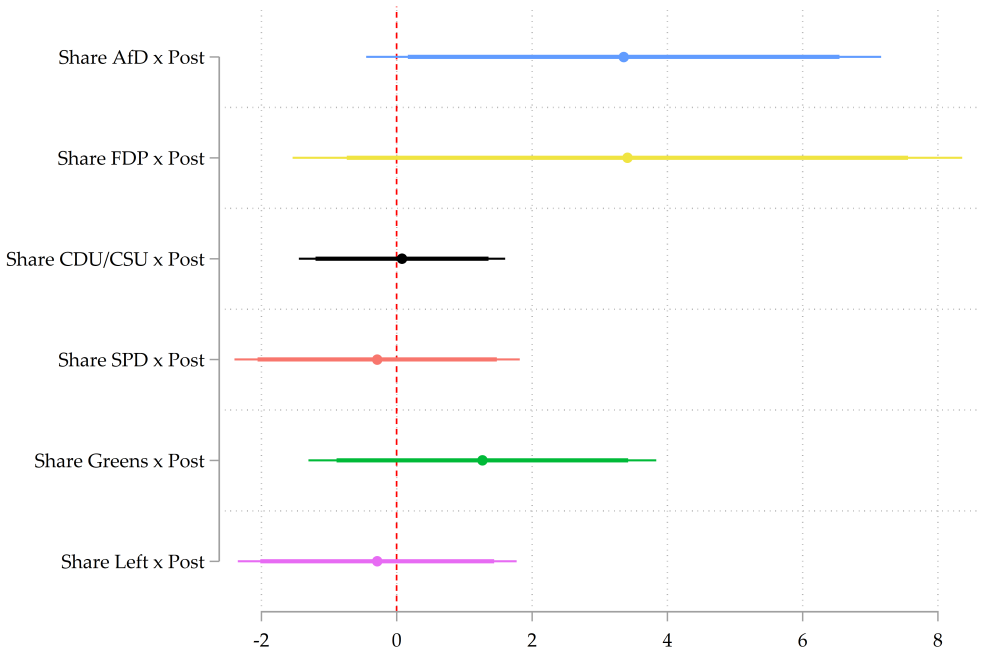
Notes: Table reports coefficients and standard errors from linear regressions as laid out in Equation 2.4. The independent variable of interest is the interaction between the (average) share of AfD members of all committees in which a politician is a full member and an indicator whether the speech was recorded in the 19th German Bundestag (2017-2021). The dependent variables are as follows: (Column 1) the standardized average cosine similarity to AfD speeches after pre-processing and tf-idf vectorization; (Column 2) the standardized average cosine similarity to speeches by Björn Höcke after pre-processing and tf-idf vectorization; (Column 3) the standardized number of sentences with words from the German-language populist dictionary by Gründl (2022); (Column 4) the standardized average cosine similarity to FDP speeches after pre-processing, tf-idf vectorization and residualizing on standardized average AfD cosine similarity. Topic controls are derived from a 20-topic LDA model. The sample comprises plenary speeches by members of the German Bundestag held between October 2013 and December 2019 with a minimum length of 100 terms from parties that were represented throughout the whole period (CDU/CSU, SPD, The Left, and Alliance90/The Greens). Standard errors clustered at the committee times electoral period level are reported in parentheses: $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$.

also the liberal non-populist FDP re-entered the Bundestag after not having been represented in the 18th legislative period between 2013 and 2017.[34] This allows us to re-estimate the baseline regression framework presented in Equation 2.4 but changing the "treatment" variable to measure the intensity of personal exposure to non-populist FDP instead of populist AfD politicians in committees.

The results of this placebo-style exercise are presented in Table 2.2. Columns (1) to (3) show the effect of relatively higher FDP exposure on our three measures of rhetorical similarity to right-wing AfD speeches, analogous to the structure of Table 2.1. We do not find any of the estimated coefficients to be significantly different from zero. Column (4) now additionally tests whether relatively higher FDP exposure in committees also makes other politicians use more similar rhetoric to the FDP. In order to avoid that FDP similarity partially absorbs similarity to the AfD as well, we first residualize average FDP cosine similarity on average AfD cosine similarity. Intriguingly, when regressing this residualized FDP cosine similarity on the share of FDP commit-

[34]The FDP (*Free Democratic Party*) is the main liberal political party in Germany and typically associated with the center or center-right of the political spectrum. The FDP has been a traditional established force of the German party system since the end of World War II, having been represented in the Bundestag since 1949 and having served repeatedly as junior coalition partner in both CDU/CSU-led (1949–1956, 1961–1966, 1982–1998, 2009–2013) and SPD-led (1969–1982, since 2021) governments. In 2013, it failed to meet the 5% electoral threshold for parliamentary representation for the first time in its history, but was reelected in 2017.

Figure 2.4: General Accommodation Effects



Notes: Figure shows coefficients and confidence intervals (90 and 95 percent) from separate linear regressions as laid out in Equation 2.4. The full corresponding regression results can be found in Table 2.B.6. For each estimated coefficient the variables are defined as follows: the independent variable of interest is the interaction between the (average) share of the respective party members of all committees in which a politician is a full member and an indicator whether the speech was recorded in the 19th German Bundestag (2017-2021). The dependent variables is the standardized average cosine similarity to speeches of that respective party after pre-processing and tf-idf vectorization. The sample comprises plenary speeches by members of the German Bundestag held between October 2013 and December 2019 with a minimum length of 100 terms from parties that were represented throughout the whole period (CDU/CSU, SPD, The Left, and Alliance90/The Greens), excluding members of the respective party. Standard errors are clustered at the committee times electoral period level.

tee members, we also do not find a statistically significant effect with the coefficient being close to zero. The absence of any effect for FDP exposure seems to suggest that the observed contagion effects are indeed specific to being exposed to (right-wing) populist rhetoric and ideology.

General Accommodation We can further explore the idea that politicians – consciously or unconsciously – accommodate their language in general to *any* exposure and interaction with colleagues of a different ideology who are using dissimilar rhetoric. We therefore extend our difference-in-differences framework to analyze potential contagion effects for all parties represented in the Bundestag. The AfD and FDP (re)entered the Bundestag in September 2017, meaning that previous AfD and FDP exposure in committees was zero. For the other parties, our treatment captures the *change* in relative committee exposure between electoral periods.

The main coefficients from this exercise are visually represented in Figure 2.4.[35] Most importantly, we find that *only* direct exposure to right-wing populist AfD members significantly affects the rhetoric employed by politicians of other parties slanting the language into the AfD's direction. On the other hand, relatively higher exposure to politicians of the established parties such as the CDU/CSU and SPD does not lead MPs to adopt their rhetoric.[36] The remaining cases of the Green and Left party are insightful as speakers from these parties exhibit the most distinctive rhetoric with respect to the AfD as can be seen in Figure 2.1. We might expect that contagion effects are especially salient for parties using more distinct language from the average Bundestag politician. However, we also do not find significant effects on rhetorical similarity with higher exposure to committee members from these parties. This contributes to our assessment that the estimated contagion effects seem to be specific to exposure to right-wing populism.

Speech Length In our baseline specification, we restricted our sample to speeches with a minimum length of 100 terms in order to select sufficiently long speeches which should better capture distinctively right-wing rhetoric used by the AfD.[37] In Table 2.3 we therefore repeat our main difference-in-differences analysis as laid out in Equation 2.4 for different restrictions on the minimum number of terms in a speech. Reassuringly, the estimated coefficients remain largely stable for all three employed rhetorical similarity measures. Only in the case of no speech length restrictions – potentially containing many short (non-ideological) remarks – and when restricting our sample to contain mostly longer speeches – significantly reducing the sample size – do the estimated coefficients become smaller and lose statistical significance.

2.5.3 Effect Heterogeneities

Our results so far have shown that politicians adapt their own rhetoric in reaction to being directly exposed to newly arriving colleagues using a radically different right-wing language. In the following, we want to provide some suggestive evidence under *what conditions* and *why* political actors might revert to such changes in their publicly displayed language use. To this end, we test whether our treatment effect varies with respect to a number of observable characteristics of politicians. We therefore adapt our baseline estimation strategy to a triple difference-in-differences framework to test

[35] The corresponding regression results can be found in Table 2.B.6 in the Appendix.

[36] In the case of the FDP, in difference to the results presented in Table 2.2 we do not residualize our results on AfD similarity, as we want to compare the uncontrolled effect on speech similarity for all parties. Nevertheless, we can again not reject that the positive coefficient estimate is statistically different from zero.

[37] Indeed, as can be seen in Figure 2.A.3 in the Appendix, the more we restrict the sample to include longer speeches, the better the cosine similarity measure becomes at identifying AfD speeches and, hence, arguably at capturing distinctively right-wing rhetoric.

Table 2.3: Minimum Speech Length Restrictions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Minimum Terms	0	20	30	50	100	200	400	600
Panel A: [Std.] Cosine Similarity to AfD Speeches								
Share AfD × Post	2.423	4.374**	4.108**	3.828**	3.356*	2.800	0.796	1.521
	(1.946)	(1.924)	(1.753)	(1.783)	(1.932)	(1.791)	(2.098)	(2.621)
Observations	22,705	20,958	20,442	19,396	17,383	14,750	12,754	8,497
Panel B: [Std.] Cosine Similarity to Höcke Speeches								
Share AfD × Post	2.782*	4.128***	4.085***	3.896***	3.868***	3.633***	2.247	2.975
	(1.542)	(1.374)	(1.301)	(1.311)	(1.321)	(1.339)	(1.621)	(1.872)
Observations	22,705	20,958	20,442	19,396	17,383	14,750	12,754	8,497
Panel C: [Std.] Populist Dictionary Score								
Share AfD × Post	3.698***	4.192***	4.190***	4.379***	4.194**	4.869***	4.539**	5.395**
	(1.384)	(1.493)	(1.507)	(1.574)	(1.630)	(1.636)	(1.888)	(2.458)
Observations	23,216	20,958	20,442	19,396	17,383	14,750	12,754	8,497
Topic Controls	✓	✓	✓	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓	✓	✓	✓
Speaker FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Table reports coefficients and standard errors from linear regressions as laid out in Equation 2.4. Across all panels, the independent variable of interest is the interaction between the (average) share of AfD members of all committees in which a politician is a full member and an indicator whether the speech was recorded in the 19th German Bundestag (2017-2021). The dependent variables are as follows: (Panel A) the standardized average cosine similarity to AfD speeches after pre-processing and tf-idf vectorization; (Panel B) the standardized average cosine similarity to speeches by Björn Höcke after pre-processing and tf-idf vectorization; (Panel C) the standardized number of sentences with words from the German-language populist dictionary by Gründl (2022). Throughout columns (1) to (8), the sample is restricted to speeches with a minimum number of terms as shown in the column head, which is the sample used to construct the respective outcome variables and standardize with mean zero and standard deviation one. Topic controls are derived from a 20-topic LDA model. The sample comprises plenary speeches by members of the German Bundestag held between October 2013 and December 2019 from parties that were represented throughout the whole period (CDU/CSU, SPD, The Left, and Alliance90/The Greens). Standard errors clustered at the committee times electoral period level are reported in parentheses: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

for effect heterogeneities in the following way:

$$\begin{aligned} Similarity_{ist} = &\beta_1 \cdot Share\ AfD\ Members_{s(i)} \times Post_t + \beta_2 \cdot Z_{s(i)} \times Post_t + \\ &\beta_3 \cdot Share\ AfD\ Members_{s(i)} \times Post_t \times Z_{s(i)} + \\ &X_i'\gamma + \delta_t + \phi_s + \epsilon_{ist} \end{aligned} \tag{2.5}$$

where $Z_{s(i)}$ is an observable characteristic of speaker s giving speech i. All other variables are defined in the same way as described in Equation 2.4. Our main coefficient of interest in this exercise is given by β_3 which tells us how the within-speaker effect of relative AfD exposure on rhetorical similarity differs by characteristic Z.

Table 2.4 shows the results of estimating Equation 2.5 with our preferred outcome measure of rhetorical similarity – standardized average cosine similarity to AfD

speeches – for four different speaker-specific characteristics.[38] In column (1), we compare how the treatment effect of relatively higher exposure to AfD members in committees differs for female relative to male speakers. As can be seen by the positive coefficient estimate for β_3, we find that female politicians are significantly more likely to slant their political rhetoric towards the AfD when having relatively more contact with AfD colleagues in committees. In fact, as can be seen by the estimate for β_1, the effect is not significantly different from zero for male speakers. This is consistent with existing research from social psychology and communication science that has highlighted differences in the communicative behavior of men and women. In particular, some studies suggest that female speakers are more prone to accommodate their communication style and match their language patterns to their conversation partners (Giles and Ogay 2007; Palomares et al. 2016). In column (2), we study whether the contagion effect differs by the age of a politician. Older politicians with more experience and seniority might be less susceptible to accommodation. However, we do not find that the effect of higher AfD exposure on rhetorical similarity differs by age.

Finally, in columns (3) and (4), we explore the role of political competition and electoral pressure in *strategic* changes of political rhetoric. As discussed in Section 2.2, individual accommodation to right-wing speech might follow strategic motives with respect to political support: with increasing success of right-wing populism, politicians might be able to win support from the populists' electoral base by using a similar language. We test this conjecture by adding information on the intensity of local competition in the electoral districts of Bundestag MPs.[39] Column (3) shows the effect of interacting our main treatment variable with the AfD's vote share scored in an MP's local district in the 2017 federal election, a measure of the absolute level of populist right-wing support.[40] We do not find that this differentially explains within-speaker changes in political rhetoric towards the AfD. In column (4), we instead use the absolute distance of the MP's own first vote share to that of the AfD's local candidate. Arguably, this constitutes a proxy measure for the intensity of local electoral competition with right-wing populists. Here we find that the higher the distance to the AfD vote share, i.e., the less the AfD was a direct electoral competitor in the 2017 election, the weaker the estimated contagion effect of AfD exposure on political rhetoric. In terms of the magnitude interpretation discussed in Section 2.5.1, a one standard de-

[38] In Table 2.4, all continuous interacted speaker characteristics have been standardized with mean zero and standard deviation one to ease interpretation and comparability.

[39] As not all Bundestag MPs ran as candidates in local electoral districts but sometimes only as candidates on state-wide party lists, we cannot assign all speakers in our dataset to electoral districts. Therefore, the number of observations in these estimations is slightly reduced.

[40] Note that we use the share of first votes (constituency votes for individual candidates) recorded for the AfD in these exercises, as we are interested in the role of *local* electoral competition a specific candidate is facing. Results remain unchanged when instead using the AfD share of second votes, i.e., votes for the state-wide party list instead of individual candidates.

Table 2.4: Effect Heterogeneity by Speaker Characteristics

	AfD Cosine Similarity			
	(1)	(2)	(3)	(4)
Share AfD × Post	0.575	3.382	4.228**	4.487**
	(1.702)	(2.085)	(1.938)	(1.768)
Female × Post	-0.958*			
	(0.518)			
Share AfD × Post × Female	7.658*			
	(4.054)			
Age × Post		0.176		
		(0.215)		
Share AfD × Post × Age		-1.298		
		(1.681)		
AfD Vote Share × Post			-0.026	
			(0.268)	
Share AfD × Post × AfD Vote Share			0.120	
			(2.102)	
Distance to AfD × Post				0.741***
				(0.212)
Share AfD × Post × Distance to AfD				-5.983***
				(1.638)
Topic Controls	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
Speaker FE	✓	✓	✓	✓
Observations	17,383	17,383	16,483	16,483

Notes: Table reports coefficients and standard errors from linear regressions as laid out in Equation 2.5. The dependent variable is the standardized average cosine similarity to AfD speeches after pre-processing and tf-idf vectorization. Share AfD describes the (average) share of AfD members of all committees in which a politician is a full member in the 19th Bundestag (2017-2021). Post is an indicator variable equal to 1 if the speech was recorded in the 19th Bundestag (2017-2021). Female is an indicator variable equal to 1 if the speaker is female. Age refers to the age of a speaker in years as of the opening of the 19th Bundestag (October 24, 2017). AfD Vote Share measures the first vote share of the AfD (in percent) in an MP's electoral district in the 2017 federal election. Distance to AfD measures the absolute distance of the MP's own first vote share to the AfD first vote share (in percentage points) in the 2017 federal election. All continuous interaction variables (Age, AfD Vote Share, Distance to AfD) have been standardized with mean zero and standard deviation one. Topic controls are derived from a 20-topic LDA model. The sample is restricted to plenary speeches held between October 2013 and December 2019 with a minimum length of 100 terms by speakers from parties that were represented throughout the whole period (CDU/CSU, SPD, Alliance90/The Greens, The Left). Standard errors clustered at the committee times electoral period level are reported in parentheses: $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

viation increase in distance to the AfD sizeably reduces rhetorical similarity by 0.18 (5.983×0.03) of a standard deviation. In sum, these results suggest that speakers seem to strategically adapt their political rhetoric to use more similar language to the AfD in response to higher electoral pressure from the far right.

2.6 Conclusion

The first-time entry of a right-wing populist party to the German Bundestag presented a novel situation for incumbent politicians, in particular with respect to being personally in contact with far-right AfD politicians. In this paper, we exploit quasi-exogenous variation in allocation of MPs to committees to generate individual-level variation in the intensity of such contact with the AfD. We have shown that higher exposure to the AfD has a contagious effect on the language employed by mainstream politicians in terms of converging towards a more similar right-wing rhetoric. Our results are robust to different measures of rhetorical similarity and seem to be specific to right-wing populism. Furthermore, we find some evidence that suggests strategic motives related to local electoral competition behind individual changes in political rhetoric.

A few words of caution are in order: the observed convergence in the usage of similar right-wing language does not necessarily imply that politicians also ideologically converge towards the AfD, i.e., become more right-wing populist themselves. Rather, our measures of rhetorical similarity – be they based on cosine similarity or a dictionary approach – capture how something is said (in terms of words used) and only to a certain extent what is meant (in terms of implied content). For example, we cannot rule out that politicians take up and cite phrases introduced by the AfD with another, or even opposite, political message intended. Nevertheless, our results clearly show how the novel and rather extreme AfD rhetoric finds its way into parliament and spreads even among mainstream politicians. On the one hand, this implies that even in a setting were they do not hold any formal political power, right-wing populists can exert a certain agenda-setting power. On the other hand, regardless of any potential ideological convergence, previous research has highlighted that *"words have consequences"* and even minor changes in rhetoric can already lead to changes in the acceptability of social norms and behavior even beyond the parliamentary arena (Bursztyn, Egorov, and Fiorin 2020; Djourelova 2023; Müller and Schwarz 2020, 2021).

We see at least two interesting avenues for future research departing from these observations. First, while we have analyzed contagion of right-wing rhetoric within political elites, we know less about the effects of the novel and distinctively right-wing language used by the AfD in the Bundestag on the general public. This is especially relevant as the AfD seems to deliberately target a wider audience by diffusing content and video recordings from parliamentary speeches via social media. Indeed, the AfD

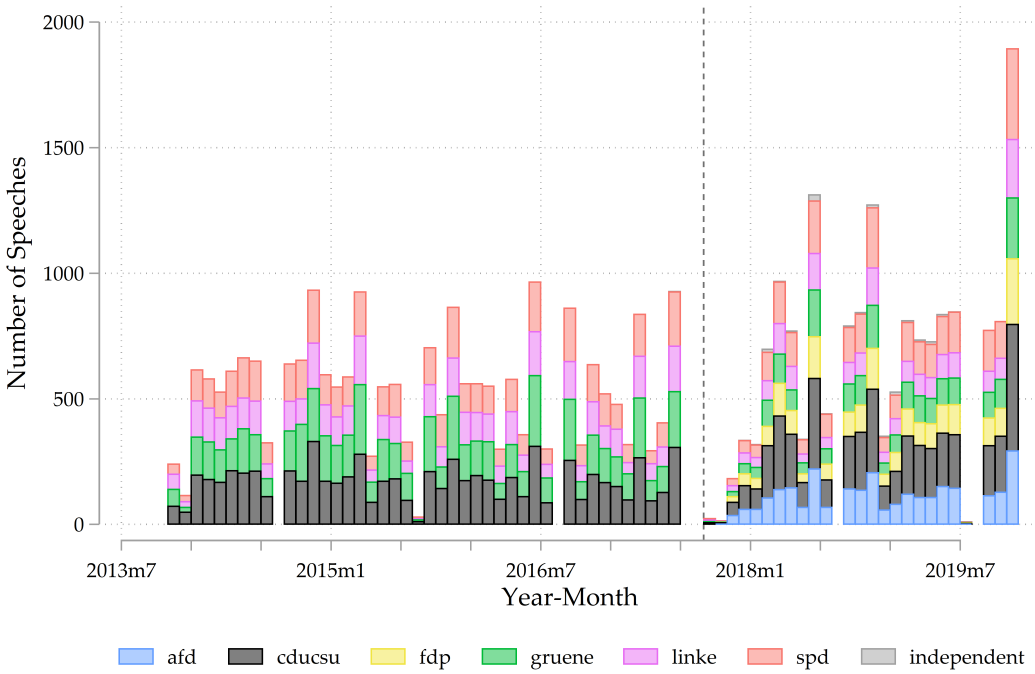
has by far the largest number of followers on various social media platforms among all parties represented in the Bundestag.[41] Second, while for our empirical analysis we have implicitly assumed that the AfD's own rhetoric remains constant at least in the short-run, it might be worthwhile to explore if and how right-wing populists themselves adopt their language when in regular contact with more moderate mainstream politicians.

[41] For example, the YouTube channel of the AfD's parliamentary group in the Bundestag has about 300,000 followers, compared to 66,000 for the Left Party, 26,000 for the Greens, 23,000 for the FDP, and 3,500 for each SPD and CDU/CSU (as of January 9, 2023). A similar ranking emerges on Facebook, where the AfD's parliamentary group has more than 250,000 followers, almost double as many as the 140,000 followers of the second-largest page by the Left Party's parliamentary group (as of January 9, 2023).

Appendix to Chapter 2

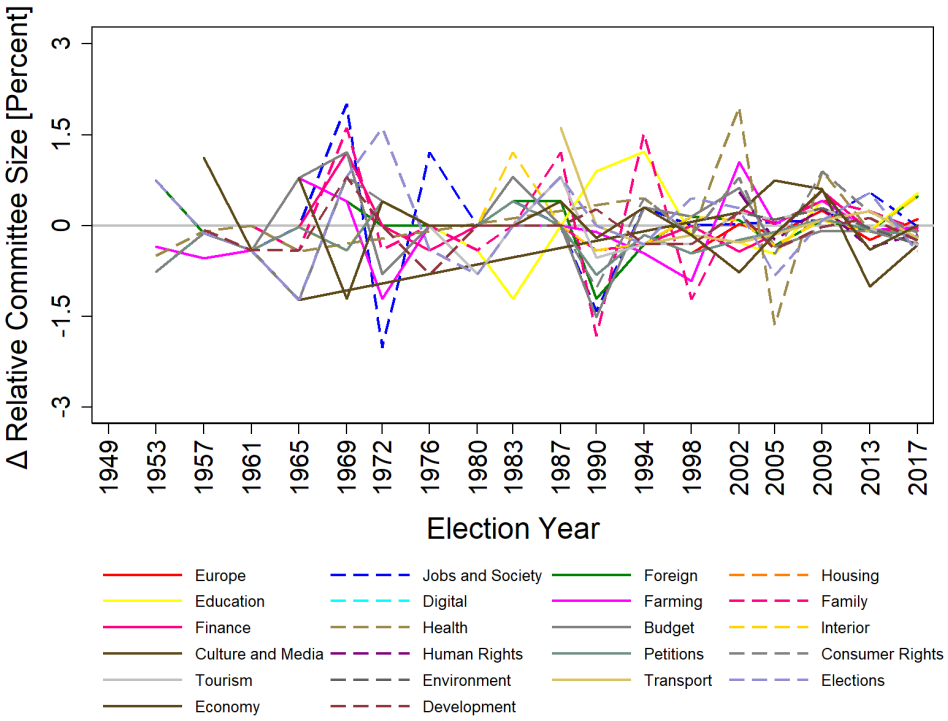
2.A Additional Figures

Figure 2.A.1: Distribution of Speeches by Month and Party



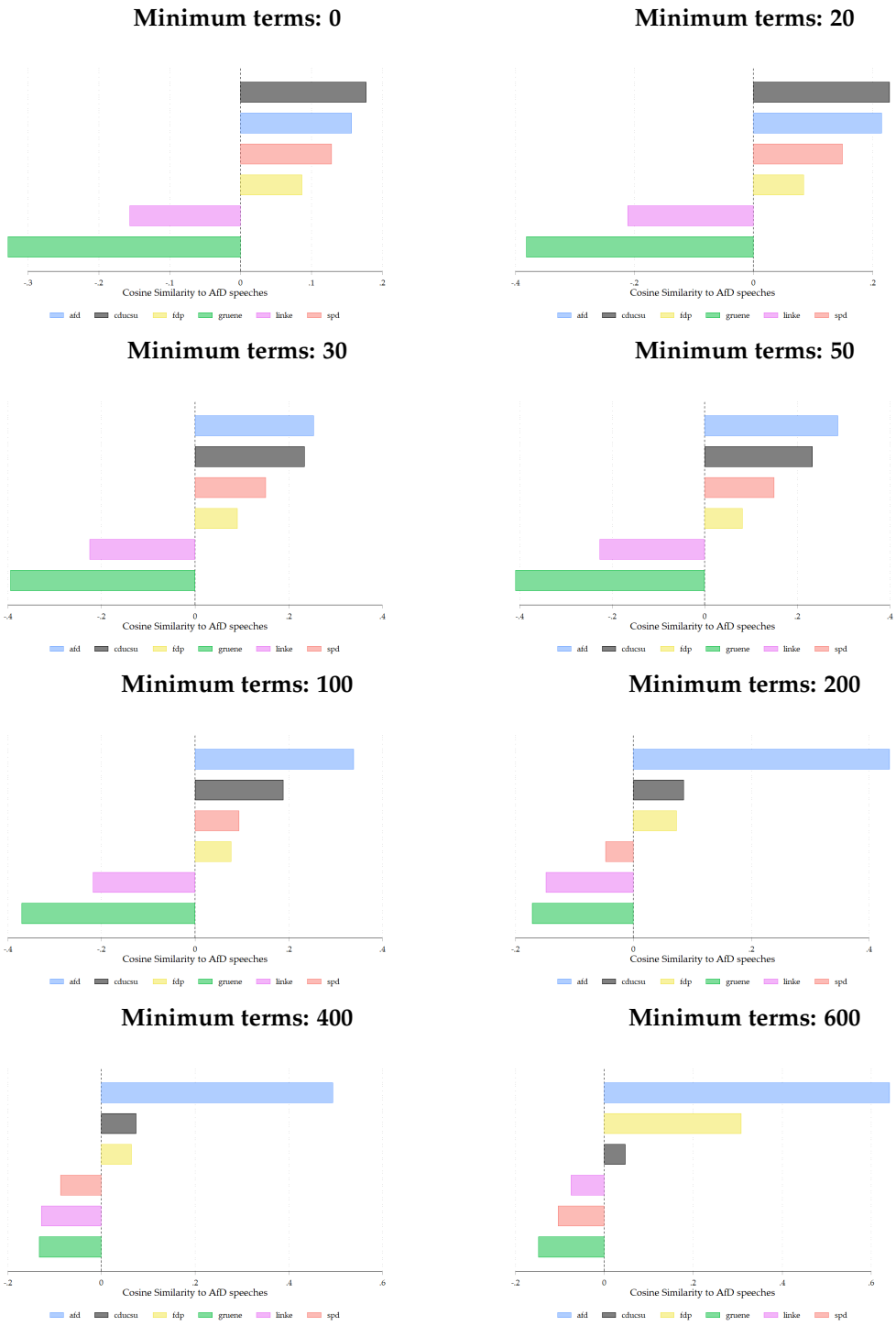
Notes: Figure shows distribution of all speeches in the German Bundestag between October 2013 and December 2019 aggregated by month and party affiliation of the speaker. "Independent" refers to non-affiliated MPs (*fraktionslos*) that do not belong to a parliamentary party group at the time of the speech.

Figure 2.A.2: Changes in Relative Committee Sizes



Notes: Graph shows percentage changes in relative committee sizes over time for all committees in the Bundestag. Data on committees for the 18th Bundestag (2013-2017) and 19th Bundestag (2017-2021) were retrieved from the website of the Bundestag (Deutscher Bundestag 2022a). Data on committees in previous legislative periods were manually extracted from the *Amtliches Handbuch des Deutschen Bundestages* ("Official Manual of the German Bundestag") (Deutscher Bundestag 1954-2017). Sizes are relative to the size of the Bundestag in the respective legislative period. As committees were reshuffled and reorganized several times over time, we manually harmonized committee names based on the committees in the 19th Bundestag.

Figure 2.A.3: AfD Cosine Similarity for different speech lengths



Notes: Graphs show the average standardized cosine similarity to AfD speeches for each party for different minimum terms restrictions on speeches. Sample includes all speeches in the German Bundestag between October 2013 and December 2019.

2.B Additional Tables

Table 2.B.1: Correlation between Similarity Measures

	AfD Cosine Similarity					
	(1)	(2)	(3)	(4)	(5)	(6)
Höcke Cosine Similarity	0.809***	0.716***	0.738***			
	(0.006)	(0.014)	(0.013)			
Populist Dictionary Words				0.212***	0.106***	0.105***
				(0.007)	(0.005)	(0.006)
Topic Controls	-	✓	✓	-	✓	✓
Month FE	-	✓	✓	-	✓	✓
Speaker FE	-	✓	✓	-	✓	✓
Without AfD & FDP	-	-	✓	-	-	✓
Observations	28,998	25,803	22,662	28,998	25,803	22,662

Notes: Table reports coefficients and standard errors from linear regressions. The dependent variable is the standardized average cosine similarity to AfD speeches after pre-processing and tf-idf vectorization. The independent variables are the standardized average cosine similarity to speeches by Björn Höcke after pre-processing and tf-idf vectorization and the standardized number of sentences with words from the German-language populist dictionary by Gründl (2022). The sample comprises all speeches that were held in the German Bundestag between 2013 and 2019 with a minimum length of 100 terms. In columns (3) and (6) we exclude all speeches by members of the AfD, the FDP as well as non-affiliated members. In columns (2), (3), (5) and (6) standard errors are furthermore clustered on the committee times electoral period level. Robust standard errors are reported in parentheses: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 2.B.2: Summary Statistics

	Mean	Std. Dev.	Min.	Max.	Obs.
PANEL A: *Similarity Measures*					
Avg. Cosine Similarity to AfD (min. 100 terms)	0.00	1.00	-2.61	5.68	29,120
Avg. Cosine Similarity to Höcke (min. 100 terms)	0.00	1.00	-2.08	7.79	29,120
Populist Dictionary Score (min. 100 terms)	0.00	1.00	-0.52	17.36	28,998
PANEL B: *Speech Characteristics*					
No. Terms	450.30	370.56	1	4513	39,310
No. Sentences	30.85	25.57	0	387	39,310
PANEL C: *Speaker Characteristics*					
Female	0.34	0.47	0	1	39,117
Age	51.10	10.44	24	81	39,117
East Germany	0.21	0.41	0	1	35,035
Academic Title	0.19	0.40	0	1	39,117
AfD First Vote Share	11.41	5.28	4	37	33,679
Distance to AfD First Vote	13.55	12.43	0	49	33,679
PANEL D: *Committee Shares by Party (19th Bundestag, 2017-21)*					
Share CDU/CSU Members (19th BT)	0.35	0.01	0.33	0.38	27,937
Share SPD Members (19th BT)	0.22	0.01	0.18	0.24	27,937
Share AfD Members (19th BT)	0.13	0.01	0.11	0.14	27,937
Share FDP Members (19th BT)	0.11	0.01	0.10	0.14	27,937
Share Left Members (19th BT)	0.10	0.01	0.07	0.12	27,937
Share Green Members (19th BT)	0.09	0.01	0.07	0.12	27,937
PANEL E: *Committee Shares by Party (18th Bundestag, 2013-17)*					
Share CDU/CSU Members (18th BT)	0.48	0.01	0.44	0.50	28,324
Share SPD Members (18th BT)	0.31	0.01	0.28	0.36	28,324
Share Left Members (18th BT)	0.10	0.01	0.07	0.13	28,324
Share Green Members (18th BT)	0.10	0.01	0.07	0.13	28,324
PANEL F: *Party Shares*					
AfD	0.07	0.25	0.00	1.00	39,310
CDU/CSU	0.30	0.46	0.00	1.00	39,310
SPD	0.20	0.40	0.00	1.00	39,310
Greens	0.21	0.41	0.00	1.00	39,310
Left	0.16	0.37	0.00	1.00	39,310
FDP	0.06	0.23	0.00	1.00	39,310
Independent MPs	0.00	0.00	0.00	0.00	39,310

Table 2.B.3: Committees in the 18th Bundestag (2013-2017)

Committee Name	Total	CDU/CSU	SPD	Linke	Greens
Economic Affairs and Energy	46	22	14	5	5
Labour and Social Affairs	41	20	13	4	4
Budget	41	20	13	4	4
Transport	41	20	13	4	4
Legal Affairs and Consumer Protection	39	19	12	4	4
Finance	37	18	11	4	4
Foreign Affairs	37	18	11	4	4
Health	37	18	11	4	4
Internal Affairs and Community	37	18	11	4	4
Environment, Nature Conservation and Nuclear Safety	36	17	11	4	4
Family Affairs, Senior Citizens, Women and Youth	36	17	11	4	4
Education, Research and Technology Assessment	34	17	11	3	3
European Union Affairs	34	17	11	3	3
Food and Agriculture	34	17	11	3	3
Defense	32	16	10	3	3
Petitions	26	12	8	3	3
Economic Cooperation and Development	21	10	7	2	2
Culture and Media Affairs	18	9	5	2	2
Sports	18	9	5	2	2
Tourism	18	9	5	2	2
Digital Agenda	16	7	5	2	2
Human Rights and Humanitarian Aid	16	7	5	2	2
Elections, Immunity and the Rules of Procedure	14	7	5	1	1

Notes: The table provides the total number of committee members as well as the total number of committee seats allocated to the different parliamentary groups in the 18th Bundestag (2013-2017).

Table 2.B.4: Committees in the 19th Bundestag (2017-2021)

Committee Name	Total	CDU/CSU	SPD	AfD	FDP	Linke	Greens
Economic Affairs	49	17	11	6	5	5	5
Labour and Social Affairs	46	16	10	6	5	5	4
Foreign Affairs	45	16	10	6	5	4	4
Internal Affairs and Community	45	16	10	6	5	4	4
Budget	44	15	10	6	5	4	4
Legal Affairs and Consumer Protection	43	15	9	6	5	4	4
Transport	43	14	10	6	5	4	4
Education, Research and Technology Assessment	42	15	9	5	5	4	4
Finance	41	14	9	5	5	4	4
Health	41	14	9	5	5	4	4
Family Affairs, Senior Citizens, Women and Youth	40	14	9	5	4	4	4
Environment, Nature Conservation, Nuclear Safety	39	13	9	5	4	4	4
European Union Affairs	39	14	8	5	4	4	4
Food and Agriculture	38	13	8	5	4	4	4
Defense	36	12	8	5	4	4	3
Petitions	28	9	6	4	3	3	3
Economic Cooperation and Development	24	9	5	3	3	2	2
Housing, Urban Development, Building, Local Government	24	9	5	3	3	2	2
Digital Agenda	21	7	5	3	2	2	2
Culture and Media Affairs	18	6	4	2	2	2	2
Sports	18	6	4	2	2	2	2
Tourism	18	6	4	2	2	2	2
Human Rights and Humanitarian Aid	17	6	3	2	2	2	2
Elections, Immunity and the Rules of Procedure	14	5	3	2	2	1	1

Notes: The table provides the total number of committee members as well as the total number of committee seats allocated to the different parliamentary groups in the 19th Bundestag (2017-2021).

Table 2.B.5: Selection into Committees

	Share AfD Members	
	(1)	(2)
Female	-0.166**	-0.125*
	(0.066)	(0.068)
Age	-0.005*	-0.006*
	(0.003)	(0.003)
East Germany	0.007	-0.020
	(0.084)	(0.134)
Academic Title	-0.082	-0.106
	(0.082)	(0.084)
AfD Vote Share		0.004
		(0.009)
Distance to AfD		0.002
		(0.003)
Constant	13.217***	13.171***
	(0.156)	(0.186)
Observations	509	486

Notes: Table reports coefficients and standard errors from linear regressions. The sample comprises all members represented in the 19th German Bundestag that were full member of at least one parliamentary committee. Members affiliated with the AfD are excluded from the sample. The dependent variable measures the average of the share of AfD members (in percent) across all committees of which a politician is a full member. Age refers to the age of a politician in years as of the opening of the 19th German Bundestag (October 24, 2017). East Germany is a dummy variable equal to 1 if the MP was elected in a state of former East Germany. Academic Title is a dummy variable equal to 1 if the MP uses a doctoral or professorial title in her name. AfD Vote Share measures the constituency vote (first vote) share of the AfD (in percent) in an MP's electoral district in the 2017 federal election. Distance to AfD measures the absolute distance of the MP's own constituency vote (first vote) share to the AfD vote share (in percentage points) in the 2017 federal election. Robust standard errors are reported in parentheses: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 2.B.6: General Accommodation Effects

	Cosine Similarity to speeches by ...					
	(1) AfD	(2) FDP	(3) CDU/CSU	(4) SPD	(5) Greens	(6) Left
Share AfD × Post	3.356*					
	(1.932)					
Share FDP × Post		3.411				
		(2.511)				
Share CDU/CSU × Post			0.079			
			(0.773)			
Share SPD × Post				-0.288		
				(1.070)		
Share Greens × Post					1.268	
					(1.305)	
Share Left × Post						-0.287
						(1.047)
Topic Controls	✓	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓	✓
Speaker FE	✓	✓	✓	✓	✓	✓
Observations	17,383	17,383	14,688	17,322	17,689	18,285

Notes: Table reports coefficients and standard errors from linear regressions as laid out in Equation 2.4. The independent variable of interest is the interaction between the (average) share of respective party members of all committees in which a politician is a full member and an indicator whether the speech was recorded in the 19th German Bundestag (2017-2021). The dependent is the standardized average cosine similarity to speeches by members of the respective party after pre-processing and tf-idf vectorization. Topic controls are derived from a 20-topic LDA model. The sample comprises plenary speeches by members of the German Bundestag held between October 2013 and December 2019 with a minimum length of 100 terms from parties that were represented throughout the whole period (CDU/CSU, SPD, The Left, and Alliance90/The Greens), excluding members of the respective party. Standard errors clustered at the committee times electoral period level are reported in parentheses: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

2.C Technical Details

Our data management is mainly done in `python` with some packages used in `R` if provided like that from the respective authors. To manage our workflow and allow for smooth integration of code from different languages we use `pytask` (Raabe 2020).

2.C.1 Pre-processing

As a first step of the pre-processing, we fix some regularly occurring errors in the raw text data where words were not separated by blanks. To fix these we use `language-tool-python`[42], the python wrapper of `LanguageTool`, an open-source grammar tool and spell checker. Next, we remove punctuation including German-specific and context-specific characters. We then remove stopwords and lemmatize the tokens. As the `nltk` database for German stopwords is very limited we use a more comprehensive set from https://github.com/solariz/german_stopwords (last accessed on 03/15/2023). For the lemmatization we use the `Hanover Tagger` (Wartena 2019), a lemmatizer and POS tagger specifically designed for the German language. We refrain from stemming as it can lead to undesired oversimplification. Especially when thinking about inclusive language only using male or using both versions of a noun might matter. Also, Gründl (2022) points out, that stemming in a German context can lead to words becoming indistinguishable (e.g., *Bürger* (citizen), *bürgen* (to vouch), and *Burg* (castle).)

2.C.2 Similarity Measures

To obtain cosine similarity measures, we use the `TfidfVectorizer` package from the `scikit learn` (Pedregosa et al. 2011) module to create the tf-idf-matrix. Further, we `transform` the matrix to obtain an array for each speech. Data frame and matrix manipulations to calculate the averaged similarity scores to each party and Höcke are done with `pandas` (McKinney 2010) and `numpy` (Harris et al. 2020).

For the populist dictionary scores we use the code provided by Gründl (2022) and his `R` packages `popdictR` (Gründl 2020b), `multidictR` (Gründl 2020a) and `regexhelpeR` (Gründl 2020c). It processes the raw text on a sentence level and uses regular expression to identify populist words or phrases. It then counts the number of sentences containing populist content. A list of the dictionary entries found in the speeches can be found in Table 2.C.1.

[42] See https://pypi.org/project/language-tool-python/ (last accessed on 03/15/2023).

2.C.3 Topic Modeling

We use `gensim` (Rehurek and Sojka 2011) and its LDA model for the LDA-Topic modeling. We prune at a 1% level. The derived topics and associated Top 20 words translated to English and in German can be found in Table 2.C.2.

Table 2.C.1: Populist dictionary entries following Gründl (2022)

Anti-elitism
so-called/sogenannte (4,696) — to finance/finanzieren (2,080) — admit/zugeben (631) — bureaucrat/bürokrat (513) — to be ashamed/schämen (467) — to deceive/täuschen (465) — audacious/dreist (183) — corrupt/korrupt (155) — to manipulate/manipulieren (141) — circles/kreisen (140) — deception/täuschung (119) — mendacious/verlogen (74) — aloof/abgehoben (71) — to mock/verhöhnen (68) — erroneously/fälschlicherweise (66) — to lecture/belehren (65) — to fiddle/tricksen (63) — dishonest/unehrlich (63) — outrageous/unverschämt (59) — to patronize/bevormunden (58) — unworldly/weltfremd (47) — far from reality/realitätsfern (47) — greedy/gierig (42) — propaganda/propaganda (42) — arrogant/arrogant (41) — disaster/desaster (39) — ludicrous/aberwitzig (38) — technocrat/technokrat (37) — to presume to do/sich anmaßen (37) — centralist/zentralisten (35) — centralistic/zentralistisch (35) — elite/elite (35) — presumptuous/anmaßend (33) — capitalist/kapitalist (31) — insanity/irrsinn (29) — encrusted/verkrustet (24) — indoctrination —instruction/belehrung (23) — lack of contact with reality/realitätsferne (23) — complacent/selbstgefällig (21) — ludicrous/wahnwitzig (21) — from above/von oben herab (19) — quixotic/lebensfremd (18) — banker/bänker (17) — dilettante/dilettantisch (17) — mafia/mafia (16) — absurdity/irrwitz (16) — speculator/spekulant (15) — out of touch with reality/realitätsfremd (14) — mob/pöbel (14) — complacent/selbstzufrieden (13) — arrogant/überheblich (12) — bosses/bosse (11) — fiddle/kungel (11) — to dare/erdreisten (9) — pedantic/oberlehrerhaft (7) — head teacher (in the meaning of a smart aleck)/oberlehrer (7) — at the expense of the Germans/zu lasten der deutschen (7) — opportunists/opportunisten (7) — to corrupt/korrumpieren (6) — remote from the people/bürgerfern (5) — disgrace/schande (4) — spineless/rückgratlos (3) — failing/versagend (3) — unprincipled/prinzipienlos (3) — haughty/hochmütig (2) — insatiable/nimmersatt (2) — remote from everyday life/lebensfern (2) — traitor to the nation/the people/volksverräter (2) — bigwig/bonze (2) — haggling/geschacher (1) — inane/hirnverbrannt (1) — pseudo-parties/pseudo-parteien (1) — government failure/staatsversagen (1) — stuck-up/hochnäsig (1) — establishment/establishment (1) — jet set/schickeria (1)

Table 2.C.1: Populist dictionary entries (continued)

Sovereignty
dictate/diktat (87) — undemocratic/undemokratisch (82) — anti-democratic/antidemokratisch (49) — allowed to say/sagen dürfen (35) — the citizens wish —want —demand/bürger fordern —möchten —mögen —verlangen —beanspruchen —wünschen (23)[43] — majority/mehrheit (10) — high-handed/selbstherrlich (9) — plebiscitary/plebiszitär (8) — the people demand —want —wish —/das volk will —fordert —möchte —mag —verlangt —beansprucht —wünscht (5) — for the —our people/für das —unser volk (2) — power-hungry/machtversessen (2) — party dictatorship/parteiendiktatur (1) — plebiscite/volksentscheid (1)
People-centrism
tradition/tradition (150) — steadfast/standhaft (28) — average german/durchschnittlicher deutscher (1) — our citizens/unsere bürger (1) — working germans/arbeitende deutsche (1)

Notes: All entries translated to English by the authors, original German version after the "/". The frequency of appearance is displayed in brackets behind the phrase. For better readability, the phrases were changed to their infinitives or non-declinated forms. The regex search patterns cover all different cases of declinations and conjugations for both singular and plural. An extensive list of regex expressions can be found in the online appendix of Gründl (2022). The categories are based on the populist ideology classification from Gründl (2022).

[43] To avoid confusion and for better readability, four different versions with different syntax from the dictionary were combined into one.

Table 2.C.2: LDA Topic modeling - Top 20 words for each topic

Topic 1
european/europäisch — europe/europa — eu — china — union — russia/russland — together/gemeinsam — national — ukraine — interest/interesse — france/frankreich — cooperation/zusammenarbeit — russian/russisch — african/afrikanisch — level/ebene — partner — germany/deutschland — greece/griechenland — great britain/großbritannien — member state/mitgliedstaat
Topic 2
topic/thema — area/bereich — address/ansprechen — minister — point/punkt — recognition/erkenntnis — be interested in/interessieren — request/nachfrage — address/angehen — discuss/diskutieren — hundred thousand/hunderttausend — responsibility/zuständigkeit — evaluate/bewerten — warn/warnen — extension/ausweitung — clock/uhr — discuss/besprechen — affect/betreffen — keyword/stichwort — to be entitled to sth./zustehen
Topic 3
climate protection/klimaschutz — co — energy/energie — climate change/klimawandel — global — goal/ziel — ecological/ökologisch — renewable/erneuerbar — expansion/ausbau — reach/erreichen — energy revolution/energiewende — green/grün — globally/weltweit — amendment/novelle — percent/prozent — science/wissenschaft — paris — net/netz — international — measure/maßnahme
Topic 4
colleague/kollegin — dear/liebe — year/jahr — large/groß — accomplish/schaffen — important/wichtig — strong/stark — considerable/deutlich — together/gemeinsam — right/richtig — provide/stellen —cordial/herzlich — to care/sorgen — example/beispiel — goal/ziel — measure/maßnahme — good/gut — country/land — show/zeigen — support/unterstützen
Topic 5
company/unternehmen — investment/investition — economy/wirtschaft — germany/deutschland — to invest/investieren —social/sozial — development/entwicklung — economic/wirtschaftlich — employment/arbeitsplatz — region — future/zukunft — infrastructure/infrastruktur — to function/funktionieren — market/markt — innovation — competition/wettbewerb — industry/industrie — business/betrieb — percent/prozent — create/schaffen

Table 2.C.2: LDA Topic modeling - Top 20 words for each topic (continued)

Topic 6
security/sicherheit — firstly/erstens — secondly/zweitens — date/datum — net/netz — thirdly/drittens — it — police/polizei — control/kontrolle — pact/pakt — perpetrator/täter — communication/kommunikation — to function/funktionieren — federal office/bundesamt — dependent/abhängig — efficient/effizient — data protection/datenschutz — withdraw/entziehen — equipment/ausstattung — judiciary/justiz
Topic 7
soldier/soldat — german armed forces/bundeswehr — mission/einsatz — female soldiers/soldatinnen — turkey/türkei — peace/frieden — armed/bewaffnet — international — nato — security/sicherheit — nation — region — conflict/konflikt — war/krieg — military/militärisch — iran — foreign minister/außenminister — humanitarian/humanitär — united/vereinter — un
Topic 8
woman/frau — work/arbeit — nursing/pflege — social/sozial — pension/rente — parents/eltern — payment/leistung — income/einkommen — wage/lohn — labor market/arbeitsmarkt — employed/beschäftigt — employee/arbeitnehmer — age/alter — statutory/gesetzlich — man/mann — welfare state/sozialstaat — percent/prozent — basic income/grundsicherung — retiree/rentner — mother/mutter
Topic 9
regulation/regelung — procedure/verfahren — case/fall — rule/regel — affected/betroffen — legal/rechtlich — authority/behörde — possibility/möglichkeit — present/vorliegend — decision/entscheidung — agriculture/landwirtschaft — interest/interesse — protection/schutz — high/hoch — person — so-called/sogenannter — public/öffentlich — legal/gesetzlich — basically/grundsätzlich — substantial/erheblich
Topic 10
law/gesetz — draft law/gesetzentwurf — hearing/anhörung — federal council/bundesrat — abolition/abschaffung — expert/experte — brandenburg — serious/seriös — to consult/beraten — state government/landesregierung — consultation/beratung — infer to from/entnehmen — agree with/zustimmen — consent/zustimmung — contain/enthalten — boss/chef — to pass/verabschieden — improvement/verbesserung — parliamentary/parlamentarisch — to introduce/einbringen

Table 2.C.2: LDA Topic modeling - Top 20 words for each topic (continued)

Topic 11
euro — billion/milliarde — year/jahr — percent/prozent — million — money/geld — country/land — budget/haushalt — federation/bund — municipality/kommune — funds/mittel — to pay/zahlen — costs/kosten — additionally/zusätzlich — minister(f.)/ministerin — tax/steuer — to increase/erhöhen — disposal/verfügung — city/stadt — research/forschung
Topic 12
human/mensch — life/leben — country/land — human right/menschenrecht — refugee/flüchtling — aid/hilfe — to help/helfen — poor/arm — million — perspective/perspektive — group/gruppe — affected/betroffen — poverty/armut — place/ort — peaceful/friedlich — situation — safe/sicher — city/stadt — escape/flucht — distress/not
Topic 13
question/frage — to believe/glauben — problem — to know/wissen — to speak/sprechen — to talk/reden — to lead/führen — to put/stellen — president/präsident — correct/richtig — year/jahr — debate/debatte — to mean/heißen — to get/bekommen — point/punkt — wrong/falsch — already/schon — big/groß — time/zeit — house/haus
Topic 14
child/kind — family/familie — education/bildung — school/schule — training/weiterbildung — bafög — north rhine/nordrhein — westphalia/westfalen — university/hochschule — disability/behinderung — to learn/lernen — specialist/fachkraft — performance/leistung — child benefit/kindergeld — quality/qualität — minister (f.)/ministerin — chance — qualification/qualifikation —daycare/kita — trained/ausgebildet
Topic 15
afd — cdu — csu — party/partei — spd — tax payer/steuerzahler — fdp — seehofer — to govern/regieren — credit/kredit — bank — to safe/retten — to sign/unterscheiden — bavaria/bayern — election campaign/wahlkampf — to defend/verteidigen — tax money/steuergeld — elections/wahlen — capital/kapital — interest/zins

Table 2.C.2: LDA Topic modeling - Top 20 words for each topic (continued)

Topic 16
germany/deutschland — german/deutsch — lady/dame — citizen/bürger — country/land — state/staat — political/politisch — policy/politik — president/präsident — world/welt — democracy/demokratie — digital — victim/opfer — freedom/freiheit — value/wert — right/recht — citizens (f.)/bürgerinnen — to show/zeigen — fight/kampf — fear/angst
Topic 17
request/antrag — fdp — german parliament/bundestag — parliamentary group/fraktion — green/grün — colleague (f.)/kollegin — spd — dear/liebe — parliament/parlament — leftist/linker — proposal/vorschlag — public/öffentlich — committee/ausschuß — to agree/zustimmen — parliamentary/parlamentarisch — votes/stimmen — debate/debatte — commission/kommission — president (f.)/präsidentin — delegated/abgeordnet
Topic 18
federal government/bundesregierung — finally/endlich — government/regierung — leftists/linke — greens/grüne — coalition/koalition — submit/vorlegen — change/änderung — to promise/versprechen — urgent/dringend — real/echt — to change/ändern — draft/entwurf — massive/massiv — to suffice/reichen — to wait/warten — to fail/scheitern — union — plan/vorhaben — reform
Topic 19
usa — contract/vertrag — negotiation/verhandlung — agreement/abkommen — us — to unite/vereinigen — evening/abend — american/amerikanisch — young people/jugendliche — relevant — recognisable/erkennbar — international — american/amerikaner — america/amerika — position/stellung — world/welt — trade/handel — to negotiate/verhandeln — state/staat — partner
Topic 20
report/bericht — supply/versorgung — information — complex/komplex — consensus/konsens — request/anfrage — restriction/einschränkung — template/vorlage — ensured/versichert — happy/glücklich — clarification/aufklärung — left-wing fraction/linksfraktion — answered/beantwortet — to inform/informieren — access/zugang — patient — digitization/digitalisierung — fund/kasse — ministry/ministerium — health insurance/krankenkasse

CHAPTER 3

War, Grievances, and Nationalism: Evidence from South Tyrol

3.1 Introduction

The history of nationalism is one of war, destruction, and violence. While nationalism can be the cause of war, it is also war that can drive nationalism (Hutchinson 2017; Wimmer 2013). On a macro-level, wars can make states consolidate power and build up state capacity; a side-effect of this process can be the developing of a sense of national community (Tilly 1975, 1993, 1994). On a micro-level, individual experiences of war can cause grievances and deep-seated feelings of nationalist resentment (Hutchinson 2017). As a result, individuals might intensify their national identity. For example, Tilly (1994, p. 141) noted how "anti-French, anti-Polish, or anti-Russian feeling[s] reinforced the desirability of becoming very German". Leaders from Adolf Hitler to Vladimir Putin have sought to exploit grievances and nationalist resentment to mobilize masses for their extreme ideologies. In order for such mobilization to be successful, leaders require a readiness of individuals to follow nationalist policy. Can war grievances make people more willing to comply with nationalist mobilization?

This paper studies the effect of personal war grievances on individuals' behavior in an episode of nationalist mass mobilization. We focus on the German-speaking population in the Italian region of South Tyrol, around half of whom emigrated to Nazi Germany in the early years of World War II (WWII). Two decades earlier, South Tyroleans were fighting as Austro-Hungarian soldiers in World War I (WWI) against, among other opponents, Italy. After the war, South Tyrol was annexed by Italy and its population subjected to an aggressive Italianization campaign. Most South Tyroleans opposed these attempts at forced assimilation; yet, the degree to which they resented Italy may have also depended on personal experiences and grievances directed at Italy.

We show that grievances resulting from fighting Italy in WWI had long-lasting effects on affected South Tyroleans. To identify enemy-specific war grievances, we exploit exogenous variation in front experience. We use individual-level data to identify families who experienced a casualty on the Italian front and compare them to families

who experienced a similar casualty on another front, e.g., on the Eastern front against Russia. We then ask whether families whose grievance is directed at Italy developed stronger anti-Italian resentment and, thereby, were more willing to comply with pro-German mass mobilization.

When Adolf Hitler came to power in Germany, he pursued a nationalist policy that sought to unify all ethnic Germans in one homogeneous nation-state (Hobsbawm 1992).[1] However, the strategic importance of an alliance with Benito Mussolini's Fascist Italy made the territorial annexation of South Tyrol impossible. As part of the so-called *South Tyrol Option Agreement* of 1939, Hitler and Mussolini instead decreed that all South Tyroleans would be forced to choose between being either German or Italian citizens. If they chose German citizenship, they had to emigrate to Germany; if they chose Italian citizenship, they were able to remain in their homeland but had to fully assimilate into Italian culture. While most individuals would have preferred to stay in South Tyrol and keep their culture and language, this was no longer possible after the Option Agreement (Steininger 2003). In essence, the Option Agreement forced South Tyroleans to make a momentous choice: either emigrate and avoid assimilation, or stay and be subjected to total Italianization.

We study whether Italy-specific war grievances mobilize individuals into emigrating to Germany. In order to answer this question, we combine novel individual-level data from two main sources. First, we collect war records on all casualties of South Tyrolean soldiers during WWI, from which we can infer their place of origin, which army unit they served in, and, crucially, *where* they served during the war.[2] Second, we match these data to individual-level emigration requests from the ensuing mass emigration after the South Tyrol Option Agreement. We hand-collect and digitize a random sample of South Tyrolean households' emigration requests, which allows us to observe who applied for emigration and who eventually migrated to Germany. We also obtain detailed information on individuals' socio-economic status, family status, and many other biographic data. We match these emigration records to the casualty data by developing a cascading linking algorithm.

Insights from social psychology suggest that negative exposure to an out-group may affect in-group cohesion and lead to negative attitudes towards the out-group (e.g., Tajfel and Turner 1979; Choi and Bowles 2007; Haidt 2013; Henrich 2020). War experiences, in particular, have been shown to affect individuals' behavior (e.g., Bauer et al. 2016). Building on these findings, we hypothesize that war grievances, depending on

[1] In his remarks on Hitler's nationalism, Hobsbawm (1992, p. 133) mentions South Tyrol as a case in point: "Subsequently Adolf Hitler, who was in this respect a logical Wilsonian nationalist, arranged to transfer Germans not living on the territory of the fatherland, such as those of Italian South Tyrol, to Germany itself, as he also arranged for the permanent elimination of the Jews."

[2] We define a casualty in a military sense as a soldier who is unavailable for service, i.e., by being wounded, captured, or killed during war.

whom they are directed at, may differentially impact individuals' identity and their attitudes towards the out-group. To investigate this hypothesis, we leverage the front on which a soldier fought as an exogenous shock to the enemy-specific component of war grievances. For causal identification, we exploit Italy's unexpected declaration of war on Austria-Hungary in 1915 as a natural experiment. While nearly all South Tyrolean soldiers initially fought on the Eastern front, they were suddenly redeployed to fight on the new Italian front. This historical event caused exogenous variation in the front to which soldiers were deployed, even though they continued to fight in the same army units under the same commanders and were not differentially selected into fighting on a different front.

Our main identifying assumption is that a casualty happening on the Italian front is not systematically related to unobserved factors affecting individuals' probability to emigrate to Germany two decades later. A prime concern is that soldiers might have been non-randomly selected into fighting on specific fronts. For example, pre-existing anti-Italian sentiment might make soldiers self-select into fighting on the Italian front and, simultaneously, affect their relatives' willingness to emigrate to Germany. We argue that this is unlikely, since exposure to the Italian front was determined by an exogenous factor: the sudden redeployment of South Tyrolean regiments after Italy's declaration of war. Nevertheless, we empirically test for (self-)selection into front exposure in a series of balancing tests, showing that soldiers who experienced a casualty on the Italian front were not systematically different from soldiers who experienced a casualty on another front. We also show that front exposure has no meaningful impacts on any other socio-economic outcomes related to the emigration decision. Furthermore, we include army unit as well as municipality fixed effects in our baseline specification, thereby holding constant potential channels of selection relating to a soldier's place of origin or army unit.

We find that personal war grievances mobilize individuals to comply with nationalist policy. Our baseline specification estimates that households holding Italy-specific war grievances were about 12 percentage points more likely to emigrate to Germany. Given an average emigration probability of 57%, this constitutes a sizable effect of 20 percent relative to the mean. This effect is identified from within the set of households who experienced any form of war grievance. As such, our empirical strategy allows us to hold fixed the effect of holding any war grievance, which might in many ways influence an individual's socio-economic outcomes (Dupraz and Ferrara 2023). Thus, we are able to isolate the effect of a war grievance being specifically directed against Italy from the – likely endogenous – effect of holding any grievance from WWI on a household's emigration decision.

We provide robustness checks relating to two main kinds of concerns. The first con-

cern is that unobservable local factors might bias the results. For example, individuals from certain municipalities might be more likely to fight on the Italian front and, at the same time, be more likely to emigrate. We control for a number of geographic fixed effects, both relating to the individual observed in the emigration decision, as well as to their father observed in WWI. Across specifications, the estimated effect of Italy-specific war grievances on emigration behavior remains stable. Second, one might be concerned that our results are driven by certain subgroups. We test this by dropping specific groups of individuals from our main analysis, e.g., individuals from municipalities with relatively higher pre-WWI Italian population or those bordering Austria or Italy. Again, we do not find that this affects our estimated coefficient of interest.

Our paper makes contributions to a broad array of research. First, our findings relate to a literature on the effects of grievances and past victimization. A consistent finding in this literature is that experiences of violence and war can make individuals more pro-social, but also more biased towards their in-group (Bauer et al. 2016; Henrich 2020, pp. 328–343; Walden and Zhukov 2020). War experiences and memories thereof can decrease trust (Conzo and Salustri 2015; Vlachos 2022), affect consumption behavior (Fouka and Voth 2022), and increase support for right-wing and militaristic politics (Grossman, Manekin, and Miodownik 2015; Kibris and Cesur 2022). War grievances, in particular, can mobilize individuals to join military groups and support violent movements (Dell and Querubin 2018; Marchais et al. 2022). To identify the effect of grievances, many papers compare individuals with varying degrees of war experience. Our study adds to this literature in two ways: first, by holding constant the severity of the grievance, we isolate who the grievance is directed at, and, second, we relate such enemy-directed grievances to individual behavior in times of nationalist mobilization. In this regard, our study is also related to work showing that past victimization can negatively affect the attitudes of victims or their descendants (Bauer et al. 2023; Dinas, Fouka, and Schläpfer 2021a,b; Fouka and Voth 2022).

Second, we contribute to a growing literature on the origins of nationalism and national identity (Rohner and Zhuravskaya 2023). For example, recent research has focused on the role of nation-building policies (Kersting and Wolf 2021), shared collective experiences such as sporting victories (Depetris-Chauvin, Durante, and Campante 2020), foreign occupation (Dehdari and Gehring 2022), and education (Clots–Figueras and Masella 2013; Cantoni et al. 2017).[3] A common feature of these papers is that they assess policies or events that *collectively* affect individuals' identity, i.e., because they live in the same geographic area or are part of the same age cohort. Our paper contributes to this literature by isolating an *individual-level* shock to nationalist sentiment. This allows us to compare individuals who are otherwise similar, e.g.,

[3]These papers complement a theoretical literature in economics on nation-building (e.g., Alesina and Reich 2015; Alesina, Reich, and Riboni 2020; Gennaioli and Voth 2015).

from the same locality or of the same age, but only differ in the shock to their negative sentiment against the former Italian war opponent.

Third, we add to a literature investigating the consequences of wars and military service on social and political movements. Some recent papers show effects of war experience – and an intergenerational transmission thereof – on individual-level behavior, such as tax compliance (Galletta and Giommoni 2023) or service in the army (Campante and Yanagizawa-Drott 2016). Another more common approach in this literature is to use local casualty shares to identify the effect of war exposure on subsequent socio-economic or political outcomes. More severe local war exposure has been linked to changing labor market conditions (e.g., Acemoglu, Autor, and Lyle 2004; Ferrara 2022), changes in the marriage market (Abramitzky, Delavande, and Vasconcelos 2011), increased socialist and fascist support in Italy (Acemoglu et al. 2022), and increased support for the Nazi party in Germany (De Juan et al. 2022; Koenig 2023). Relatedly, Ferrara and Fishback (2022) show that a stronger local casualty shock during WWI led to more anti-German attitudes in the United States, to which ethnic Germans responded by moving away. We deviate from these papers in two substantial ways. First, we use an individual-level shock and relate it to individual-level behavior, thereby going further than most existing work using locally aggregated casualty data. Second, our identifying assumption places no exogeneity restriction on *whether* but only on *where* a soldier suffered a casualty.

Fourth, we build on a theoretical literature analyzing the interplay between identity and behavior. In their seminal paper, Akerlof and Kranton (2000) introduce a framework to explain how identity impacts economic behavior. They consider identity as static and as determining individuals' preferences and actions. Shayo (2009, 2020) models identity as endogenously determined: an individual's identity can change when it becomes relatively more costly to identify with a group.[4] Since identity is never directly observable, the empirical literature on identity choices focuses on revealed-preference measures. For example, Atkin, Colson-Sihra, and Shayo (2021) expand on the framework proposed by Shayo (2009, 2020) and investigate identity-revealing food consumption choices. Other work considers self-reported (racial or ethnic) identity measures (Dahis, Nix, and Qian 2020; Jia and Persson 2021). These concepts have also been applied to investigating the circumstances under which minorities do or do not assimilate into the majority culture (Fouka 2019, 2020; Fouka, Mazumder, and Tabellini 2022). In the case of the South Tyrol Option Agreement, we measure a highly consequential revealed identity choice: whether or not an individual emigrates in order to keep their cultural and national identity. As such, this paper can

[4]Other recent advances include Besley and Persson (2018, 2021) and Bonomi, Gennaioli, and Tabellini (2021). Similarly to Shayo (2009, 2020), they consider multiple identity dimensions and explain how individuals choose one of them depending on the political climate.

also be seen as examining an extreme case of assimilation avoidance.

Last, this paper makes thematic contributions to two further research areas: on the one hand, it contributes to the literature on forced relocation and selective emigration.[5] Many papers in this field focus on the long-run consequences of forced relocation on socio-economic outcomes (e.g., Becker et al. 2020; Sarvimäki, Uusitalo, and Jäntti 2022). Recently, more attention has been paid to factors influencing individual emigration behavior (Becker and Ferrara 2019). These papers have emphasized the role of networks (Becker et al. 2022), the experience or threat of violence in the location of origin (Buggle et al. 2023; Clemens 2021), or cultural attitudes such as individualism (Beck Knudsen 2022). We contribute by investigating how personal grievances can drive individuals into emigration in a setting lying on "the spectrum between perfectly voluntary migration and forced migration" (Becker and Ferrara 2019, p. 14).[6] On the other hand, we relate to a long-standing literature examining collective action and political mobilization (Olson 1965; Tilly 1978). Many studies focus on the spread of political movements (e.g., García-Jimeno, Iglesias, and Yildirim 2022) or on factors driving selection into protest movements (Cantoni et al. 2019, 2022). Rather than focusing on the mobilization of self-selected politically active agents, we study the behavior of ordinary people in times of (nationalist) mass mobilization.

The remainder of this paper is structured as follows: Section 3.2 introduces the historical setting and Section 3.3 describes the data. Section 3.4 discusses our empirical strategy and provides evidence for the exogeneity of the proposed war grievance shock. In Section 3.5 we discuss our findings, while Section 3.6 concludes.

3.2 Historical Background

In the first half of the 20th century, nationalism reorganized the European political landscape, involving the formation of new nation-states, the rise of nationalist and fascist mass movements, as well as the relocation of ethnic communities. Historians and social scientists have pointed to the fundamental role of war at the root of this development (Roshwald 2001; Wimmer 2013; Hutchinson 2017). In particular, World War I has been called "*the* seminal catastrophe of the century" (Kennan 1981), laying the foundations for the rise of extreme nationalism culminating in World War II. We focus on an instructive – albeit relatively unknown – episode of nationalist mobilization: the mass emigration of German-speakers from South Tyrol to Nazi Germany.

[5] For a comprehensive review of this literature see Becker and Ferrara (2019) and Becker (2022).

[6] Related to our study, personal army experience may play a role in emigration: Salem and Seck (2023) show that Africans who were soldiers in the Colonial French army and were deployed in France were more likely to relocate to France after independence.

3.2.1 From World War I to Italianization

Until World War I, modern-day South Tyrol had for centuries formed part of the Habsburg-ruled Austro-Hungarian Empire.[7] While the Habsburg Empire was a distinctively multinational entity, many of its regions were locally ethnically homogeneous. The population of South Tyrol was largely German-speaking: in the 1910 census, 89.0% out of 251,451 inhabitants identified as German-speakers, compared to only 2.9% Italian-speakers.[8]

World War I On July 28, 1914, World War I began with the Austro-Hungarian declaration of war against the Kingdom of Serbia and, like millions of other men across Europe, South Tyroleans were called to arms. As part of the Central Powers, Austria-Hungary initially fought on two fronts, against Serbia in the Balkans and against Russia on the Eastern Front. A major turning point marked the declaration of war by the Kingdom of Italy on Austria-Hungary on May 23, 1915. As Italy had been part of the Triple Alliance with Germany and Austria-Hungary and initially remained neutral in 1914, this came as an unexpected shock, forcing the opening of a third front in the South. This "breach of faith, the like of which history has never seen", as proclaimed by Emperor Franz Joseph, was met with surprise and outrage by many citizens of Austria-Hungary, in many cases leading to deep-seated hatred against Italy (Di Michele 2020).[9]

To counter the imminent Italian invasion, the Austro-Hungarian military command hastily redeployed multiple army units from the fighting in the East to the Southern border, among them all of the main regiments with South Tyrolean soldiers (Glaise von Horstenau 1932). These South Tyrolean soldiers now suddenly found themselves fighting against the Italian Army on a newly emerging front spanning from the Adriatic coast to the Alps in South Tyrol.[10] Hostilities ended more than three years later in light of the increasing disintegration of the Austro-Hungarian Empire. The armistice was signed on November 3, 1918, and Tyrol was subsequently occupied by Italy.

[7] When speaking of South Tyrol, we refer to the territory of the modern-day *Autonomous Province of Bolzano - South Tyrol* in Italy. Figure 3.1a depicts the location of South Tyrol (darker red shape) before WWI within the historical boundaries of the County of Tyrol, a crown land of the Austro-Hungarian Empire.

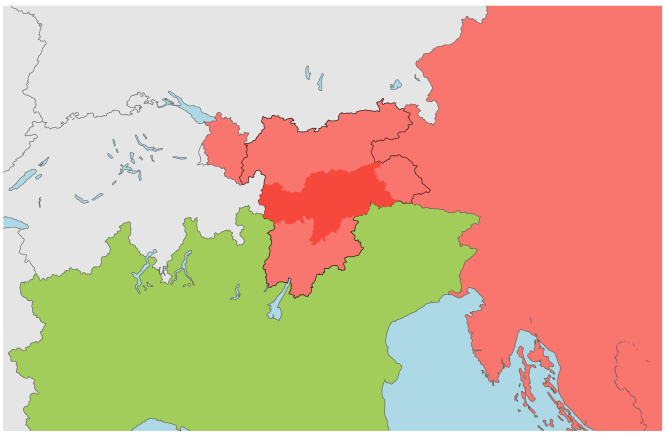
[8] The remaining population included 3.8% Ladins, an ethnic and linguistic minority in northern Italy, and 4.3% speakers of other languages. (Autonomous Province of South Tyrol Provincial Statistics Institute 2021). In the Habsburg Empire, affiliation to an ethnic group or nation was determined by language use (Stergar and Scheer 2018).

[9] An example of the mood at the time can be seen in Panel (A) of Appendix Figure 3.A.2, which shows an anti-Italian propaganda postcard from WWI.

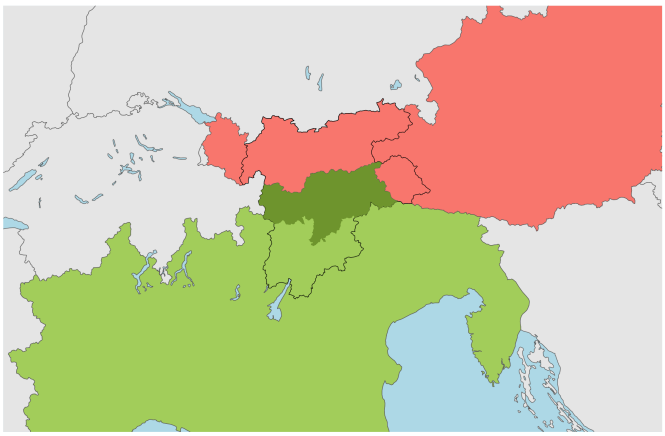
[10] Figure 3.A.1 in the Appendix shows the course of the frontline that soon evolved into a stalemate with a fierce trench and tunnel warfare involving heavy losses, in particular in the high-altitude Alpine sectors that became known as the *"White War"* (Thompson 2008).

Figure 3.1: Political Maps of South Tyrol 1914-1939

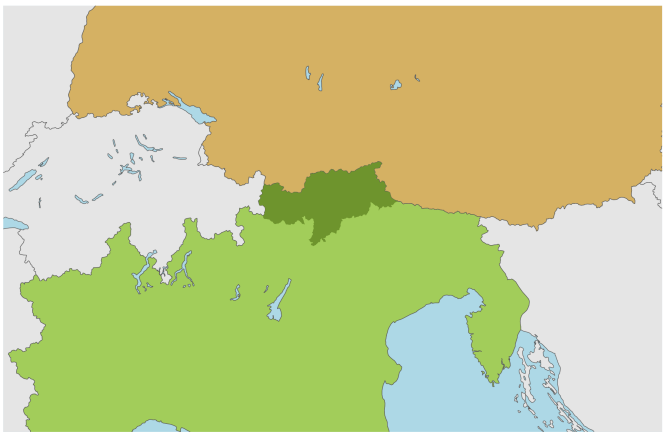
(a) 1914: Before World War I



(b) 1920: After World War I



(c) 1939: South Tyrol Option Agreement



Notes: Maps show historical political borders as of 1914, 1920, and 1939. Italy is shown in green, Austria-Hungary/Austria in red, and Nazi Germany in brown. Territory of modern-day South Tyrol is highlighted in a darker shade. Borders of the historical County of Tyrol are indicated by a black line. Shapefiles provided by the Census Mosaic Project: https://censusmosaic.demog.berkeley.edu/data/historical-gis-files (last accessed on 03/15/2023).

Annexation and Italianization After the signing of the Treaty of Saint-Germain-en-Laye in 1919, South Tyrol was ceded to Italy, while North and East Tyrol became part of the newly formed Republic of Austria.[11] The annexation of German-speaking South Tyrol and other Italian-speaking Austrian possessions was promised to Italy by the Western Allies as a reward for entering the war on their side (Grote 2012).[12] As a result, the German-speaking population of South Tyrol – who for centuries formed the majority – suddenly found itself a minority under the rule of their former war opponent Italy.

The rise to power of Italian Fascism under Benito Mussolini in 1922 not only marked a turning point for European nationalism but also significantly deteriorated the situation for the German-speaking population in South Tyrol. Fascist ideology was deeply rooted in Italian nationalism and imperialism propagating unification through forced cultural assimilation of all ethnic minorities under Italian rule (Kallis 2002). This policy of Italianization was immediately implemented in South Tyrol in order to erase all traces of German culture among the local population (Di Michele 2008; Grote 2012; Steininger 1997b, 2003). Specific measures included the introduction of Italian as the only official language, the establishment of an exclusively Italian schooling system, and the dismissal of all German-speaking officials from public service.[13] In an attempt to eradicate its German identity, all geographic places were given newly created Italian names, including the name of the region itself, which was changed from *South Tyrol* to *Alto Adige* (Grote 2012).[14] Fascist authorities went so far as to draw up lists with Italianized versions of all fore- and surnames in South Tyrol.[15] In addi-

[11] Figure 3.1b depicts the post-war border between Italy and the newly formed Republic of Austria.

[12] Italy's claim on the predominantly German-speaking South Tyrol was justified by the so-called "natural boundary theory" promoted by Italian nationalists, based on the idea that every ethnically homogeneous nation-state should be marked by natural borders. In fact, the new border between Austria and Italy was exactly defined by the main ridge and watershed of the Alps. Although this result stood in stark contrast to the ideas of self-determination promoted by U.S. president Woodrow Wilson in post-WWI peace negotiations, the Western allies ultimately accepted Italy's annexation of South Tyrol. Apart from nationalist ideology, historians have also highlighted the military-strategic and economic advantages for Italy of controlling the Brenner Pass that forms the main gateway over the Eastern Alpine range (see e.g. Grote 2012).

[13] The blueprint for the Italianization campaign were the 32 *"Provvedimenti per l'Alto Adige"* ("Measures for the Alto Adige") presented by radical Italian nationalist and Fascist politician Ettore Tolomei in 1923. A complete list is provided in Grote (2012, p. 37). From a theoretical perspective, these provisions are in line with models of nation-building policies by homogenization as in Alesina and Reich (2015).

[14] The name *Alto Adige*, literally meaning *Upper Adige* referring to Italy's second-longest river, explicitly emphasizes the geographical connectedness to the Italian lands in the South, while the designation *South Tyrol* naturally highlights the connection to formerly unified Tyrol and Austria in the North. The basis of renamings formed the *"Prontuario dei nomi locali dell'Alto Adige"* ("Reference Work of Place Names of Alto Adige"), also developed by Tolomei, featuring Italian toponyms for each village, river, forest, and mountain in South Tyrol. These Italian translations are to this day in official use in South Tyrol (bilingually with the German version) and continue to be a divisive issue (Mumelter 2017).

[15] Although the Italianization of all German names formed part of Tolomei's 32 provisions, it was never comprehensively implemented and only individuals actively seeking assistance from the government were forced to Italianize their name in order to be eligible (Grote 2012, p. 38).

tion, this de-nationalization campaign was complemented by the targeted settlement of migrants from across Italy to decrease the predominance of the German-speaking population. The intense efforts to marginalize their cultural identity were met with backlash and fierce resistance by the German-speaking population in South Tyrol (Di Michele 2008). The best-known example concerns the establishment of clandestine *Katakombenschulen* ("catacomb schools"), secretly providing education to children in German by dismissed school teachers.[16]

3.2.2 The South Tyrol Option

After Adolf Hitler and the Nazi Party came to power in Germany, aggressive expansionism dominated its foreign policy. The *"Heim ins Reich"* (literally: "Back to the Empire") ideology sought to unify all ethnic Germans in one "Greater German Empire".[17] Many South Tyroleans greeted Hitler's Pan-Germanic course with enthusiasm and saw Germany as a protector of their culture against the cultural oppression faced in Italy (Grote 2012, p. 65).[18] These expectations were fueled by the first results of Nazi Germany's expansionism with the annexations of Austria and the Sudetenland in 1938. However, the strategic importance of an alliance with Fascist Italy made it politically infeasible for Hitler to seek territorial annexation of South Tyrol.[19] At the same time, the continued resistance of South Tyroleans against the – ultimately failed – Italianization campaign posed an increasing political problem for Mussolini (Di Michele 2008).

On June 23, 1939, the two dictators reached the *South Tyrol Option Agreement* on a relocation of the German-speaking population, concluding that "if the South Tyrolean issue was not going to go away, so the people must" (Grote 2012, p. 67). Essentially, the agreement presented South Tyrolean Germans with a choice, the so-called "Option": they could either leave for Germany to become citizens of the Reich and retain their cultural identity, or stay in their homeland but be subjected to total Italian assimilation. The Option Agreement was immediately announced to the public and the German-

[16] Recent research has investigated similar cases of cultural backlash to forced assimilation by German immigrants in the United States (Fouka 2020), Muslim immigrants in France (Abdelgadir and Fouka 2020), or by Muslims in Indonesia (Bazzi, Hilmy, and Marx 2022). An encompassing theoretical framework on educational resistance to cultural assimilation is provided by Carvalho, Koyama, and Williams (2022).

[17] The importance of the *"Heim ins Reich"* ideology for the case of South Tyrol is exemplified by Panel (B) of Appendix Figure 3.A.2, showing a contemporaneous Nazi propaganda poster.

[18] Hannah Arendt pointed out that it would be a mistake to see the behavior of South Tyrolean emigrants as an example of mere "fanatic nationalist sentiment"; rather "these people no longer felt sure of their elementary rights if these were not protected by a government to which they belonged by birth" (Arendt 1962, p. 292).

[19] In fact, Hitler had made it clear early on that he would be willing to sacrifice South Tyrol for the "greater good of Germany" and formally rejected any claims in the 1936 Axis treaty with Fascist Italy. Nevertheless, pro-Nazi organizations in South Tyrol, such as the *Völkische Kampfring Südtirols*, continued to agitate for reunification with Germany and Austria throughout the 1930s (Grote 2012, pp. 65–66).

speaking population in South Tyrol was given a six-month period until December 31, 1939 to make their choice.[20]

Every head of household had to declare their decision in favor of or against accepting German citizenship, with a non-response considered as opting for Italy.[21] Those who opted in favor of Germany were promised financial compensation for loss of material possession as well as relocation to new settlement areas for South Tyroleans. Historians estimate that around 84% of eligible South Tyrolean households opted in favor of Germany. In the end, however, only around 75,000 individuals, i.e., slightly fewer than half of declared opters, actually left for Germany in the belief of never being able to return (Wedekind 2003).[22]

The Option confronted many South Tyroleans with a difficult choice, leading to the division of society into so-called *"Dableiber"* (literally "stayers", i.e., those in favor of remaining in South Tyrol) and *"Optanten"* ("opters", i.e., those willing to leave for Germany), with rifts running through villages, families, and friendships. The Option was strongly felt as a choice about national and cultural identity. Pamphlets agitating in favor or against the Option make this evident:

> *"The decision is difficult, but not for a moment doubtful, because we know what we owe to the call of our German blood [...] we sacrifice the land for the great goal, the great, holy German Reich."*

> *"It is about emigration or staying in the land, about home or foreign country. The choice cannot be difficult. [...] Whoever therefore signs the white slip of paper [ballot paper for opting for Italy] gives his vote to the homeland."*[23]

The Option Agreement and mass emigration effectively came to a premature end in September 1943 when Nazi Germany occupied South Tyrol after the overthrow of Mussolini's Fascist regime and Italy's armistice with the Allies. After the end of

[20] Technically, the Option was applied to all ethnic Germans (*Volksdeutsche*) with Italian citizenship which included a number of small German-speaking enclaves in other Italian provinces such as Trentino, Belluno, Vicenza, and Udine as well as those residing outside of Italy. However, the approximately 235,000 German-speaking South Tyroleans, i.e., those living in the Province of Bolzano, formed more than 90% of the eligible population (Wedekind 2003). Residents of South Tyrol with German citizenship were left with no choice but to resettle (Lutt 2016).

[21] While the majority of heads of household were men who co-decided for their wives and underage children, there also was a significant number of female eligible opters, such as widows or unmarried women. A more detailed discussion of the exact administrative procedure of the Option will be given in Section 3.3.

[22] Aggregated statistics on the results of the Option should be treated with caution. Given the lack of a surviving central register, and the difference between eligible households and actual individuals, historical estimates vary substantially and are hard to verify (Alexander, Lechner, and Leidlmair 1993, pp. 24–25).

[23] Pamphlets quoted from Steininger (1997a, pp. 402–404) and translated by the authors.

WWII, South Tyrol remained a part of Italy, despite Austrian efforts for reunification (Steininger 2006). In 1948, Italy offered everyone who had opted for Germany to regain Italian citizenship, after which an estimated 20,000 South Tyroleans returned (Lutt 2016). After another three decades of ethnic conflict, culminating in a terrorist campaign, the South Tyrolean issue was eventually resolved in the 1970s with the granting of extensive autonomy rights to the South Tyrol province and the German-speaking population.[24]

3.3 Data

3.3.1 Data Sources

World War I Casualty Lists We gather individual-level data on WWI combat experience from casualty records of the Austro-Hungarian army. Specifically, we combine two digitized historical sources to maximize coverage. First, we obtain the complete list of fallen Tyrolean soldiers from the *Tiroler Ehrenbücher* (Tyrolean Honor Books, henceforth *Ehrenbücher*; Tiroler Landesmuseen 2014). They were compiled after the end of WWI with the explicit aim of recording all dead soldiers from Tyrol and were published in 120 volumes in 1930. Entries for each of the 23,756 fallen soldiers contain detailed personal information: date and place of birth, place of residence, occupation,[25] family status, military unit (regiment and company) in which the soldier served during the war as well as date, place, and cause of death. Figure 3.A.5 in the Appendix shows an exemplary entry from the *Ehrenbücher*.

Second, we complement these records with the *Verlustlisten Österreich-Ungarns* (Casualty Lists of Austria-Hungary, henceforth *Verlustlisten*; Verein für Computergenealogie 2021). These lists were published almost daily throughout WWI by the Austro-Hungarian war ministry. They recorded all recent casualties of the armed forces, i.e., soldiers that had died, been wounded, or captured by the enemy. Again, all entries contain information on the date and place of birth, military unit (regiment and company) as well as the date and place of the casualty. In total, these lists contain around 2.7 million entries, whereby we focus on those 32,380 soldiers recorded as originating from Tyrol. While these contemporaneous casualty lists likely underreport the number of dead soldiers with respect to the more complete retrospectively collected *Ehrenbücher*, they provide the important advantage of including potentially surviving, i.e., wounded or captured, soldiers in our analysis. Figure 3.A.6 in the Appendix shows an exemplary entry from the *Verlustlisten* for a soldier in Italian captivity.

[24] An extensive discussion of post-1945 South Tyrol with an empirical study of the terrorist attacks can be found in Belmonte (2022) and Belmonte and Di Lillo (2021).

[25] We manually classified all entries into four categories: skilled, semi-skilled, and unskilled labor as well as farming-related occupations.

In order to obtain an enemy-specific measure of war experience, we locate each entry in the casualty lists to the specific WWI battle front. We code a binary variable that captures whether a soldier was killed, wounded, or captured on the Italian front, or on one of the other fronts the Austro-Hungarian armed forces were fighting on. In order to assign a casualty list entry to a battle front, we proceed in two steps: first, we manually code whether an explicitly mentioned place of the recorded event lies on the Italian front. Second, for entries lacking this information, we rely on historical accounts by Glaise von Horstenau (1932) to infer the front to which a soldier's military unit (regiment) was deployed at the time of the recorded entry date. Nearly all Tyrolean soldiers were transferred from the Eastern to the Italian front over the summer of 1915 to counter the Italian attack.[26] We leverage this front-specific combat exposure to identify enemy-specific war grievances.

Emigration Requests The procedure to leave South Tyrol for Germany confronted eligible heads of households with a two-step emigration process (Alexander, Lechner, and Leidlmair 1993). In a first step, after the announcement of the Option Agreement in June 1939, they had to declare their intent to opt for Germany or Italy until December 31, 1939 at the local Italian municipal office. The remainder of the process was handled by the *Amtliche Deutsche Ein- und Rückwanderungsstelle* (ADERSt, Official German Immigration and Repatriation Office), an administrative authority specifically created for facilitating the mass emigration of South Tyroleans. The ADERSt proceeded by assigning all opting individuals a unique identification number and prepared a personal file for each opting household to collect any further documentation and correspondence. In the second step of their decision, opters had to go to the local ADERSt office to formally initiate the emigration procedure by renouncing Italian citizenship and signing a request for German citizenship. Once the value assessment of property for compensation was completed and the request processed, households received details on their departure. All emigrants were first brought to Innsbruck in Austria, then part of Germany, where they were centrally registered and temporarily housed before traveling on to their final destination on their own.

Our dataset consists of a random sample of 2,388 ADERSt emigration requests that were manually digitized at the *Archivio di Stato di Bolzano/Staatsarchiv Bozen* and first used in Lochmann (2020).[27] While these files often vary substantially in scope and in-

[26] As can be seen in Figure 3.A.7 and Figure 3.A.8 in the Appendix, the resulting distribution of casualty list entries shows how Tyrolean soldiers first fought on the other fronts, mainly in the East against Russia, before mostly being moved to the Italian Front after May 1915.

[27] An estimated 100,000 of these files still exist corresponding to the universe of files from the ADERSt offices in Meran, Brixen, and Bruneck, while those of the office in Bozen have only fragmentarily survived. The files are held in boxes of which 25 were randomly drawn, stratified by ADERSt office. On average, each box contains 93.5 files ($sd = 32.5$). A more detailed description of the sampling procedure as well as of the digitization process can be found in Lochmann (2020).

cluded documents, all of them contain the so-called *Abwanderungsantrag* (emigration request). This three-page form lists an extensive number of personal information on the head of the household and their family members.[28] The files contain information about the place and date of birth, residential address, family status, occupation[29], religion, citizenship, ethnicity, all former places of residence, and, if applicable, military, criminal, and health records of the applying head of household. Furthermore, the form records the full name, date and place of birth of the wife and all children of the applicant. Importantly, also the full name, place of residence and, if applicable, date of death of the opter's parents are recorded. The final page records financial assets and property as well as self-reported information on the favored emigration destination. Crucially, the file allows us to infer whether the respective household followed through with their initial option intention: a stamp on the cover sheet or first page of the request with the letter *"A"* for *"abgewandert"* (emigrated) indicates that the file had been closed and the listed persons definitely emigrated to Germany (Lutt 2016, p. 81). As an example, a full scan of an emigration request is provided in Figures 3.A.9-3.A.12 in the Appendix.

Historical Census Data We complement our individual-level data with a number of variables measured on a more aggregate level. We digitize the pre-WWI census for Tyrol (k. k. Statistische Zentralkommission 1907), containing information on the municipality-level (*Gemeinde*) of the resident population and its distribution by gender and ethnicity, as well as information on the local economic structure.

As our data sources use information from different years and South Tyrolean municipalities were subject to various administrative reforms over time, we link all municipalities to their corresponding municipality as of 1940, i.e., to the time of the Option.[30] To track changes of administrative borders over time, we rely on historical information compiled by *Storia dei Comuni*.[31] For all 104 municipalities, we furthermore code whether they share a border with other Italian provinces or Austria from historical shapefiles provided by the *Geoportal Südtirol* (Autonome Provinz Bozen 2022). Moreover, we manually harmonize all reported municipalities in the casualty lists and emigration records in order to complement them with municipal-level characteristics.

[28] While most of the heads of households were men, e.g. as by default husbands decided for their wives, women could be entitled to the Option in the case of unmarried women of full age or widows. Indeed, in our sample of personal files, we observe 32.18% (n = 752) female opters.

[29] In the same way as for the casualty records, we manually classified all entries into four categories: skilled, semi-skilled, and unskilled labor as well as farming-related occupations. Furthermore, we manually checked whether the individual currently works in their learnt profession.

[30] As part of a nation-wide policy of centralization, the Fascist regime merged many municipalities throughout the 1920s and 1930s, reducing the total number of administratively independent municipalities in South Tyrol from 205 in 1900 to 104 in 1940 (Di Michele 2008, p. 229)

[31] The website can be accessed via http://www.elesh.it/storiacomuni (last accessed on 06/02/2022).

3.3.2 Linking Casualty Lists to Emigration Requests

We develop a cascading algorithm to match individuals in the emigration records with entries in the WWI casualty lists. We, first, match the father of an individual to the *Ehrenbücher* that contains exclusively dead soldiers, and, second, to the *Verlustlisten* where he might have been recorded as killed, wounded, or captured. Then, we identify sufficiently old male opters themselves in the *Verlustlisten* among soldiers recorded as wounded or captured but having survived the war. We match two records if an individual's full name, home town and, if available, year of birth perfectly align in both datasources.[32] A more detailed description of all steps in the matching procedure is given in Section 3.C.

We match 244 emigration requests to entries in the WWI casualty lists. Of these, 210 relate to cases where we could identify the father of an individual in the casualty lists, while we matched 35 opting individuals with themselves.[33] To gauge the accuracy of our matching rate, consider the following comparison: among the 1,985 emigration requests which include complete information on fathers, we find 126 (6.3%) dead fathers. This corresponds approximately to the 5.6% death rate among the male Tyrolean population during World War I.[34]

Our final dataset, therefore, consists of 2,338 opters to which we matched individual-level information on war grievances for 244 individuals as well as aggregate-level data as described in Section 3.3.1. Summary statistics on all used variables are presented in Table 3.B.1 in the Appendix.

3.3.3 Who emigrated to Germany?

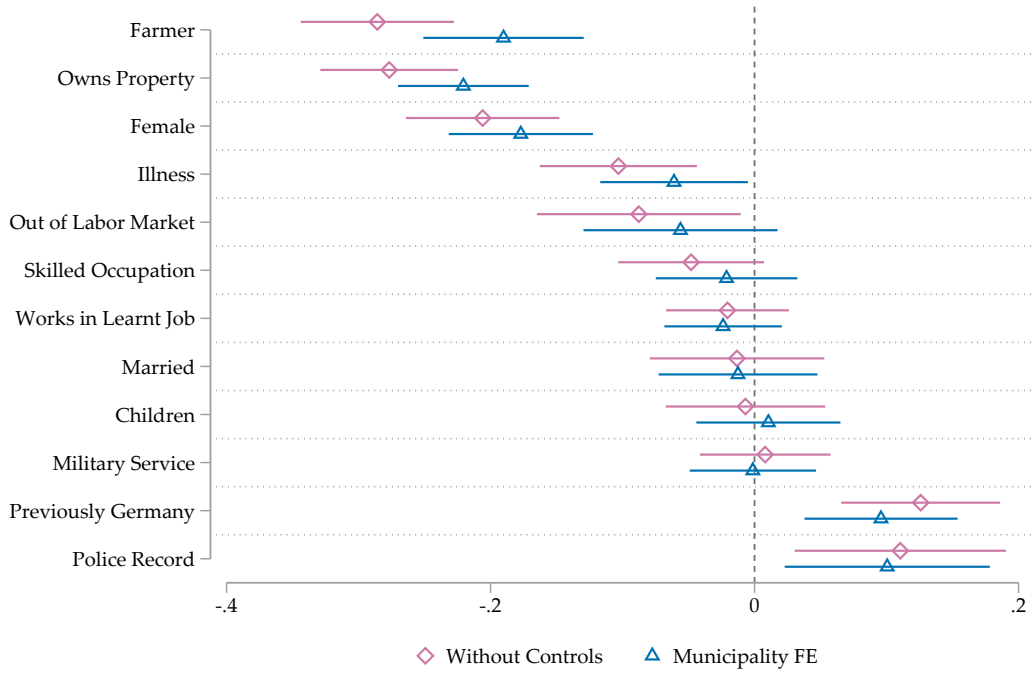
Before discussing our empirical strategy to identify the effect of war grievances, we provide descriptive evidence on the main determinants of emigrating to Germany. This exercise serves two purposes: first, it highlights that there are a number of meaningful factors that drive an individual's emigration behavior. Second, from an econometric perspective, this sheds light on some of the potential confounders driving emigration that our variation in enemy-specific war grievances must be orthogonal to.

[32] Due to our relatively small sample size, we decided against using probabilistic or fuzzy matching algorithms to keep the matching accuracy as high as possible. Matching via year of birth is only possible when looking up opters themselves in the casualty lists, as the option files do not provide the date of birth for the parents. To account for the missing age information in case of linking fathers and sons, we only consider potential matches if they pass a sanity check of a minimum age difference of 16 years.

[33] These individual numbers do not add up to 244 as there is one case where both the opter himself was wounded during World War I as well as his father was recorded as having fallen. Table 3.C.1 in the Appendix summarizes the matching results.

[34] This share is derived by dividing the 23,756 entries in the *Ehrenbücher*, constituting our best estimate of the total number of fallen soldiers, by the total Tyrolean male population of 422,726 in the 1900 census (k. k. Statistische Zentralkommission 1907, p. 158), arguably a – if at all overestimated – proxy for the pool of potential soldiers between 1914 and 1918. This corresponds well to twice of the death rate among the *total* Tyrolean population of 2.7% reported in contemporaneous accounts by Winkler (1919).

Figure 3.2: Determinants of Emigration to Germany



Notes: This figure plots coefficients from a regression of household *i*'s emigration behavior on the full set of individual-level controls: being female, having previous military experience, having illnesses, having a police record, having previously migrated to Germany, having children, being married, owning property, being out of the labor market, current job being equal to learnt job, being a farmer, and working in a skilled occupation. Furthermore, all regressions control for age and its square. Two sets of coefficients are reported: the first set reports coefficients of a multivariate regression on the aforementioned variables (red markers), and the second set (blue markers) reports coefficients from a regression additionally controlling for municipality fixed effects. 95% confidence intervals are reported. Full regression results can be found in columns (1) and (2) of Table 3.B.2.

Overall, 57% of opters in our dataset eventually moved to Germany. Figure 3.2 visualizes the results of a linear regression of a household emigrating on a number of individual-specific variables obtained from the emigration requests.[35] The red markers show the point estimates and corresponding 95% confidence intervals from a joint regression on all stated covariates, in addition to linearly and quadratically controlling for a person's age.[36]

The strongest predictors of ultimately not emigrating to Germany are economic factors, with working in a farming-related occupation or owning property being associated with a 20 percentage points lower probability to emigrate. Similarly, female heads of household as well as those with a recorded illness are significantly less likely to leave for Germany. In contrast, having previously lived in Germany, potentially prox-

[35] The corresponding results in form of a regression table can be found in Table 3.B.2 in the Appendix.

[36] As can be seen in Column (2) of Table 3.B.2, the effect of age is negative but decreasing, suggesting that both relatively young and old people are more likely to emigrate than middle-aged opters.

ying for existing personal ties, and having a police record, i.e., likely having been in conflict with Italian authorities, is positively correlated with the probability to follow through with emigration. When adding municipality fixed effects to account for baseline differences in emigration behavior between municipalities, results remain largely unchanged (blue markers).[37]

In sum, this discussion highlights that numerous economic and demographic factors influenced the difficult choice South Tyroleans had to make. It should, however, be noted that these estimates are merely correlational, with no claim to causality. In the remainder of this paper, we are interested in isolating and causally identifying the effect of an individual-level war shock on an individual's identity and behavior. We now turn to discussing how we leverage South Tyrol's history during World War I to exploit variation unrelated to any other factors influencing the emigration decision.

3.4 Identification Strategy

3.4.1 Empirical Setup

The South Tyrol Option and the ensuing mass emigration represents a large and unprecedented episode of nationalist mobilization. Whether an individual decided to remain in Italy or emigrate to Germany was influenced by a multitude of factors, ranging from the socio-economic background to family status and networks. Crucially, this choice forced individuals to reveal their preference with respect to their national identity: emigrating to Germany was the "price" they had to pay for retaining their "Germanness".[38] We focus on the role of war grievances on individual behavior in this episode of pro-German (and anti-Italian) nationalist mass mobilization. In particular, we are interested in the effect of war grievances being directed against Italy, as opposed to other enemies in World War I.

We propose an individual-level grievance shock based on exogenous variation in the front exposure of South Tyrolean soldiers. To isolate such Italy-specific grievances from other confounding factors, we exploit Italy's unexpected entry to World War I as a natural experiment. In the first step, we identify those individuals in the emigration

[37] One might furthermore be concerned that the observed emigration behavior was also driven by administrative factors. For example, the historical literature mentions that the process could take longer for farmers and landowners due to the necessary value assessment (Alexander, Lechner, and Leidlmair 1993, p. 51). In our main specification, we can directly control for these factors. Moreover, we can test whether our results are driven by the respective ADERSt officers in charge of the requests – whose signatures appear on the emigration requests – that might be handling different types of cases or work under different decision parameters. As shown in column (3) of Table 3.B.2, after the inclusion of officer fixed effects, being a farmer, being female, owning property, age, and having a police record still remain the strongest predictors for emigration. Later, we also show that our main results are robust to the inclusion of ADERSt officer fixed effects.

[38] See also the discussion of contemporaneous debates among South Tyroleans in Section 3.2.2.

requests whose relative or who themselves fought in World War I and experienced a violent war event (see Section 3.3.2). In a second step, within this set of people, we then identify those whose war-related grievance stems from fighting on the Italian front and, by extension, having suffered violence caused by Italians.[39]

In line with insights from social psychology (e.g., Choi and Bowles 2007; Haidt 2013; Henrich 2020; Tajfel and Turner 1979), we hypothesize that having endured such an Italy-specific casualty, i.e., a negative exposure to the out-group, intensifies resentment against Italians. If such grievances mobilize individuals into complying with nationalist policy, we expect that the treatment group, i.e., those whose war grievance is directed against Italy, has higher rates of emigration than the control group, i.e., those whose war grievance is directed at another enemy. Importantly, this approach allows us to isolate the enemy-specific component of an individual's war grievances, thereby holding fixed having incurred any war grievance in the first place. In other words, we compare otherwise identical individuals who only differ in the fact that they hold specifically anti-Italian grievances.

Our main identifying assumption is that a casualty happening on the Italian front is not systematically related to unobserved factors affecting individuals' probability to leave for Germany two decades later. In other words, those emigrating to Germany were not systematically more or less likely to have been violently exposed on the Italian front.[40] One might worry that individuals were able to select into different fronts; yet, this is historically implausible. As explained in Section 3.2.1, all South Tyrolean soldiers were initially fighting on the Eastern front. These South Tyrolean soldiers were then among the first to be recalled to fight the Italian army on the newly opening Southern front. This sudden and rapid redeployment led to most subsequent casualties occurring on the Italian front.[41] As such, selection into different fronts is ruled out in this historical setting, allowing for exogenous variation in front exposure. We will provide further evidence underscoring these claims below (see also Section 3.4.2).

The causal effect of anti-Italian war grievances on emigrating to Germany is estimated in the following regression:

$$\begin{aligned} Emigrate_{ima} = \gamma_1 \cdot WarGrievance_i + \gamma_2 \cdot WarGrievance_i \times ItalianFront_i \\ + X_i'\beta + \theta_m + \mu_a + \epsilon_{ima} \end{aligned} \tag{3.1}$$

[39] Note that we call such a violent war event a "war grievance" in order to distinguish it conceptually from other grievances people may have held. Crucially, we understand an Italy-specific war grievance as intensifying the degree of anti-Italian sentiments in individuals.

[40] We will discuss a number of potential violations of this identifying assumption and provide evidence underscoring our claims in Section 3.4.2 below. Furthermore, in Section 3.5.2 we report that the estimated size of our main treatment coefficient is stable across a number of different specifications taking potential alternative explanations into account. This is consistent with the treatment being exogenous.

[41] This is visualized by the distributions of recorded casualty entries by front in Figure 3.A.7 and Figure 3.A.8 in the Appendix.

where $Emigrate_i$ is a binary variable taking value 1 if household i accepted German citizenship and emigrated to Germany. $WarGrievance_i$ is a dummy variable taking value 1 if opter i's family experienced a casualty during WWI; in other words, individual i was matched to a soldier in our linking procedure. The interaction term $WarGrievance_i \times ItalianFront_i$ takes value 1 only if said casualty was recorded on the Italian front. X_i contains an array of individual- and municipal-level controls. Finally, we include fixed effects for the municipality of residence m of individual i (θ_m) as well as for the army unit a in which the soldier incurring the WWI casualty served in (μ_a).

The estimate of γ_1 can be interpreted as the percentage point change in emigration probability resulting from holding any war grievance. This effect is, however, not straightforward to interpret, as it captures numerous financial, emotional, and other factors at once.[42] It is furthermore likely biased due to selection into suffering a casualty or even fighting in the war in the first place. One must, for example, have been of a certain age and healthy enough to have been drawn into the armed forces. Due to these endogeneity problems, we do not formulate a hypothesis regarding this estimate and remain agnostic about its direction and size. In our analysis, we treat this effect as a baseline control variable, as our main estimate is conditional on the effect of $WarGrievance_i$, i.e., having experienced some casualty in the first place.[43]

Our main coefficient of interest is γ_2 and captures the *differential* effect of a casualty happening on the Italian front. This allows us to hold constant the fact that an individual holds any war grievance and isolate the emotional component of such grievances being directed at Italy. We hypothesize that this effect is positive, meaning that anti-Italian war grievances make individuals more likely to ultimately emigrate to Germany.[44]

3.4.2 Threats to Identification

To identify the Italy-specific component of war grievances, we make the identifying assumption that suffering a casualty on the Italian front is orthogonal to other unobserved factors driving an individual's emigration decision. Fighting in World War I is, of course, not a random event. Moreover, in the absence of randomized assignment of soldiers to a front, it is in principle possible that soldiers selected themselves or

[42] Dupraz and Ferrara (2023) show that losing a father in war has long-lasting effects on individuals' socio-economic outcomes. By identifying from within households who have experienced a casualty in WWI, we can hold these channels fixed and isolate the enemy-specific direction of the grievance.

[43] Note that in specifications including army unit fixed effects, the baseline effect of holding any war grievance cannot be estimated, since the sum of army unit fixed effects is perfectly collinear with experiencing a casualty.

[44] In the main specification, we do not distinguish between different types of war events. We later show that our results are not affected by controlling for the specific kind of event at the root of Italy-specific grievances, i.e., whether one's father was killed, wounded, or captured, or whether the individuals themselves were wounded or captured.

were systematically selected into specific fronts or army units. Such selection would compromise our identification strategy, if it were systematically related to soldiers' or their children's later emigration behavior.

We offer four arguments to rule out such potential selection concerns. The first argument is historical in nature: before May 1915, all South Tyrolean casualties occurred on other fronts, most of them on the Eastern Front against Russia. With the sudden opening of the Italian front in 1915, the seven main regiments South Tyrolean soldiers were fighting in were suddenly redeployed to the South.[45] As can be seen in Figures 3.A.7 and 3.A.8 in the Appendix, soon afterwards a large majority of casualties were indeed recorded on the Italian front. Given the unexpected nature of Italy's entry into the war and the sudden and nearly complete reassignment of South Tyrolean soldiers, systematic selection of soldiers into fighting on a specific front was highly improbable. As such, we argue that, conditional on having participated in the war, the front a soldier fought on is orthogonal to other factors influencing their families' emigration behavior two decades later.

Second, we show that South Tyrolean soldiers who experienced a WWI casualty in WWI but did so on different fronts, were not systematically different from each other. Drawing on soldier-level data from the universe of fallen South Tyrolean soldiers, we test whether pre-war characteristics predict dying on the Italian front (Table 3.1). Soldiers who died on the Italian front, relative to other fronts, do not differ with respect to their socio-economic background and come from similarly sized towns that have a comparable ethnic and economic structure. There are three variables that appear to predict dying on the Italian front: the soldier's age, marital status, and being part of the reserve force. These factors are the mechanical result of the development of the war: only with the outbreak of the war against Italy were soldiers from reserve units drafted en masse to ensure the defense of Austrian territory in the South.[46] These men were on average older and more likely to be married. When omitting soldiers from reserve units from the balancing table (see Appendix Table 3.B.3), we find that observable individual characteristics are well balanced; only age continues to be significantly different, with those dying (later) on the Italian front being, as expected, slightly younger.[47] We conclude that the location where an individual experienced a

[45]The Austro-Hungarian armed forces recruited soldiers based on territorial principles, such that regiments were comprised of soldiers from the same districts. Most Tyrolean soldiers served in the 4 *Kaiserjäger* and 3 *Landesschützen* regiments (Wandruszka and Urbanitsch 1987). Indeed, in our WWI casualty records about 80% of entries can be attributed to these seven regiments.

[46]More specifically, we observe two types of reserve forces in our data, the Austrian Imperial *Landsturm* and the Tyrol-specific *Standschützen*. Both were militia-type units that could be called up as the "last reserve" in case of an imminent threat to Tyrol's territorial integrity, as happened in May 1915 when their first order was to hold the Italian front until the arrival of the regular troops. Importantly, reserve soldiers fought in separate units from the regular forces throughout the war (Glaise von Horstenau 1932). This allows us to later single them out and control for potential bias by including army unit fixed effects.

[47]In our main results, we always control for the birth year of individuals as well as army unit fixed

Table 3.1: Balancing of Soldiers on Italian vs. Other Fronts

	Other Front	Italian Front	Diff.	P-val.	Obs.
Individual characteristics					
Birth Year	1885.69	1886.05	-0.36*	(0.07)	6718
Farmer	0.46	0.46	-0.00	(0.86)	6489
Skilled Occupation	0.30	0.32	-0.01	(0.24)	6489
Married	0.24	0.28	-0.04***	(0.00)	6202
Reserve Force	0.06	0.26	-0.20***	(0.00)	6145
Municipal characteristics					
Total Population	4452.31	4662.04	-209.73	(0.14)	7210
Population Density (per ha)	27.24	29.21	-1.97	(0.38)	6869
Share Male (%)	50.09	49.99	0.10**	(0.03)	7210
Share Italian (%)	3.51	3.52	-0.00	(0.98)	7210
Share German (%)	88.52	89.08	-0.56	(0.28)	7210
Share Ladin (%)	5.50	4.79	0.71	(0.16)	7210
Share Taxable Land (%)	85.13	85.45	-0.32	(0.48)	6869
No. Factories	0.14	0.15	-0.01	(0.28)	7210
Cattle per capita	0.34	0.34	0.00	(0.70)	7210

Notes: This table reports results from two-group mean-comparison tests by treatment status, i.e., whether a war grievance happened on Italian or another front. Columns report group averages, the difference between averages, p-values for mean equality, and the number of observations. The sample includes all dead South Tyrolean soldiers from *Ehrenbücher* for whom we were able to locate their death on a specific front. Significance levels: *** p<0.01, ** p<0.05, and * p<0.1.

violent war event is unrelated to a number of soldier-specific observable characteristics. Given that the same army units and even the same individual soldiers served first on the Eastern and then on the Italian front, this result is not surprising and in line with historical evidence.

Third, we show that having suffered an Italy-specific war grievance does not predict other individual-level outcomes relating to the emigration decision. We report conditional averages of our main control variables in Table 3.2, depending on their treatment status. We find very few systematic differences between treated and untreated individuals. Importantly, this holds for socio-economic outcomes which strongly predict emigration behavior, such as being a farmer or owning property. While this analysis does not directly test for selection into front exposure, it shows that the treatment has no meaningful effects on other variables.

Fourth, we address potential sources of selection directly in our main regressions. One concern is that individuals were drafted into certain army units based on geographic location. If this correlates with other geographic determinants of emigration, our results might be biased. In fact, the Austro-Hungarian Army drafted soldiers into army

effects. Later on, we also show that our results are robust to control for the birth year of fathers as well as for the direct exclusion of reserve force soldiers from our sample.

Table 3.2: Summary Statistics by Treatment of Matched Opters

	Other Front	Italian Front	Diff.	P-val.	Obs.
Emigrated to Germany	0.48	0.60	-0.13**	(0.05)	242
Birth Year	1907.50	1905.88	1.62	(0.21)	242
Female	0.26	0.28	-0.02	(0.70)	242
Military Service	0.55	0.55	-0.00	(0.98)	242
Illness	0.11	0.12	-0.01	(0.80)	242
Police Record	0.04	0.02	0.02	(0.36)	242
Previously Germany	0.06	0.07	-0.01	(0.81)	242
Children	0.31	0.36	-0.05	(0.37)	242
Married	0.28	0.28	0.00	(0.94)	242
Owns Property	0.29	0.24	0.05	(0.39)	222
Out of Labor Market	0.03	0.09	-0.06*	(0.05)	236
Works in Learnt Job	0.78	0.81	-0.03	(0.60)	242
Farmer	0.52	0.51	0.01	(0.87)	242
Skilled Occupation	0.32	0.29	0.02	(0.69)	242

Notes: This table reports averages of the control variables used in the main analysis, conditional on treatment status, i.e., whether the war grievance happened on the Italian front or on another front. Differences and p-values from two-group mean-comparison tests are included. These averages are drawn from the sample of all opters who we could link to a violent war experience and for which we could locate the event to a front of WWI. While we matched 244 opters to WWI casualty records, for 2 individuals we could not assign their casualty to a battlefront.

units based on their municipality of origin (Wandruszka and Urbanitsch 1987). To control directly for such potential selection effects, we show that our baseline results are robust to including army unit as well as municipality fixed effects.[48] Any potential selection into army units based on municipalities is therefore controlled for. Moreover, common shocks affecting individuals in the same locality, army unit, or age cohort are of no concern once these fixed effects are included. Any remaining bias must stem from a systematic individual-level relationship between holding Italy-specific war grievances and emigrating to Germany. Additionally, Figure 3.A.3 in the Appendix shows that there is no obvious geographic pattern in the local share of casualties recorded on the Italian front.[49]

3.5 Results

3.5.1 Main Results

In Table 3.3, we report our main results from estimating Equation (3.1). Throughout all five columns, we find a positive and significant effect for a war grievance stem-

[48] We return to these tests in Section 3.5.2 when we discuss Table 3.4. In Section 3.5.2, we discuss further robustness tests addressing concerns that certain subgroups might be driving our results. Our results are robust to these different specifications.

[49] Figure 3.A.4 additionally shows the geographic distribution of matched individuals by treatment status, i.e., from which front the war grievance originated.

ming from the Italian front, relative to it happening on another front. Column (1) reports the estimates from a regression without any controls. The first coefficient estimate of -0.097 (standard error: 0.048) captures the effect of holding any war grievance. Individuals whose fathers or who themselves experienced a casualty in WWI were 9.7 percentage points less likely to emigrate to Germany than individuals who did not hold war grievances. As discussed in Section 3.4.1, we do not interpret this effect causally as it may well be biased. The second coefficient estimate of 0.126 (standard error: 0.063) captures the *differential* effect of having experienced a casualty on the Italian front. Under our identifying assumption, we interpret this effect causally: holding an Italy-specific war grievance causes individuals to be 12.6 percentage points more likely to emigrate to Germany. Relative to a baseline emigration probability of 57%, this amounts to a 20 percent increase. Note again that this effect is identified from within those individuals who hold any war grievance, enabling us to disentangle whom the war grievance is directed at.

The remainder of the table shows that this result remains stable with the inclusion of various control variables and fixed effects. The estimated effect of Italy-specific war grievances on an individual's observed emigration behavior lies consistently between 0.11 and 0.15. That is, relative to other households who hold comparable war grievances, those war grievances which are directed specifically against Italy increase the probability of moving to Germany by 11 to 15 percentage points. In column (2), we add individual-level controls and control non-parametrically for age by adding a dummy variable for all observed birth years of opters.[50] While the estimate for the main war grievance effect decreases and becomes insignificant, the estimate for the Italian interaction remains unaffected.

One concern is that soldiers were systematically selected into army units in a way that correlates with their later emigration behavior. While this is unlikely – soldiers were drafted into army units based on predetermined geographic areas, but not assigned to different fronts based on this selection – we nevertheless address these concerns. To control for potential selection into military units, we control for army unit fixed effects in column (3). In doing so, we hold fixed soldiers' battle exposure and, thereby, abstract away from other army unit-specific experiences, e.g., whether certain units participated in particularly intense battles. The coefficient increases slightly in magnitude but, importantly, remains qualitatively similar and significant.[51]

Finally, columns (4) and (5) address the worry that unobserved local conditions simultaneously affect serving on the Italian front as well as emigration. In column (4), we

[50] The set of individual-level control variables corresponds to all determinants of the emigration decision discussed in Section 3.3.3.

[51] Note that from this specification onward the baseline effect of holding any war grievance can no longer be estimated as it is perfectly collinear with the sum of the army unit fixed effects.

Table 3.3: Main Results

	Emigrated to Germany				
	(1)	(2)	(3)	(4)	(5)
War Grievance	-0.097**	-0.065			
	(0.048)	(0.046)			
War Grievance × Italian Front	0.126**	0.119*	0.145**	0.146**	0.109*
	(0.063)	(0.061)	(0.064)	(0.066)	(0.058)
Individual Controls	-	✓	✓	✓	✓
Municipality Controls	-	-	-	✓	-
Birth Year FE	-	✓	✓	✓	✓
Army Unit FE	-	-	✓	✓	✓
Municipality FE	-	-	-	-	✓
Observations	2338	2032	2032	1680	2016
Adjusted R^2	0.001	0.246	0.248	0.248	0.391
Dependent Variable Mean	0.572	0.570	0.570	0.518	0.569

Notes: This table reports estimates of Equation (3.1). The dependent variable is the observed emigration behavior of household *i*. The first explanatory variable captures whether household *i* holds a war grievance, i.e., whether the individual or their father experienced a casualty in WWI. The second explanatory variable captures whether the war grievance of household *i* happened on the Italian front. Individual-level controls include the following variables: being female, having previous military experience, having illnesses, having a police record, having previously migrated to Germany, having children, being married, owning property, being out of the labor market, current job being equal to learnt job, being a farmer, and working in a skilled occupation. Municipal-level controls include the following variables: total population, population density, male population share, Italian population share, German population share, Ladin population share, taxable area, number of factories, and cattle per capita. Birth year fixed effects control non-parametrically for opting individual *i*'s birth year. Army Unit fixed effects are dummies for the regiment in which the soldier, i.e., opting individual *i* or *i*'s father, was assigned to during World War I at the time of the casualty. Municipality fixed effects are dummies for the municipality individual *i* lives in. We report heteroskedasticity-robust standard errors. Significance levels: *** p<0.01, ** p<0.05, and * p<0.1.

report estimates from a regression where we include municipality-level controls on the local demography (e.g., share of Italian speakers, population density) as well as economic activity (e.g., number of factories, share of cultivated area) measured before World War I. Again, the estimates remain very similar. In column (5), we go further and include fixed effects for the municipality an individual lives in, in order to control for any unobservable local aspects that might correlate with front exposure and emigration behavior. While our estimated coefficient of interest is slightly reduced in this more demanding specification, it remains qualitatively similar.

Across specifications, the estimated effect of holding an Italy-specific war grievance on observed emigration behavior remains stable. This is consistent with historical evidence that there was no systematic (self-)selection into experiencing a war event on the Italian front. Under our identifying assumption – that a war event happening

on the Italian front is not systematically related to unobserved factors affecting one's emigration probability – we interpret these coefficients causally. We find strong evidence for anti-Italian war grievances driving individuals' willingness to comply with the nationalist mass mobilization to emigrate to Germany.

3.5.2 Robustness Checks

In this section, we show that our estimated effect is robust to testing for a series of potential concerns. In particular, under non-random selection into treatment, the causal interpretation of our estimated effect would no longer hold. However, any such concern would have to argue that specific soldiers were selected into experiencing a casualty on the Italian front. The concern would simultaneously have to maintain that these individuals or their descendants were also more likely to emigrate to Germany. In Section 3.4.2, we already addressed this concern by arguing that Italian front exposure is exogenous to soldier characteristics. We now directly test for a series of these concerns by augmenting our main specification.

In the first set of robustness checks, we focus on potential selection into treatment. One might, for example, be concerned that biographic, geographic, or administrative factors determined the front a soldier fought on. While this is not problematic in itself, it would bias our results if these factors were also systematically related to individuals' emigration behavior. In our main results, reported in Table 3.3, we already added fixed effects controlling for army unit assignment, municipality, and age. In Table 3.4, we document that the estimated effect of war grievances on emigration behavior remains stable under more demanding fixed effects specifications.[52]

We have so far controlled for potential selection based on age and municipality of the opting individuals themselves. One might, however, be concerned that these specifications do not adequately address selection into treatment based on fathers' characteristics. To rule out that selection happened based on fathers' origin, we include fixed effects for fathers' hometowns in column (2). Similarly, one might be worried that opters' fathers were selected into fighting on the Italian front based on their age. We test for this concern by including fixed effects for the birth year of opters' fathers in column (3). In both regressions, the result remains stable, with it becoming even larger in magnitude and more significant in column (3).

Moreover, one might be concerned that individuals from certain areas and certain age cohorts differ in their readiness to emigrate to Germany. While we partly address this in column (1), we might also want to control for the interaction of these indicators as, for example, one might be concerned with local cohorts behaving differently. If these

[52]Column (1) corresponds to the strictest specification of our main results (column (5) in Table 3.3) and is included for reference.

Table 3.4: Robustness – Alternative Fixed Effects

	Emigrated to Germany			
	(1)	(2)	(3)	(4)
War Grievance × Italian Front	0.109*	0.114*	0.178**	0.171**
	(0.058)	(0.063)	(0.073)	(0.081)
Individual Controls	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓
Army Unit FE	✓	✓	✓	✓
Municipality FE	✓	✓	✓	-
Father Municipality FE	-	✓	✓	✓
Father Birth Year FE	-	-	✓	✓
Birth Period × Municipality FE	-	-	-	✓
Observations	2016	1639	1630	1459
Adjusted R^2	0.391	0.419	0.424	0.457
Dependent Variable Mean	0.569	0.532	0.533	0.530

Notes: The table reports coefficients and standard errors from regressions as laid out in Equation (3.1) The dependent variable is the observed emigration behavior of household *i*. The explanatory variable captures whether household *i* holds a war grievance, i.e., the individual or their father experienced a casualty in WWI, and whether the war grievance happened on the Italian front. Individual-level controls are identical to those reported in the main results (Table 3.3). Birth year fixed effects control non-parametrically for opting individual *i*'s birth year. Army Unit fixed effects are dummies for the regiment in which the soldier, i.e., opting individual *i* or *i*'s father, was assigned to during WWI at the time of the casualty. Municipality fixed effects are dummies for the municipality individual *i* lives in. Column (1) reports the estimated coefficients from our preferred specification (Column (5) in Table 3.3). Column (2) includes fixed effects for the municipality of the father of individual *i*. Column (3) includes birth year fixed effects for the father of individual *i*. Column (4) includes a fixed effect for every birth period times municipality cell; these birth periods are constructed by binning 5-year cohorts of birth years. We report heteroskedasticity-robust standard errors. Significance levels: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1.

local cohorts were also differentially impacted by front exposure during WWI, this might bias our results. In column (4), we report results from a regression in which we add fixed effects for the interaction between municipality and binned 5-year birth cohorts. Again, results remain largely unchanged.

In the Appendix, we perform tests addressing two further concerns. We have argued that our results isolate the effect of the direction of a grievance, while holding any other effects of war grievances fixed. However, there are different kinds of war grievances, relating to *who* exactly and *how* someone was affected in WWI. In Table 3.B.4, we show that our results are robust to adding fixed effects for the source of the casualty, i.e., whether the father or the opting individual themselves was affected. Relatedly, controlling for the specific type of casualty, i.e., whether the respective soldier was killed, wounded, or captured, also does not significantly change our results. Furthermore, we show that holding fixed the year in which a casualty happened does

not affect our results. This allows us to compare individuals whose casualty occurred in the same year but on different fronts, thereby addressing the concern that Italy-specific war grievances might capture the effect of casualties happening later in the war. Finally, one might be concerned that the bureaucratic process of the option procedure systematically drives our findings. In Table 3.B.5, we show that our results are unaffected by the inclusion of fixed effects for the specific ADERSt office at which the request was filed and for the ADERSt officer who handled the specific case.

A second set of potential concerns relates to certain subgroups driving our effects, which we address in Table 3.5.[53] First, one might be worried that individuals with a higher exposure to Italians drive our results. For example, if individuals from villages with a higher share of Italians harbor stronger anti-Italian sentiments to begin with, it would be conceivable that they are more likely to self-select into fighting on the Italian front as well as leaving Fascist Italy for Germany in 1940. In column (2), we drop all individuals from municipalities with an above-median share of Italian population as recorded in the pre-WWI census, which does – despite the significant drop in sample size – not change the estimated effect.

A related argument could be made about individuals from border regions exhibiting different behavior. People living closer to the Austrian (after 1938 German) border might per se have stronger connections to Germany, affecting both their likelihood to emigrate as well as to fight on the Italian front. When dropping individuals from municipalities bordering Austria, the reported effect in column (3) remains almost unchanged. Analogously, dropping individuals from municipalities on the border to the rest of Italy, with potentially closer relations with Italians, does not affect our main coefficient of interest as shown in column (4). In column (5), we drop all individuals from border regions, again leaving our results unchanged.

Finally, in column (6) we return to the issue of soldiers drafted from the reserve force. As discussed in Section 3.4.2, these soldiers were only deployed en masse when the Italian front opened in 1915 and differed significantly by age and marital status, as compared to regular soldiers. While the inclusion of army unit fixed effects in our main specification already assured that our results are estimated only from *within* army units, we might still be worried that reserve soldiers drive our results. Nevertheless, when dropping all matched individuals linked to soldiers from the reserve force from our sample, we find, if at all, more positive and significant effects of Italy-specific war grievances on emigration behavior.

[53] Again, column (1) provides the strictest specification from our main findings in Table 3.3 for reference.

Table 3.5: Robustness – Subsamples

	Emigrated to Germany					
	(1) Full Sample	(2) $\leq$ Mdn Italian	(3) No AT Border	(4) No IT Border	(5) No Border	(6) No Reserve
War Grievance $\times$ Italian Front	0.109*	0.136*	0.148**	0.113*	0.137*	0.145**
	(0.058)	(0.077)	(0.071)	(0.061)	(0.071)	(0.061)
Individual Controls	✓	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓	✓
Army Unit FE	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Observations	2016	1020	1641	1871	1553	1983
Adjusted R^2	0.391	0.401	0.432	0.404	0.444	0.396
Dependent Variable Mean	0.569	0.476	0.547	0.570	0.547	0.568

Notes: The table reports coefficients and standard errors from regressions as laid out in Equation (3.1) The dependent variable is observed emigration behavior of household *i*. The explanatory variable captures whether household *i* holds a war grievance, i.e., the individual or their father experienced a casualty in WWI, and whether the war grievance happened on the Italian front. Individual-level controls are identical to those reported in the main results (Table 3.3). Birth year fixed effects control non-parametrically for opting individual *i*'s birth year. Army Unit fixed effects are dummies for the regiment in which the soldier, i.e., opting individual *i* or *i*'s father, was assigned to during WWI at the time of the casualty. Municipality fixed effects are dummies for the municipality individual *i* lives in. Column (1) reports the estimated coefficients from our preferred specification (Column (5) in Table 3.3). Column (2) drops individuals from municipalities with a pre-WWI above-median share of Italian population (1.61%). Column (3) drops individuals from municipalities with a border to the rest of Italy. Column (4) drops individuals from municipalities with a border to Austria. Column (5) drops individuals from all border-municipalities. Column (6) drops individuals whose casualty was reported as part of the reserve force (*Landsturm* or *Standschützen*). We report heteroskedasticity-robust standard errors. Significance levels: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1.

3.6 Conclusion

Nationalism has been a powerful force in history. In particular, feelings of national identity and resentment have the power to mobilize large masses of individuals. In this paper, we have studied a unique case of nationalist mobilization: the mass emigration of almost half of South Tyrol's population to Nazi Germany. Crucially, this event was driven by deep feelings about national identity and cultural belonging. We ask whether grievances from a war long passed can influence individuals' behavior in such episodes of nationalist mass mobilization. We exploit South Tyrol's history as a unique natural experiment to tackle this question. Using exogenous variation in soldiers' front experience during World War I, we can isolate the effect of war grievances directed at Italy, thereby abstracting away from other financial or emotional effects of war grievances. We find that holding such Italy-specific war grievances increases an individual's probability to emigrate to Germany by around 10 to 15 percentage points. We show that this effect is stable in a series of robustness tests.

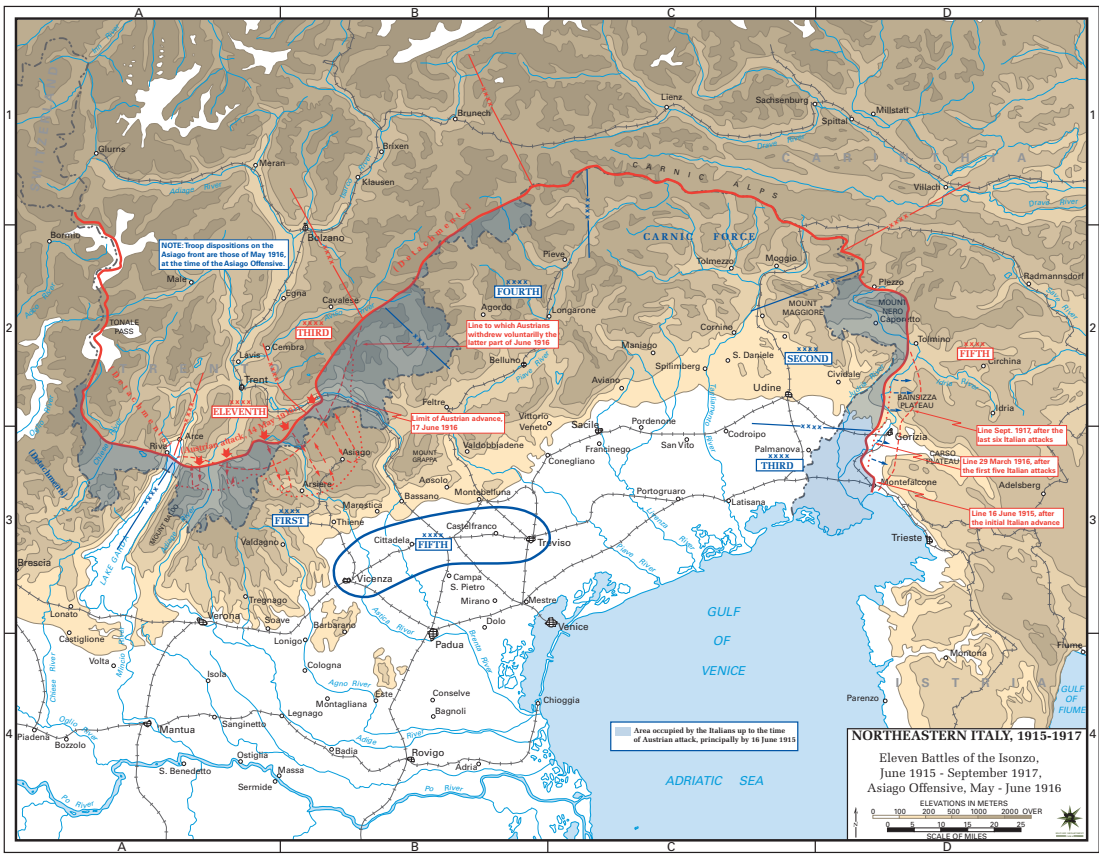
The results of this paper contribute to a growing research agenda examining the roots of nationalism and, more broadly, the determinants of successful nation-building poli-

cies (Rohner and Zhuravskaya 2023). We show how war grievances and feelings about the out-group are powerful and persistent drivers of individual behavior in times of extreme nationalist mass mobilization. Our results highlight the long-lasting negative imprints wars can leave on the identity and behavior of ordinary citizens.

Appendix to Chapter 3

3.A Additional Figures

Figure 3.A.1: Italian Front during World War I



Notes: This map shows the course of the Italian front between June 1915 and September 1917. South Tyrol is located in the northwestern part of the map. Source: https://www.westpoint.edu/academics/academic-departments/history/world-war-one (last accessed on 03/10/2022).

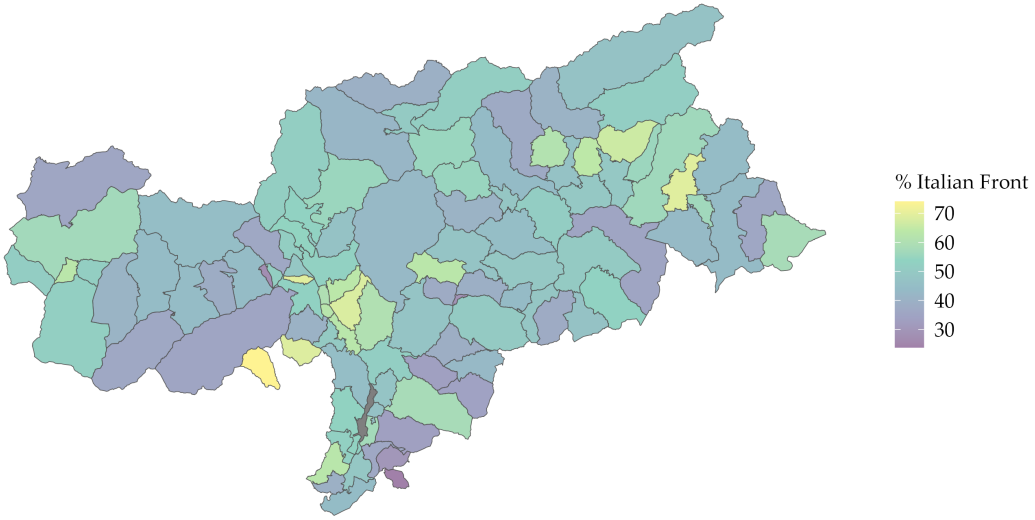
Figure 3.A.2: Propaganda Poster

(a) Anti-Italian WWI Postcard

(b) "Heim ins Reich" Propaganda Poster

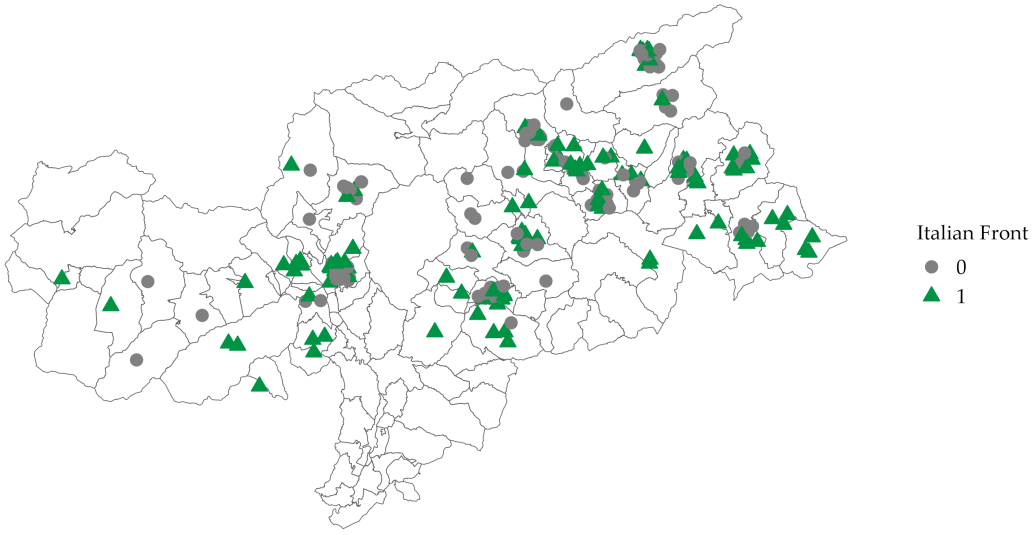
Notes: Panel (A) shows an Austro-Hungarian postcard from World War I with a German, a Bosnian, an Austrian, and a Hungarian soldier (from left to right). The text reads "May God punish unfaithful Italy!" Source: Europeana 1914-1918 (2011). Panel (B) shows a Nazi propaganda poster advertizing the mass relocation of South Tyroleans to Nazi Germany. It emphasizes the pan-German ideology that South Tyroleans were members of the German nation and, thus, by emigrating to Germany were fulfilling their duty to their nation. The German text on the top reads "Greater Germany is calling!", while the bottom explicitly states the *"Heim ins Reich!"* ("Back to the Empire!") parole. Source: Obermair (2021)

Figure 3.A.3: Geographic distribution of Italian front share by municipality



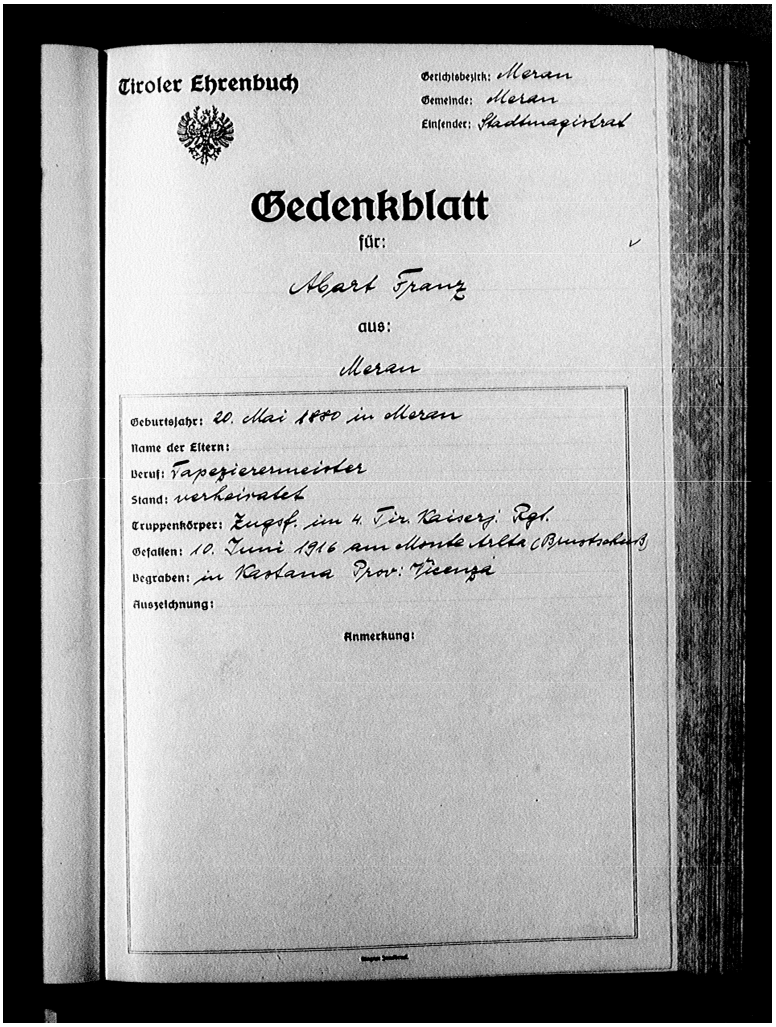
Notes: This figure shows a municipality-level map of South Tyrol as of 1940. The municipalities are shaded according to the relative share of soldiers who died on the Italian Front in WWI among all soldiers recorded in the *Ehrenbücher* originating from this municipality.

Figure 3.A.4: Geographic distribution of Italian front treatment (only matched)



Notes: This figure shows a municipality-level map of South Tyrol as of 1940. Every matched individual is plotted with a dot within the borders of their municipality of residence. Green triangles indicate that the individual is assigned to the "Italian front" treatment group, and grey circles indicate that it was assigned to the "Other fronts" control group.

Figure 3.A.5: Exemplary entry in *Ehrenbücher*



Notes: The entry shows the memorial page for Franz Abart from Meran. He was born on May 20, 1880, in Meran, was married and worked as wallpaperer. He served as platoon leader in the 4th regiment of the *Tiroler Kaiserjäger* ("Imperial Tyrolean Riflemen"). He was killed on June 10, 1916, at Monte Arlta, Italy. Source: Tiroler Landesmuseen (2014)

Figure 3.A.6: Exemplary entry in *Verlustlisten*

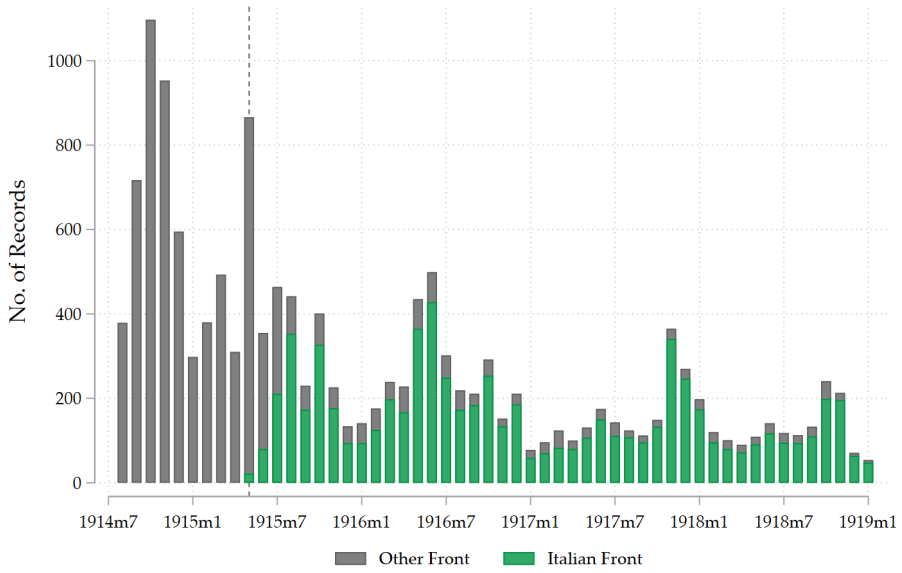

1882; kriegsgef., Nowo Nikolajewsk, Gouvernement Tomsk, Rußland.

Ennemoser Alois, ResJäg., TJR. Nr. 3, 5. Komp., Tirol, Meran, Platt, 1881; kriegsgef., Volterra, Italien.

Erben Josef, ResFahrkan., FHR. Nr. 10, ehem. Nr. 9, Haub-MunK. Nr. 6, Böhmen, Hohenelbe, Ochsengraben, 1886; kriegsgef. Tscherkassy, Gouvernement Kiew, Vereinsspital

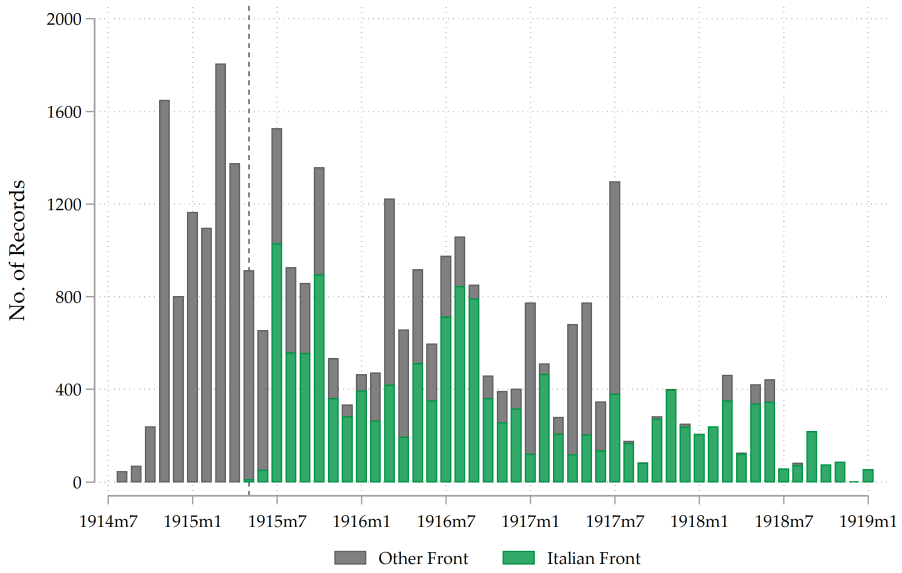
Notes: Entry for Alois Ennemoser recorded in *Verlustliste* No. 397 from May 22, 1916. He originated from the municipality of Platt in the district of Meran in Tyrol and served as *Reservejäger* (reserve rifleman) in Company No. 5 of Regiment No. 3 of *Tiroler Kaiserjäger* ("Imperial Tyrolean Riflemen") and was held in war captivity in Volterra, Italy. Source: Verein für Computergenealogie (2021)

Figure 3.A.7: Distribution by Front Status (*Ehrenbücher*)



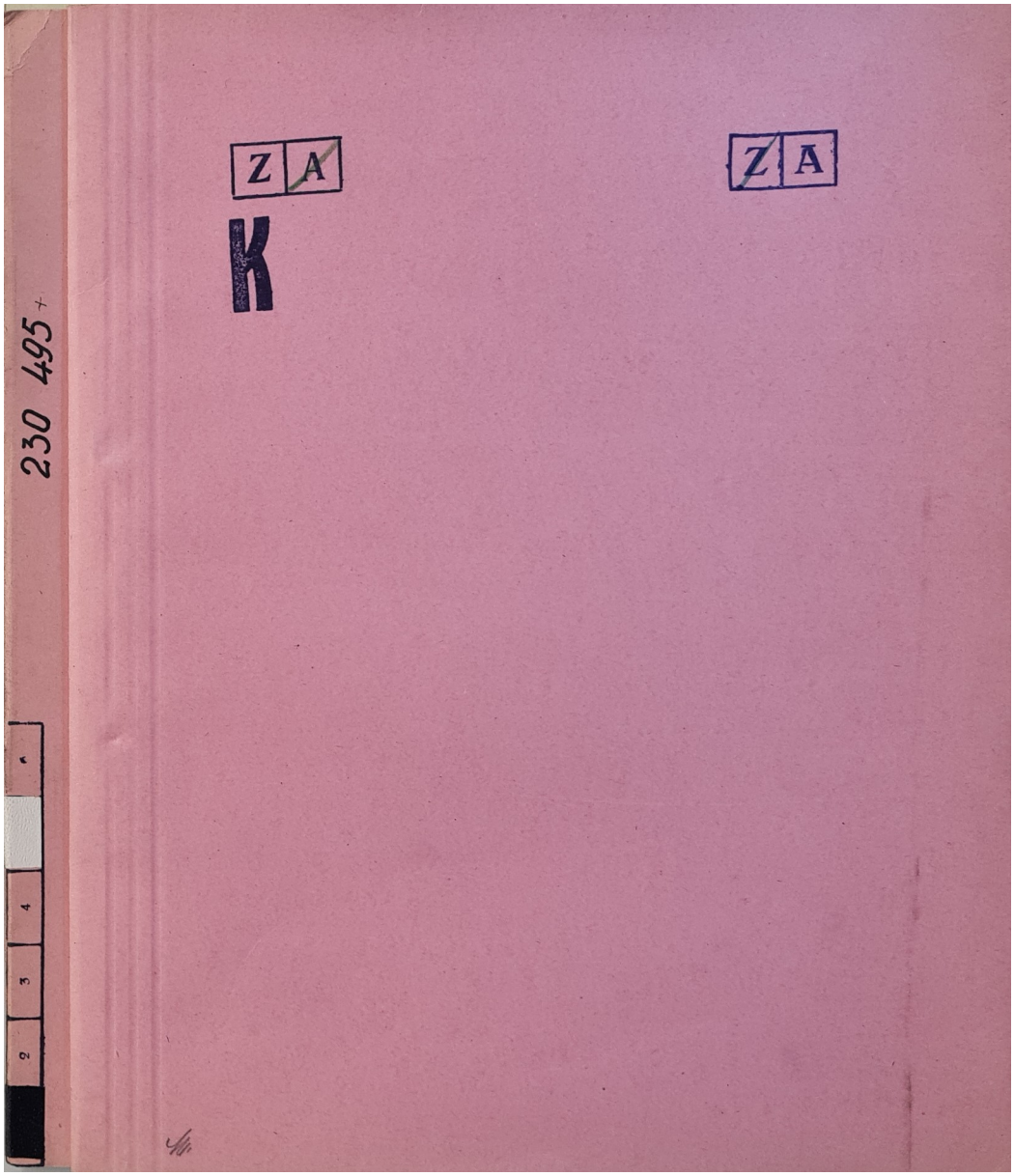
Notes: This figure shows the distribution of entries in the *Ehrenbücher* by the front on which they were recorded. We include all entries for which we could infer the front as follows: (i) explicit mention of place of death and (ii) deployment at the time of death of military unit (regiment) in which soldier was recorded to have served from Glaise von Horstenau (1932). The dashed vertical line indicates the Italian war entry in May 1915.

Figure 3.A.8: Distribution by Front Status (*Verlustlisten*)



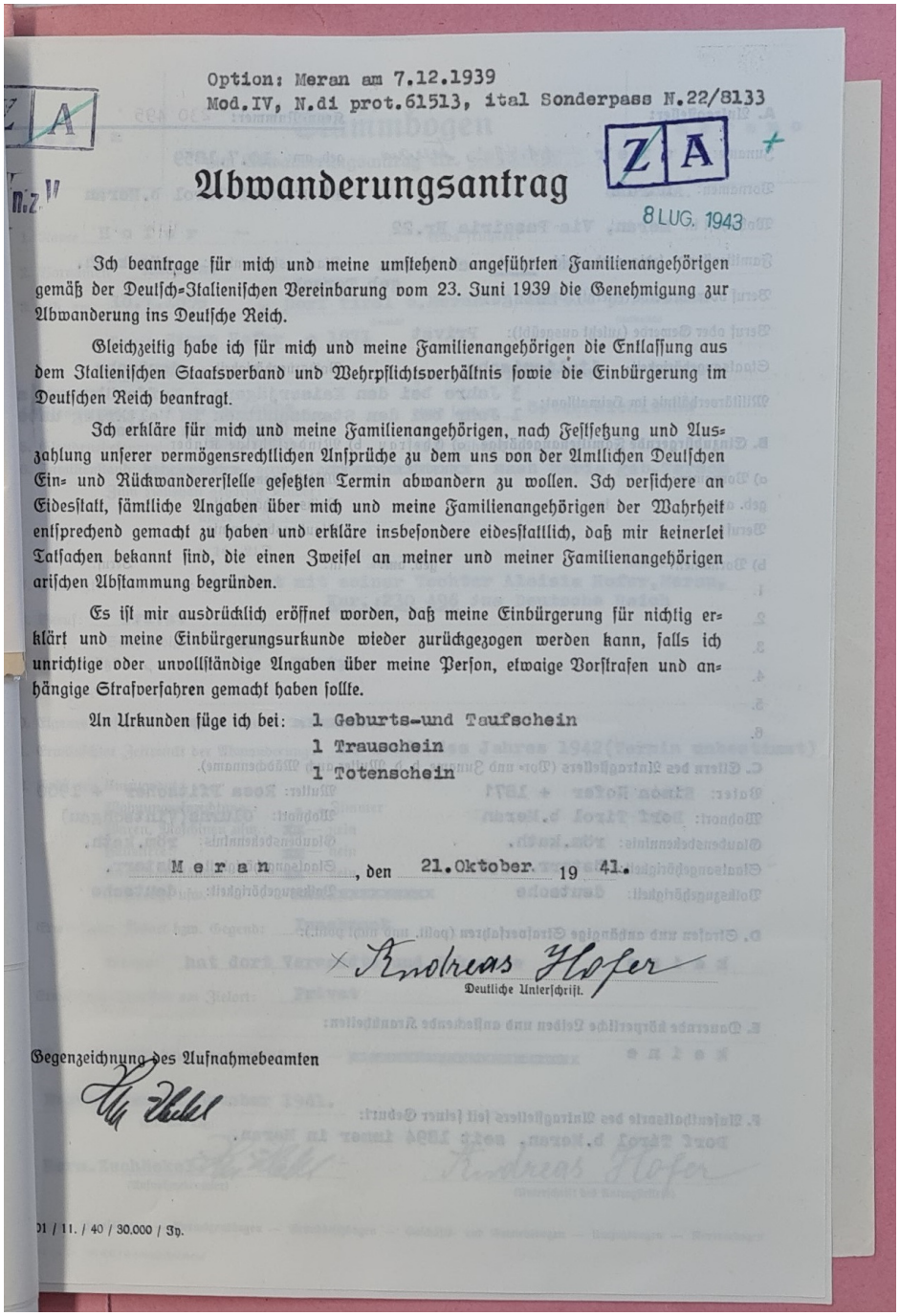
Notes: This figure shows the distribution of entries in the *Verlustlisten* by the front on which they were recorded. We include all entries for which we could infer the front as follows: (i) explicit mention of place of recorded event and (ii) deployment at time of recorded event of military unit (regiment) in which soldier was recorded to have served from Glaise von Horstenau (1932). The dashed vertical line indicates the Italian war entry in May 1915.

Figure 3.A.9: Cover Sheet of Personal Folder



Notes: This figure and the following figures (Figures 3.A.10-3.A.12) show scans of an exemplary option file from our dataset. In particular, this figure shows the cover sheet of the file folder of one option file. Among others, the individual opter's unique case number (here: 230 495) can be seen as well as the stamps indicating their emigration status (here: in upper-right corner the stamp "A" for *"abgewandert"* (emigrated) can be seen. This stamp superseded the stamp "Z" for *"zurückgestellt"* (shelved) that all files received while being processed)

Figure 3.A.10: Emigration Request – Page 1



Option: Meran am 7.12.1939
Mod.IV, N.di prot.61513, ital Sonderpass N.22/8133

Z A

Abwanderungsantrag

Z A

8 LUG. 1943

Ich beantrage für mich und meine umstehend angeführten Familienangehörigen gemäß der Deutsch-Italienischen Vereinbarung vom 23. Juni 1939 die Genehmigung zur Abwanderung ins Deutsche Reich.

Gleichzeitig habe ich für mich und meine Familienangehörigen die Entlassung aus dem Italienischen Staatsverband und Wehrpflichtsverhältnis sowie die Einbürgerung im Deutschen Reich beantragt.

Ich erkläre für mich und meine Familienangehörigen, nach Festsetzung und Auszahlung unserer vermögensrechtlichen Ansprüche zu dem uns von der Amtlichen Deutschen Ein- und Rückwandererstelle gesetzten Termin abwandern zu wollen. Ich versichere an Eidesstatt, sämtliche Angaben über mich und meine Familienangehörigen der Wahrheit entsprechend gemacht zu haben und erkläre insbesondere eidesstattlich, daß mir keinerlei Tatsachen bekannt sind, die einen Zweifel an meiner und meiner Familienangehörigen arischen Abstammung begründen.

Es ist mir ausdrücklich eröffnet worden, daß meine Einbürgerung für nichtig erklärt und meine Einbürgerungsurkunde wieder zurückgezogen werden kann, falls ich unrichtige oder unvollständige Angaben über meine Person, etwaige Vorstrafen und anhängige Strafverfahren gemacht haben sollte.

An Urkunden füge ich bei: 1 Geburts-und Taufschein
1 Trauschein
1 Totenschein

Meran, den 21. Oktober 19 41.

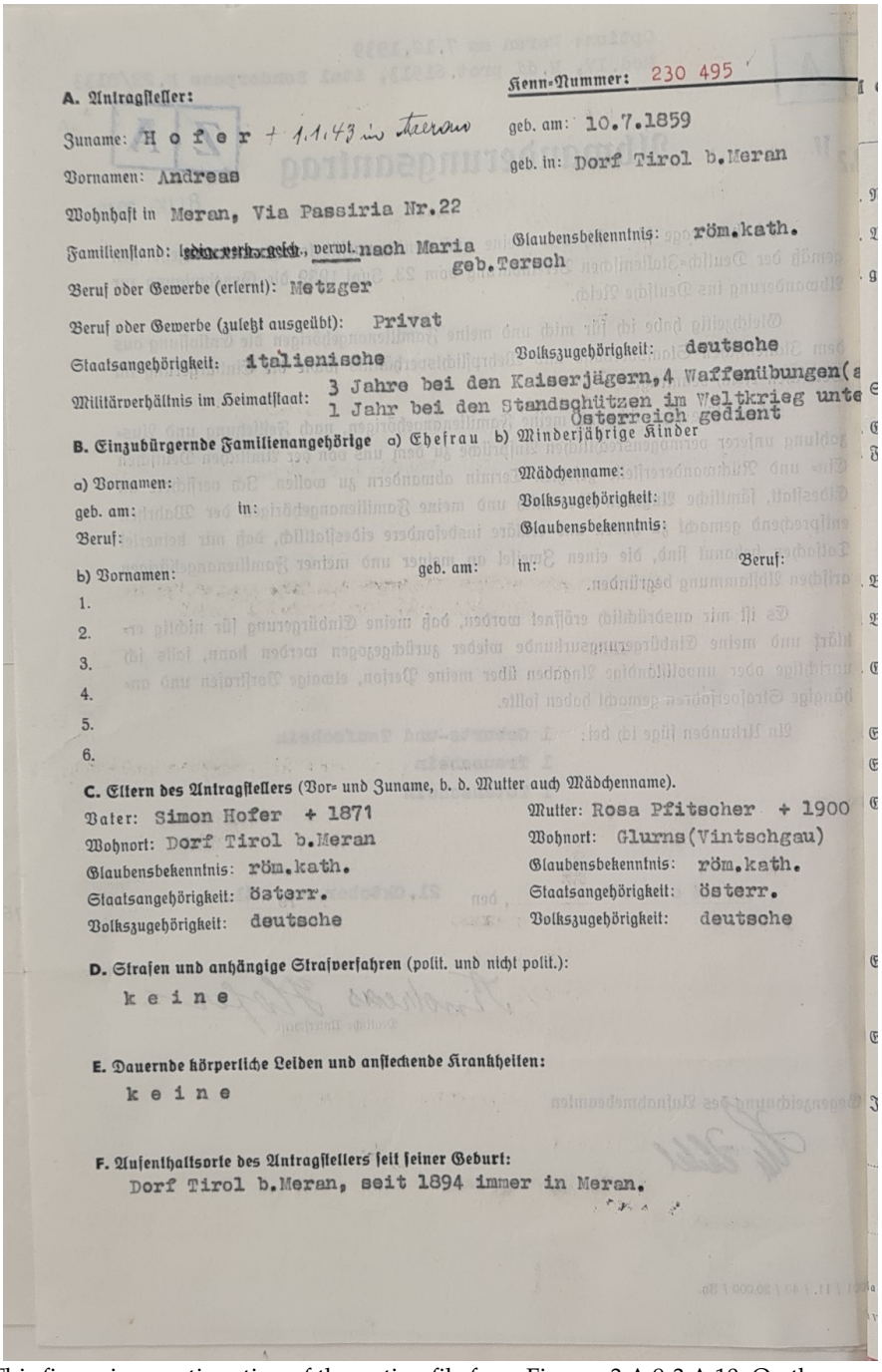
× Andreas Hofer
Deutliche Unterschrift.

Gegenzeichnung des Aufnahmebeamten

01 / 11. / 40 / 30.000 / Sy.

Notes: This figure is a continuation of the option file from Figure 3.A.9. It shows the first page of the *Abwanderungsantrag* (Emigration Request) an individual filed when finally applying for German citizenship and emigration to Germany. On this page the individual declares their intent to leave for Germany. The municipality of the office, the date, and the signatures of the officer and individual are visible. The stamps indicating the emigration status (here: A for *"abgewandert"* (emigrated)) are included again on this page.

Figure 3.A.11: Emigration Request – Page 2



Kenn-Nummer: 230 495

A. Antragsteller:

Zuname: H o f e r + 1.1.43 in Meran — geb. am: 10.7.1859

Vornamen: Andreas — geb. in: Dorf Tirol b.Meran

Wohnhaft in Meran, Via Passiria Nr.22

Familienstand: ~~ledig, verh., geld.~~, verwt. nach Maria geb.Tersch — Glaubensbekenntnis: röm.kath.

Beruf oder Gewerbe (erlernt): Metzger

Beruf oder Gewerbe (zuletzt ausgeübt): Privat

Staatsangehörigkeit: italienische — Volkszugehörigkeit: deutsche

Militärverhältnis im Heimatstaat: 3 Jahre bei den Kaiserjägern,4 Waffenübungen(a
1 Jahr bei den Standschützen im Weltkrieg unte
Österreich gedient

B. Einzubürgernde Familienangehörige a) Ehefrau b) Minderjährige Kinder

a) Vornamen: — Mädchenname:

geb. am: in: — Volkszugehörigkeit:

Beruf: — Glaubensbekenntnis:

b) Vornamen: geb. am: in: Beruf:

1.
2.
3.
4.
5.
6.

C. Eltern des Antragstellers (Vor- und Zuname, b. d. Mutter auch Mädchenname).

Vater: Simon Hofer + 1871	Mutter: Rosa Pfitscher + 1900
Wohnort: Dorf Tirol b.Meran	Wohnort: Glurns(Vintschgau)
Glaubensbekenntnis: röm.kath.	Glaubensbekenntnis: röm.kath.
Staatsangehörigkeit: österr.	Staatsangehörigkeit: österr.
Volkszugehörigkeit: deutsche	Volkszugehörigkeit: deutsche

D. Strafen und anhängige Strafverfahren (polit. und nicht polit.):

k e i n e

E. Dauernde körperliche Leiden und ansteckende Krankheiten:

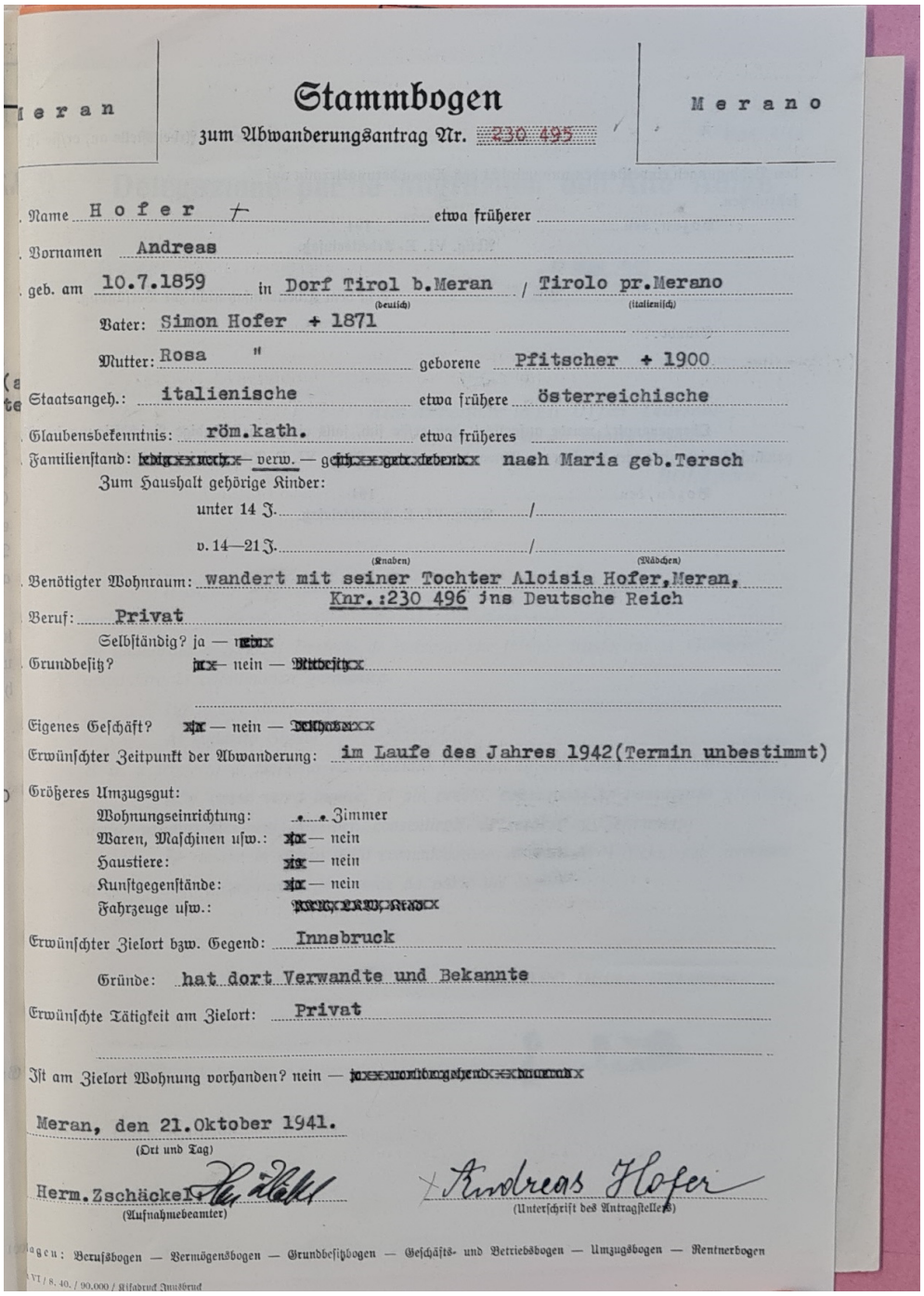
k e i n e

F. Aufenthaltsorte des Antragstellers seit seiner Geburt:

Dorf Tirol b.Meran, seit 1894 immer in Meran.

Notes: This figure is a continuation of the option file from Figures 3.A.9-3.A.10. On the second page of the Emigration Request the main biographic details of the individual are recorded. It includes six sections on (A) the applicant, (B) co-applying family members, (C) parents of the applicant, (D) criminal records, (E) health records, (F) all places of residence since birth.

Figure 3.A.12: Emigration Request – Page 3



Stammbogen
zum Abwanderungsantrag Nr. 230 495

eran — Merano

Name Hofer † etwa früherer
Vornamen Andreas
geb. am 10.7.1859 in Dorf Tirol b.Meran (deutsch) / Tirolo pr.Merano (italienisch)
Vater: Simon Hofer + 1871
Mutter: Rosa " geborene Pfitscher + 1900
Staatsangeh.: italienische etwa frühere österreichische
Glaubensbekenntnis: röm.kath. etwa früheres
Familienstand: ledig — verw. — geschieden nach Maria geb.Tersch
Zum Haushalt gehörige Kinder:
unter 14 J. (Knaben) / (Mädchen)
v. 14—21 J. (Knaben) / (Mädchen)
Benötigter Wohnraum: wandert mit seiner Tochter Aloisia Hofer, Meran, Knr.:230 496 ins Deutsche Reich
Beruf: Privat
Selbständig? ja — nein
Grundbesitz? ja — nein — Mitbesitz
Eigenes Geschäft? ja — nein — Teilhaber
Erwünschter Zeitpunkt der Abwanderung: im Laufe des Jahres 1942(Termin unbestimmt)
Größeres Umzugsgut:
Wohnungseinrichtung: . . Zimmer
Waren, Maschinen usw.: ja — nein
Haustiere: ja — nein
Kunstgegenstände: ja — nein
Fahrzeuge usw.: Kraftwagen, Krad
Erwünschter Zielort bzw. Gegend: Innsbruck
Gründe: hat dort Verwandte und Bekannte
Erwünschte Tätigkeit am Zielort: Privat
Ist am Zielort Wohnung vorhanden? nein — ja, vorübergehend, ständig

Meran, den 21.Oktober 1941.
(Ort und Tag)

Herm.Zschäckel
(Aufnahmebeamter)

† Andreas Hofer
(Unterschrift des Antragstellers)

Sagen: Berufsbogen — Vermögensbogen — Grundbesitzbogen — Geschäfts- und Betriebsbogen — Umzugsbogen — Rentnerbogen

VI / 8. 40. / 90.000 / Hitadruck Innsbruck

This figure is a continuation of the option file from Figures 3.A.9-3.A.11. On the third and final page to be filled in, information on the applicant's financial situation, family status, employment status are listed and further solicitations regarding the emigraton are recorded. It finishes with the date and place of the application and the name and signature of the officer in charge of the file and the applicant.

3.B Additional Tables

Table 3.B.1: Summary Statistics

	Mean	Median	Std. Dev.	Min.	Max.	Obs.
PANEL A: *Individual-level Variables*						
Emigrated to Germany	0.57	1.00	0.49	0.00	1.00	2,338
War Grievance	0.10	0.00	0.31	0.00	1.00	2,338
Female	0.32	0.00	0.47	0.00	1.00	2,337
Age	38.72	34.00	16.79	13.00	88.00	2,320
Military Service	0.44	0.00	0.50	0.00	1.00	2,318
Illness	0.14	0.00	0.35	0.00	1.00	2,315
Police Record	0.06	0.00	0.24	0.00	1.00	2,315
Previously Germany	0.13	0.00	0.34	0.00	1.00	2,338
Children	0.27	0.00	0.44	0.00	1.00	2,337
Married	0.26	0.00	0.44	0.00	1.00	2,332
Owns Property	0.26	0.00	0.44	0.00	1.00	2,157
Out of Labor Market	0.14	0.00	0.35	0.00	1.00	2,257
Works in Learnt Job	0.74	1.00	0.44	0.00	1.00	2,294
Farmer	0.41	0.00	0.49	0.00	1.00	2,338
Skilled Occupation	0.31	0.00	0.46	0.00	1.00	2,338
PANEL B: *Municipal-level Variables*						
Total Population	5274.23	2461.00	6062.82	152.00	23513.00	2,313
Population Density (per ha)	9.33	0.67	51.74	0.03	408.88	1,946
Share Male (%)	48.97	49.20	2.17	45.58	57.11	2,313
Share Italian (%)	1.95	1.61	2.65	0.00	24.59	2,313
Share German (%)	91.96	96.63	13.76	0.30	100.03	2,313
Share Ladin (%)	1.55	0.00	12.12	0.00	99.70	2,313
Share Taxable Land (%)	90.11	95.89	13.65	12.50	99.19	1,946
No. Factories	0.03	0.00	0.25	0.00	3.00	2,313
Cattle per capita	0.28	0.22	0.24	0.00	1.10	2,313
PANEL C: *Within Matched Individuals*						
Italian Front	0.53	1.00	0.50	0.00	1.00	242
Matched Father	0.86	1.00	0.35	0.00	1.00	242
Matched Opter	0.14	0.00	0.35	0.00	1.00	242
Died	0.52	1.00	0.50	0.00	1.00	242
Wounded	0.29	0.00	0.46	0.00	1.00	242
Captured	0.19	0.00	0.39	0.00	1.00	242

Notes: This table reports summary statistics of the covariates used in our analysis. Panel A reports summary statistics on the individual-level (household-level) variables drawn from the emigration requests. Panel B reports summary statistics on the municipalities of individuals in the option request data, where the municipality-level data is drawn from the 1900 census. Panel C reports variables conditional on being matched to records in the World War I casualty lists.

Table 3.B.2: Determinants of Emigration

	Emigrated to Germany		
	(1)	(2)	(3)
Female	-0.206***	-0.177***	-0.089***
	(0.030)	(0.028)	(0.024)
Age	-0.024***	-0.024***	-0.014***
	(0.004)	(0.004)	(0.003)
Age^2	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Military Service	0.008	-0.001	0.003
	(0.025)	(0.024)	(0.022)
Illness	-0.103***	-0.061**	-0.025
	(0.030)	(0.029)	(0.026)
Police Record	0.110***	0.101**	0.064**
	(0.041)	(0.040)	(0.032)
Previously Germany	0.126***	0.096***	0.041
	(0.031)	(0.030)	(0.027)
Children	-0.007	0.011	0.031
	(0.031)	(0.028)	(0.024)
Married	-0.013	-0.013	-0.009
	(0.034)	(0.031)	(0.028)
Owns Property	-0.277***	-0.221***	-0.114***
	(0.027)	(0.025)	(0.024)
Out of Labor Market	-0.088**	-0.056	-0.002
	(0.039)	(0.037)	(0.035)
Works in Learnt Job	-0.020	-0.024	-0.030
	(0.024)	(0.023)	(0.020)
Farmer	-0.286***	-0.190***	-0.095***
	(0.030)	(0.031)	(0.028)
Skilled Occupation	-0.048*	-0.021	-0.018
	(0.028)	(0.027)	(0.024)
Municipality FE	-	✓	✓
ADERSt Officer FE	-	-	✓
Observations	2038	2006	1934
Dependent Variable Mean	0.571	0.568	0.568
R^2	0.242	0.403	0.557

Notes: This table reports results from a regression of observed emigration behavior on the set of control variables we use in our main regression shown in Equation (3.1). The only difference concerns the age variable: while the main results tables (e.g., Table 3.3) control for age non-parametrically, this regression imposes a functional form on age. It includes age linearly and as a squared term. Column (1) reports estimates from a regression on these covariates only. Column (2) controls for municipality fixed effects. Column (3) control for ADERSt officer fixed effects. For a more detailed discussion, see the table notes in Table 3.B.5. Significance levels: $***$ $p<0.01$, $**$ $p<0.05$, and $*$ $p<0.1$.

Table 3.B.3: Balancing of Soldiers by Front (excluding Reserve Force)

	Other Front	Italian Front	Diff.	P-val.	Obs.
Individual characteristics					
Birth Year	1886.55	1888.38	-1.83***	(0.00)	4819
Farmer	0.48	0.48	-0.00	(0.98)	4672
Skilled Occupation	0.28	0.28	-0.00	(0.72)	4672
Married	0.21	0.20	0.01	(0.38)	4484
Municipal characteristics					
Total Population	4135.48	4545.72	-410.24**	(0.01)	5148
Population Density (per ha)	27.30	29.62	-2.32	(0.39)	4984
Share Male (%)	50.15	50.04	0.12**	(0.02)	5148
Share Italian (%)	3.40	3.78	-0.37*	(0.09)	5148
Share German (%)	88.52	89.55	-1.03	(0.10)	5148
Share Ladin (%)	5.92	4.20	1.72***	(0.00)	5148
Share Taxable Land (%)	0.85	0.86	-0.01	(0.16)	4984
No. Factories	0.14	0.15	-0.01	(0.46)	5148
Cattle per capita	0.35	0.34	0.01	(0.26)	5148

Notes: This table reports results from two-group mean-comparison tests by treatment status, i.e., whether a war grievance happened on Italian or another front. Columns report group averages, the difference between averages, p-values for mean equality, and the number of observations. The sample includes all dead soldiers from *Ehrenbücher* that originate from municipalities in South Tyrol and for which we were able to locate their death on a specific front. Soldiers that served in reserve force regiments (*Landsturm* and *Standschützen*) are excluded from the sample. Significance levels: $***$ $p<0.01$, $**$ $p<0.05$, and $*$ $p<0.1$.

Table 3.B.4: Robustness – Types of War Grievances

	Emigrated to Germany				
	(1)	(2)	(3)	(4)	(5)
War Grievance × Italian Front	0.109*	0.114**	0.107*	0.130*	0.143*
	(0.058)	(0.058)	(0.062)	(0.071)	(0.082)
Individual Controls	✓	✓	✓	✓	✓
Birth Year FE	✓	✓	✓	✓	✓
Army Unit FE	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓
Casualty Source FE	-	✓	-	-	✓
Casualty Type FE	-	-	✓	-	✓
Casualty Year FE	-	-	-	✓	✓
Observations	2016	2016	2016	2016	2016
Adjusted R^2	0.391	0.392	0.390	0.392	0.392
Dependent Variable Mean	0.569	0.569	0.569	0.569	0.569

Notes: The table reports estimates of Equation (3.1). The dependent variable is the observed emigration behavior of household *i*. The explanatory variable for which a coefficient is reported captures whether household *i* holds an Italy-specific war grievance. Individual-level controls are identical to those in the main results (Table 3.3). Birth year fixed effects control non-parametrically for individual *i*'s birth year. Army Unit fixed effects capture the regiment in which the soldier, i.e., *i* or *i*'s father, was fighting in. Municipality fixed effects are dummies for the municipality individual *i* lives in. Column (1) reports the estimated coefficients from our preferred specification (Column (5) in Table 3.3). Column (2) includes fixed effects for the source of *i*'s war grievance, i.e., whether casualty occurred to *i* itself or $i's$ father. Column (3) includes fixed effects for the type of *i*'s war grievance, i.e., whether casualty refers to the respective person having died, been wounded, or been captured. Column (4) includes fixed effects for the year when *i*'s war grievance occurred, i.e., in which year during World War I the casualty was recorded. Column (5) includes the three aforementioned types of fixed effects simultaneously. We report heteroskedasticity-robust standard errors. Significance levels: $***$ $p<0.01$, $**$ $p<0.05$, and $*$ $p<0.1$.

Table 3.B.5: Robustness – ADERSt Officer & Offices

	Emigrated to Germany		
	(1)	(2)	(3)
War Grievance × Italian Front	0.109*	0.106*	0.093*
	(0.058)	(0.059)	(0.048)
Individual Controls	✓	✓	✓
Birth Year FE	✓	✓	✓
Army Unit FE	✓	✓	✓
Municipality FE	✓	✓	✓
ADERSt Office FE	-	✓	✓
ADERSt Officer FE	-	-	✓
Observations	2016	2013	1942
Adjusted R^2	0.391	0.393	0.536
Dependent Variable Mean	0.569	0.569	0.570

Notes: The table reports estimates of Equation (3.1). The dependent variable is observed emigration behavior of household *i*. The explanatory variable for which a coefficient is reported captures whether household *i* holds an Italy-specifc war grievance. Individual-level controls are identical to those in the main results (Table 3.3). Birth year fixed effects control non-parametrically for individual *i*'s birth year. Army Unit fixed effects capture the regiment in which the soldier, i.e., *i* or *i*'s father, was fighting in. Municipality fixed effects are dummies for the municipality individual *i* lives in. Column (1) reports the estimated coefficients from our preferred specification (Column (5) in Table 3.3), and is included for reference. Column (2) includes fixed effects for the ADERSt office handling *i*'s case. Column (3) includes fixed effects for the ADERSt officer in charge of *i*'s case. We report heteroskedasticity-robust standard errors. Significance levels: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1.

3.C Linking Procedure

There are three types of potential matches between individuals in our emigration requests and soldiers in the WWI casualty lists:

1. *Option Files (Father)* $\leftrightarrow$ *Ehrenbücher (Father)*: We might expect to identify the father of an individual filing the option request among the soldiers recorded in the *Ehrenbücher* (all dead)

2. *Option Files (Father)* $\leftrightarrow$ *Verlustlisten (Father)*: We might expect to identify the father of an individual filing the option request among the soldiers reported in the *Verlustlisten* (dead, wounded, or captured)

3. *Option Files (Opter)* $\leftrightarrow$ *Verlustlisten (Opter)*: We might expect to identify the individual filing the option request itself among the potentially surviving soldiers reported in the *Verlustlisten* (wounded or captured)

For each of these potential linkages, we proceed in the following steps:

(i) We run a first round of fuzzy string matching by firstname and lastname using the command `reclink` in Stata (Blasnik 2010) with a minimum similarity score of 0.9 to identify potential misspellings of names which we manually correct.

(ii) We perform a perfect string match with the command `merge` using the last name, first name, home town, and – only in case (3) – the birth year of an individual. We exclude all observations from the option files that have at least one missing entry in one of these variables. There are three potential definitions of a home town in our option file dataset that correspond to the definition of home town (*Heimatort*) used in the casualty lists:[54] the birthplace of the opter, the place of residence of the opter, or the (last) place of residence of the father. We sequentially perform three rounds of perfect string matches using different operationalizations of home town, removing the already matched opters after each round and thereby explicitly enforcing a matching hierarchy. In the case of looking up fathers in the casualty lists (cases (1) and (2) above), the hierarchy is as follows: (i) place of residence of father, (ii) birthplace of opter, (iii) place of residence of opter. In the case of looking up opters themselves (case (3) above), the hierarchy is instead: (i) birthplace of opter, (ii) place of residence of opter, (iii) place of residence of father. In order to account for the various administrative

[54] In the Austro-Hungarian Empire, the notion of *Heimatort* (home town) or *Heimatberechtigung* (right of domicile) refers to the municipality to which an individual is legally affiliated to. This does not necessarily equal the place of birth but refers to the locality to which the individual or its family traditionally belonged (von Hirschhausen 2009).

changes in municipalities between World War I and the time of the Option before and during World War I, we manually link all mentioned municipalities in our three datasets to their corresponding municipality as of 1940 using information from *Storia dei Comuni* (see Section 3.3.1).[55]

(iii) After each match, we perform a number of checks: in case of matches to fathers ((1) and (2)), where we do not have information on the birth year in the opter file, we perform a sanity check in terms of age difference and only keep matched pairs where the difference between father and opter child is at least 16 years. After each match we manually compare, if arising, duplicates in matching pairs, i.e., multiple entries from casualty lists matched to the same opter, and manually decide which one is likely the correct match based on the other observable characteristics. In the end, only unique combinations of opter and entry in the casualty list remain. We save these matching pairs in separate files.

(iv) We combine all resulting separate files into one file containing all matched combinations between individual opters/their fathers and their respective counterpart in the casualty list.

Table 3.C.1: Matching Results

Option requests	**2,338**
father known	*1,985*
Matched Father	**210**
dead	126
wounded	49
captured	33
Matched Opter	**35**
wounded	23
captured	12
Total Matches	**244**

Notes: This table reports the number of matches by source and type of the casualty. One individual was matched both via opter itself (wounded) and via its father (dead), such that entries do add up to more than 244.

Finally, we add the information from matched casualty list entries back to the main opter database, separately for each linking combination. This allows us to assess the final number of 244 opters that could successfully be linked to an entry in at least one casualty list. As shown in Table 3.C.2, we see that there are 33 cases where we locate the father of an individual opter in both *Ehrenbücher* and *Verlustlisten*. There is also 1

[55]The website can be accessed via http://www.elesh.it/storiacomuni (last accessed on 06/02/2022).

case where both the opter himself was found in the *Verlustlisten* as well as his killed father in the *Ehrenbücher*.

Table 3.C.2: Datasource of Matches

Source	**Cases**
OPT (Father) ↔ VL (Father)	101
OPT (Father) ↔ EB (Father)	75
OPT (Opter) ↔ VL (Opter)	34
OPT (Father) ↔ EB (Father) + VL (Father)	33
OPT (Opter) ↔ EB (Father) + VL (Opter)	1
Total	**244**

Notes: Abbreviations for datasource as follows: OPT = Option Files, EB = Ehrenbücher, VL = Verlustlisten.

CHAPTER 4

Loyalty and Identity: Evidence from the Habsburg Army in World War I

"A country of nationalities cannot wage war without danger to itself."

Kasimir Felix Badeni*

4.1 Introduction

Wars pose great challenges to states and societies, requiring vast amounts of economic, political, and human resources. Yet, "guns are not enough to win wars; one also needs motivated soldiers" (Alesina, Reich, and Riboni 2020, p. 382). The extent to which states can motivate their citizens to fight for their country, asking for the ultimate sacrifice of their own lives, crucially depends on their loyalty. This holds particularly true for societies that are characterized by large degrees of ethnic diversity. Does loyalty and the willingness to fight in a war differ by a soldier's ethnic identity?

I shed light on this question by examining the combat motivation of ethnic minority soldiers fighting for Austria-Hungary during World War I.[2] The Habsburg Empire provides a unique setting to study the relation between ethnic identity and loyalty during war: its armed forces strongly reflected the multi-national character of the Empire, with soldiers from 11 ethnic groups serving side-by-side in non-segregated mixed units. In this paper, I will focus on the specific case of soldiers from the Austrian region of Tyrol, which offers two important advantages: first, the population of Tyrol was characterized by a dichotomous split into a German-speaking majority, the dominant ethnic group of the Empire, and a large Italian-speaking minority. Second, Italy unexpectedly entered the war against Austria-Hungary in 1915, providing the opportunity to analyze how this affected the exerted war effort by Italian minority soldiers.

*Minister-President of Austria (1895-1897), cited from Lyall (2020, p. 253).

[2]The Austro-Hungarian Empire was a state in central Europe ruled by the Habsburg dynasty. In accordance with historical terminology, I will interchangeably use the terms Austria-Hungary, Austro-Hungarian Empire, and Habsburg Empire.

For my empirical analysis, I use novel individual-level data from the universe of wartime casualty records for Tyrolean soldiers in the Austro-Hungarian Army. Apart from reporting biographic information, the casualty lists differentiate between a soldier being killed, wounded, or captured by the enemy. I complement this dataset with information on soldiers' municipality of origin as well as the military unit in which they served. Importantly, the casualty lists do not include any data on the ethnicity or language spoken by a soldier. I, therefore, implement a novel measure based on language classification from natural language processing to infer from the name of a soldier whether he was likely German or Italian.

I find robust evidence for a differential war effort exerted by Italian minority soldiers relative to those of the German majority. I begin by establishing significant differences in casualty outcomes on an aggregate level: Italian-majority municipalities observe less dead and wounded soldiers per capita, whereas relatively more soldiers end up in captivity. I confirm these results in my main individual-level analysis: Italian soldiers are relatively more likely to end up in captivity than to get killed or wounded, compared to their German comrades-in-arms. This differential is not sensitive to controlling for selection based on the place of origin or sorting into military units being exposed to different battle intensities. The results are robust to using alternative measures of soldiers' ethnicity as well as the exclusion of Ladin-speaking areas where the prediction algorithm does not perform well.

In a final set of results, I investigate heterogeneities in the observed differential effects on war outcomes. Two sets of factors seem particularly relevant: first, local ethnic composition matters, with Italian soldiers from homogeneous and less diverse municipalities being more likely to end up in captivity. Second, Italy's entry to World War I left a decisive impact on Italian soldiers' behavior. I only observe a difference in casualty outcomes after the Italian war declaration and only for Italian soldiers fighting on the front against Italy. Although I cannot fully rule out alternative explanations relating to discriminatory treatment of minority soldiers at the front, these results are in line with an interpretation of increasing disloyalty due to mistrust among ethnic groups and the alienation of Italian soldiers (Alesina, Reich, and Riboni 2020; Lyall 2020).

My paper speaks to a number of research agendas in economics, political science, and history. First, I add to an active literature investigating why individuals exert effort in wars. In principle, one would not expect rational agents to risk their lives in war due to the potentially enormous costs and free-riding problems (Alesina, Reich, and Riboni 2020, p. 391). Nevertheless, history has shown repeatedly how soldiers willingly incur hardship, wounds, and even death for their side. Existing work has highlighted the role of peer effects (Ager et al. 2022), leadership (Dippel and Ferrara 2022), threat

or use of punishment (Chen 2017; Lyall and Zhukov 2021), religious beliefs (Beatton, Skali, and Torgler 2019), state repression (Rozenas, Talibova, and Zhukov 2022), and public goods provision (Caprettini and Voth 2022) for the effort exerted by soldiers. A number of studies have explicitly studied the role of diversity and discrimination in this regard: in their seminal work, Costa and Kahn (2003, 2008) show how soldiers serving in more homogeneous units were less likely to shirk in the American Civil War. Conversely, Indacochea (2019) finds that increases in racial integration in the U.S. Army improved survival rates during the Korean War. Closest to my case, Talibova (2021) shows how Russian WWI veterans from ethnic minority groups were more likely to join the revolutionary Red Army in the ensuing Russian Civil War.[3] My paper adds to this literature in two ways: first, by introducing the case of the Habsburg Empire with its exceptionally large ethnic diversity and, second, by examining a situation where ethnic minority soldiers found themselves fighting against their national "mother state".

Second, my work builds on existing theoretical work on the interplay between identity, nationalism, and war. Alesina, Reich, and Riboni (2020) recently provided a theoretical framework on how states resort to nationalism in order to motivate soldiers during war. My historical setting allows me to empirically test their hypothesis that ethnic minority soldiers might put less effort into fighting due to feeling excluded by the regime's negative nation-building.[4] The importance of group status and interethnic distance also plays a crucial role in models of endogenous (national) identity formation (Shayo 2009, 2020). In a wider sense, I also speak to theoretical work on the salience of ethnic conflict (Esteban and Ray 2008, 2011).

Finally, this paper makes a thematic contribution by studying the role of nationalization in the Austro-Hungarian Empire during World War I. A traditional view has held that the Habsburg Empire was a "prison of peoples" (Wandruszka and Urbanitsch 1980, p. XVI) that collapsed due to the spread of nationalism among its ethnic groups; a process that was significantly catalyzed by World War I. Although recent historiography has somewhat challenged this perspective, Cole (2018, p. 117) notes that there continues to be a lack of empirical research on the role of World War I for national identification. My paper contributes to this research gap by providing, to the best of my knowledge, the first quantitative analysis of individual-level behavior among minority soldiers in the Habsburg Army. More generally, this study sheds light on the promising setting of the Austro-Hungarian Empire to study questions of nationalism

[3] Relatedly, other papers have highlighted the negative consequences of racial discrimination (Qian and Tabellini 2022) and assimilation policies (Fouka 2020) for volunteering in the U.S. Army during World War II.

[4] In fact, Alesina, Reich, and Riboni (2020, p. 403) explicitly mention (South) Tyrol as a promising case for future research on the role of nation-building and war in border regions.

and identity, that has so far received scarce attention among economic historians.[5]

This paper proceeds as follows: Section 4.2 provides historical background on Italians in the Habsburg Empire serving as soldiers during World War I. Building on existing theoretical frameworks, I derive concrete hypothesis on the combat motivation of Italian soldiers. In Section 4.3, I introduce the WWI casualty data and describe the language classification approach to predict soldiers' ethnicity. Section 4.4 presents the main municipal- and individual-level results, whereas Section 4.5 shows heterogeneities in the observed differences in war effort. Section 4.6 concludes.

4.2 Background

4.2.1 Historical Setting

Italians in the Habsburg Empire

On the eve of World War I, the Habsburg-ruled Austro-Hungarian Empire was one of Europe's major powers, being the second-largest geographic entity and third-most populous country.[6] The Empire's presumably most distinctive characteristic was its large ethnic diversity, making it the prototype of a multi-national state (*Vielvölkerstaat*) of its time. Crucially, only less than half of the population belonged to the two dominant ethnic groups of Germans (23.4%) and Hungarians (19.8%).[7] The remaining majority of the population belonged to several non-core ethnic groups: Czechs (12.5%), Serbs, Croats, and Bosnians (10.9%), Poles (9.7%), Ukrainians (denominated Ruthenians, 7.8%), Romanians (6.3%), Slovaks (3.9%), Slovenes (2.4%), Italians (1.5%), and other ethnicities (2.1%).[8]

The approximately 800,000 Italian-speakers formed the smallest officially recognized ethnic minority in Austria-Hungary.[9] They predominantly lived in two areas: in the so-called Austrian Littoral (*Österreichisches Küstenland*) on the Adriatic coast, more

[5] Existing work has studied the role of institutions (Becker et al. 2016), economic development (Komlos 1989), emigration (Jagadits 2021), trade (Schulze and Wolf 2008, 2012), and economic inequality (Erfurth 2022) in the Habsburg Empire.

[6] The Empire's territory spanned 676,000 km^2 encompassing 51.4 million inhabitants, making it only second in area to the Russian Empire and third in population to Russia and the German Empire (k. k. Statistische Zentralkommission 1912, p. 38).

[7] The Austro-Hungarian Compromise (*Ausgleich*) of 1867 had established the dual monarchy of Austria-Hungary, with the German-dominated Austrian Empire (*Cisleithania*) and the formally independent Hungarian-dominated Kingdom of Hungary (*Transleithania*), combined under the common rule of the Habsburg Emperor.

[8] These numbers refer to the results of the 1910 census as aggregated by Taylor (1990, p. 286). Note that the Habsburg institutions determined affiliation to an ethnic group or nation by language use (Stergar and Scheer 2018). Figure 4.A.1 in the Appendix shows the geographic distribution of ethnic groups across the Empire as of 1910.

[9] Until the mid-19th century more than 5 million Italians (ca. 15%) had lived in the Austrian Empire. After Austrian defeat in the Second and Third Italian War of Independence, Lombardy (1859) and Venetia (1866) fell to the newly formed Kingdom of Italy.

than 350,000 Italian-speakers lived in the city of Trieste and the counties of Gorizia-Gradisca and Istria. About 380,000 Italians lived in the southern part of the crown land of Tyrol – called *Trentino* by Italians after its capital *Trient/Trento* or Italian Tyrol (*Welschtirol*) by Germans[10] – in the southwestern corner of the monarchy.[11] The empirical analysis in this paper will focus on Italian-speaking soldiers from Tyrol and, therefore, be the focus of the remainder of this section.

The ethno-linguistic composition of Tyrol represented a somewhat special case within the Empire. While in many regions the population was also very diverse on a local level, there existed a relatively clear language border separating the German and Italian population in Tyrol. Figure 4.1a visualizes the geographic distribution of the Tyrolean population. In total, the 1910 census recorded 57.3% Germans and 42.1% Italians among the 916,261 inhabitants (k. k. Statistische Zentralkommission 1912, p. 59).[12] As the Italian-speaking Trentino had for centuries formed part of Tyrol and the Habsburg-ruled territories, Di Michele (2020, pp. 23–24) highlights that the local population often harbored a historically grown attachment to the Empire and a distinct regional identity that did not necessarily comply with the concept of a unified Italian nation. Nevertheless, in line with general tendencies across the Empire, the immediate pre-war period saw an intensification of nationalized disputes over education and autonomy (Di Michele 2020, pp. 33–36). By contrast, many Austrians held deep-seated resentments against Italians, not least since the territorial losses to unified Italy and an increasing geopolitical rivalry.[13]

The Austro-Hungarian Army

The Austro-Hungarian Army strongly reflected the multi-ethnic nature of the Empire's population.[14] Consequently, "diversity management" (Überegger 2018, p. 16) became one of the central tenants of Habsburg military policy. In order to prevent separatist tendencies, army units were explicitly recruited to be ethnically mixed (Wandruszka and Urbanitsch 1987, p. 97). This was also reflected in the daily military service: while German was the official language of command, if a regiment was comprised of at least 20% soldiers from one ethnic group, they were entitled to use their

[10] As noted by Di Michele (2020, p. 28), the name of the region itself formed part of national disputes, with Austrian authorities never accepting the denomination *Trentino*.

[11] The remaining Italian-speakers lived in the Hungarian-ruled city of *Rijeka/Fiume* (ca. 25,000) and in Dalmatia (ca. 18,000). All population figures are retrieved from Corsini (1980).

[12] Moreover, an estimated 28,000 Ladin speakers lived in Tyrol, an ethnic minority of Rhaeto-Romance descent, that were, however, jointly recorded with Italian speakers in the censuses (Bihl 1980).

[13] Referring to Franco-German relations, some observers have even coined a "hereditary enmity" between Austria and Italy (Gatterer 1972).

[14] Technically speaking, the Austro-Hungarian Army consisted of three separate branches: the Empire-wide *k.u.k.* (*kaiserlich und königlich*, "imperial and royal") Common Army, the Austrian *k.k.* (*kaiserlich-königlich*, imperial-royal) *Landwehr*, and the Hungarian *k.u.* (*königlich-ungarisch*, royal-hungarian) *Honvéd*. Furthermore, there existed a separate common Navy (*k.u.k. Kriegsmarine)*.

native tongue as a regimental language.[15] Likewise, officers were obliged to learn all regimental languages of the units they commanded (Scheer 2020). While the ethnic composition of troops was largely representative of the general population, the officer corps was much less diverse with a strong overrepresentation of Germans.[16]

Historians have drawn mixed conclusions regarding the success of these anti-nationalist policies. On the one side, the Habsburg Army served as an important identification tool and was partly successful in instilling a common imperial identity beyond ethnic and national divisions among officers and soldiers (Deák 1990; Cole 2014). On the other side, the explicit classification into ethno-linguistic categories – disregarding other identities people might have held – actually increased national identification among soldiers (Stergar and Scheer 2018). Finally, many observers have highlighted the detrimental consequences of communication problems due to the parallel use of 11 languages for battlefield performance (see e.g., Di Michele 2020, p. 58).[17]

World War I

The First World War began with the Austro-Hungarian declaration of war against Serbia on July 28, 1914. Together with 2.7 million men, about 27,000 Italian-speaking Tyroleans were initially called to serve on the front.[18] Contrary to some fears among the military command, the mobilization of Italian soldiers (as those of other ethnic minorities) proceeded in an orderly manner without significant numbers of draft evasions or protests (Di Michele 2018, p. 52). The Italian soldiers from Trentino were predominantly assigned to the Tyrolean regiments, i.e., the 4 *Kaiserjäger* regiments of the *k.u.k.* Common Army and the 3 *Landesschützen* regiments of the Austrian *Landwehr*. All of these units were comprised of approximately 60% German and 40% Italian soldiers, accurately reflecting their population shares (Di Michele 2020, p. 62).

The situation for Italian soldiers changed significantly with the Italian declaration of war against Austria-Hungary on May 23, 1915. As Italy had previously been part of the Triple Alliance with Germany and Austria-Hungary, this "breach of faith, the like of which history has never seen" (Emperor Franz Joseph) sparked massive anti-Italian resentment among Austrians. Although the military command had already before critically assessed many soldiers from ethnic minorities, their distrust towards Ital-

[15] While every soldier had to learn a set of about 80 orders in German, the general communication among soldiers was conducted in the regimental languages (Di Michele 2020, p. 58).

[16] The 1910 military census reports the following composition of regular soldiers: Germans (25.2%), Hungarians (23.1%), Czechs (12.9%), Serbo-Croats (9.0%), Poles (7.9%), Ukrainians (7.6%), Romanians (7.0%), Slovaks (3.6%), Slovenes (2.4%), and Italians (1.3%). In contrast, more than 70% of officers were Germans (k. u. k. Kriegsministerium 1911, pp. 145–146).

[17] Stefan Zweig captures this impression in his memoir *The World of Yesterday*: "The Babelish language confusion caused an additional hardship for the soldiers of every imaginable nationality who had been thrown into this rolling coffin" (Zweig (1948, p. 337), as translated by Scheer (2020, p. 221)).

[18] Di Michele (2020, p. 61) estimates that this would rise to 55,000 soldiers from Trentino until 1918.

ian soldiers reached new heights. Di Michele (2020, p. 89) quotes an internal army report exemplifying this: "No South Tyrolean soldier of Italian tongue must be considered absolutely trustworthy [translated by author, emphases in original]". At the same time, the opening of a new front against Italy meant that Trentino itself became a battleground. Many of the Tyrolean regiments were among the first to be redeployed from the Eastern to the Italian front.[19] Despite orders from the high command that Italian soldiers should not be deployed there, a significant amount of them continued to fight alongside their German comrades on the Southern front due to the near-catastrophic lack of trained personnel (Di Michele 2018, p. 57).

Contemporary accounts by Italian soldiers reflect the worsening of their situation after May 1915. For example, the soldier Alfonso Tomasi wrote in his diary: "The Austrians have always treated us with contempt, but there it seemed like raging hatred. [...] Italy had declared war on Austria, and 'we were to blame' [translated by author, emphasis in original]" (Di Michele 2020, p. 93). Consequently, Überegger (2003) reports how the increasing distrust and sometimes outward discrimination contributed to an increasing desertion rate among Trentino soldiers.[20] In sum, these negative experiences increased a feeling of "otherness" and distance to the Habsburg institutions and Austrians in general among many Italian soldiers (Di Michele 2020, p. 104).

The war ended with the military defeat and abolition of the Habsburg monarchy in November 1918. Subsequently, the Paris Peace Conference sealed the break-up of the multi-ethnic Empire into various national successor states; the Italian-speaking Trentino as well as German-speaking South Tyrol were ceded to Italy.

4.2.2 Conceptual Framework

To frame my empirical analysis on the behavior of ethnic minority soldiers during WWI, I draw on existing theoretical work in economics and political science.

Alesina, Reich, and Riboni (2020) introduce a theoretical framework to examine how states motivate soldiers to exert effort in war. They describe how with the advent of mass armies in the 19th century, states increasingly resorted to nationalism to induce war effort from the population. Their model distinguishes two types of nation-building policies: on the one hand, states can instill positive national sentiment by creating a common identity through education and cultural homogenization. On the other hand, governments may resort to "negative nation-building" by stigmatizing opponents with aggressive propaganda.[21] The authors explicitly highlight the reac-

[19] For more details, see also the discussion in Section 3.2 of Chapter 3.

[20] While desertion legally constituted a capital offence, Überegger (2003) highlights that the death penalty was only very rarely actually executed. This might, in turn, have contributed to lowering the stakes of deserting in the first place.

[21] In their model, the choice of the type of nation-building depends on fiscal capacity. States with

tion by ethnic minorities as a potential downside of such negative nationalism: "Soldiers from regions without any national identity may not put much effort into fighting, or may even break away to join the enemy" (Alesina, Reich, and Riboni 2020, p. 384). With respect to the treatment of ethnic minorities in the Habsburg Empire, Cole (2018, p. 127) mirrors these theoretical implications and points out how repressive state actions and the rigid categorization into national groups constituted a form of "negative nationalization". As discussed in the previous section, this might have been especially true for Italian soldiers in the Habsburg Army after their home country entered into war with their Italian "mother state".[22]

Relatedly, Lyall (2020) provides a comprehensive historical and empirical study of "divided armies" consisting of different ethnic groups. He argues that the degree of "military inequality", i.e., ethnic diversity *within* armies, strongly affects their battlefield performance. A key mechanism constitutes the reluctance of repressed and marginalized "non-core" ethnic groups to risk their life for the regime. In particular, the extent of pre-war discrimination erodes minority soldiers' beliefs about state legitimacy as well as interethnic trust, leading to motivation and coordination problems during battle. In a case study of the Austro-Hungarian Army during World War I, Lyall (2020, pp. 306–316) shows how deep-seated ethnic mistrust and increasing indiscipline among non-core ethnic minorities hampered battlefield performance. Specifically, he points out how "facing the same enemy on the same front, and armed with the same tactics and weapons, k.u.k. Army units displayed considerable variation in their desertion rates depending on the ethnic composition of a given unit." (Lyall 2020, p. 300).

Building on these existing frameworks, I formulate concrete hypotheses about the observable war effort exerted by Italian soldiers in the Habsburg Army during World War I. First, soldiers from the Italian non-core minority were less loyal to the Empire and, therefore, less willing to risk their lives. Relative to German soldiers, they were *less* likely to die or get seriously injured, but *more* likely to desert and end up in enemy captivity. Second, these differences in war effort should have *intensified* after the entry of Italy to World War I. The ensuing increase in interethnic mistrust and "negative nationalism" alienated Italian soldiers and reduced their willingness to fight for the Habsburg Empire.

relatively lower fiscal capacity cannot compete in terms of costly mass public goods needed for positive nation-building and, therefore, have to resort to negative nationalism.

[22] Panel (A) of Figure 3.A.2 in Chapter 3 of this thesis shows an explicit example for the aggressive anti-Italian war propaganda in Austria-Hungary after May 1915.

4.3 Data

4.3.1 Data Sources

World War I Casualty Lists My data on soldiers in World War I is retrieved from the universe of war-time casualty records of the Austro-Hungarian Army. The so-called *Verlustlisten Österreich-Ungarns* (Casualty Lists of Austria-Hungary; Verein für Computergenealogie 2021) were published in 709 volumes on a regular basis throughout the war by the Austro-Hungarian war ministry. In total, they contain 2.7 million entries, of which 35,261 soldiers are recorded as originating from Tyrol. They will provide the baseline sample for the empirical analysis in this paper.

Crucially, the casualty lists report the reason for the "loss" of a soldier, i.e., whether he was killed, wounded, or captured.[23] These outcomes constitute my main measure of exerted war effort: I hypothesize that loyal soldiers exert higher war effort and are more likely to get killed or endure serious injuries (see also Ager et al. 2022; Beatton, Skali, and Torgler 2019); whereas, disloyal soldiers will be less willing to risk their life and more likely to end up in captivity.[24]

Entries in the casualty lists contain a number of biographic information for each soldier: his full name, year of birth, hometown (crown land, district, and municipality)[25], military rank, army unit (regiment and company) as well as place and type of casualty. Only for a subset of about 24% of entries an exact date for the casualty is recorded; for the remaining entries I infer the date by using the publication date of the respective casualty list.[26] In order to capture soldiers serving in the same military unit, I manually harmonize the provided entries on army units in accordance with available historical records on the Austro-Hungarian Army's order of battle.[27] After cleaning, I am left with 262 distinct regiments consisting of 1,147 companies in which Tyrolean soldiers served, with about 85% of soldiers serving in one of the 7 main Tyrolean regiments.[28]

[23] Note that in a military sense, a casualty refers to a soldier being ruled out of battle, which does not necessarily imply death.

[24] In my sample, 21.1% of soldiers are reported to have died, 49.2% as being seriously wounded, and 31.5% as being captured (see Table 4.B.1). This adds up to more than 100%, as there are a few cases in which soldiers were reported to have been captured as well as wounded or deceased. Excluding those cases does not affect results.

[25] The term *Heimatort* (home town) used in the casualty lists refers to the location in which a soldier exerts his *Heimatberechtigung* (right of domicile). Although often corresponding to the birthplace, it refers to the locality to which individuals or their families traditionally belonged (von Hirschhausen 2009).

[26] The average difference between the publication and exact casualty date in the subsample with both information is 136 days with a median difference of 80 days.

[27] I mainly use the information provided by Ehnl (1934) and complement it with information from two private websites: https://web.archive.org/web/20211027104849/http://www.austro-hungarian-army.co.uk/orb14.htm and http://www.austrianphilately.com/dixnut/index.htm (both last accessed on 03/12/2023).

[28] On average, a regiment consisted of 3,000-4,000 soldiers divided into 4 battalions with about 1,000 men further split up into 4 companies with about 250 men (Rauchensteiner 2014, p. 53).

Appendix Figure 4.A.2 shows an exemplary entry in the casualty lists with a German and Italian soldier from Tyrol.

Austro-Hungarian Census I complement this individual-level data with aggregate information on the level of municipalities (*Gemeinden*) from the 1900 Austro-Hungarian census for Tyrol (k. k. Statistische Zentralkommission 1907). For each municipality, the census records the resident population and its distribution by gender, ethnicity (as measured by language groups), and religion as well as information on the local economic structure. I match the census data to the casualty lists via the reported and manually harmonized district (*Bezirk*) and municipality (*Gemeinde*) information. Overall, I am able to successfully match 87.9% of all casualty list entries to a census municipality.

4.3.2 Measuring Ethnicity of Soldiers

The casualty lists do not directly contain information on the ethnicity or language spoken by a soldier. As discussed in Section 4.2.1, the Habsburg institutions classified ethnic groups based on language use (Stergar and Scheer 2018). I, therefore, develop a novel measure based on language classification of names to predict whether a soldier in my dataset was likely German or Italian.

I employ a text classification algorithm from natural language processing to predict the language of the full name of a soldier. More precisely, I rely on the character-based *n*-gram text categorization package `textcat` for `R` developed by Hornik et al. (2013).[29] The basic idea of the approach is to slice the words of a text of unknown language into character *n*-grams, i.e., sequences of 1, 2, ..., *n*-1 characters for a word of length *n*, and calculate the relative frequency distribution of all *n*-grams. This "document profile" is then compared to pre-trained "language profiles" derived in the same way from text of known language. The algorithm predicts the language of the best fitting "language profile". I run the `textcat` approach – restricted on the language profiles for German and Italian – on the reported full name of a solider and receive a prediction of the name being either in German or Italian.[30]

Overall, my preferred measure predicts 61.35% of soldiers to be German and 38.65% to be Italian.[31] This corresponds well to the relative Tyrolean population shares of 57.3% Germans and 42.1% Italians in the 1910 census. Figure 4.1b plots the geographic distribution of ethnicities when aggregating the share of Italian-predicted soldiers to the

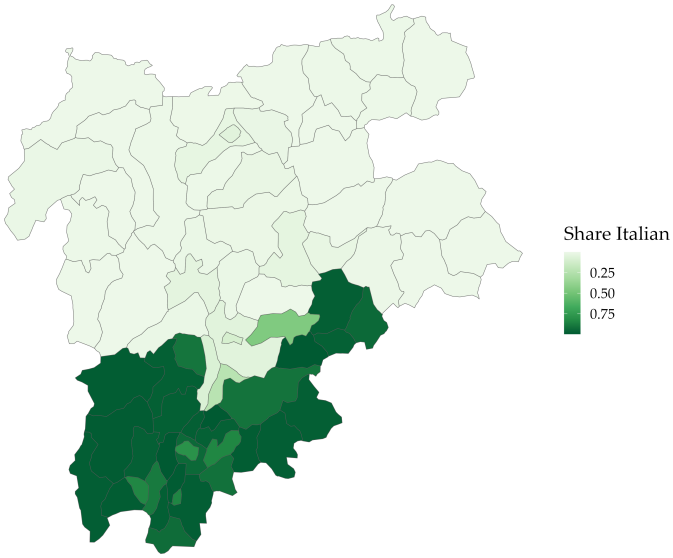
[29] For my predictions, I employed version `1.0-7` of `textcat` (Hornik et al. 2020).

[30] The `textcat` package comes with two sets of language profiles, the `ECIMCI` profiles for 26 European languages and the `TC_byte` profiles for 74 languages. I will standardly be using the `ECIMCI` profile as it achieves a better prediction in terms of the share of predicted Italians relative to the population share. I later report robustness checks with the alternative `TC_byte` profiles.

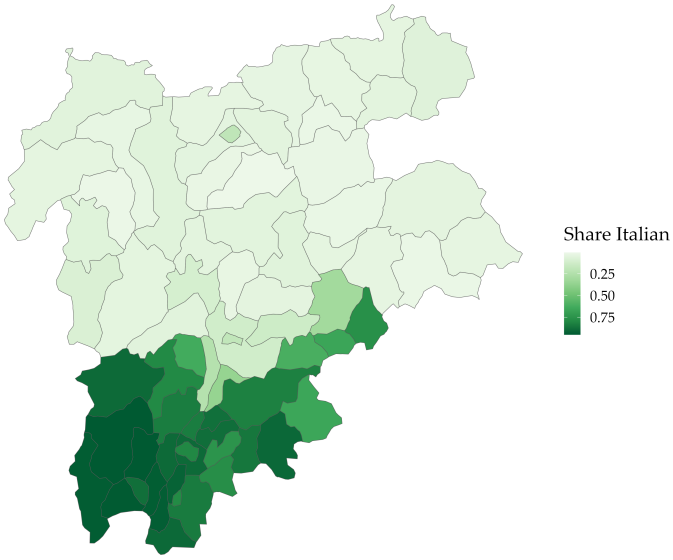
[31] The twenty most common surnames among soldiers predicted as German and Italian, respectively, are reported in Table 4.B.2 in the Appendix.

Figure 4.1: Ethnicity Maps of Tyrol

(a) Share of Italian Population (1900 Census)



(b) Predicted Share of Italian Soldiers (WWI Casualty Lists)



Panel (A) shows the share of the Italian-speaking population by judicial district (*Gerichtsbezirk*) as recorded in the 1900 Austro-Hungarian census for Tyrol (k. k. Statistische Zentralkommission 1907). Panel (B) shows the share of soldiers predicted to be Italian among all soldiers listed in the casualty lists as originating from a municipality (*Gemeinde*) within the judicial district (*Gerichtsbezirk*). Table 4.B.4 reports the corresponding values. Shapefiles are retrieved from GISta Hungarorum: https://www.gistory.hu/g/en/gistory/otka (last accessed on 03/10/2022).

level of their district of origin. Reassuringly, a clear geographic pattern emerges in line with the relatively sharp linguistic border between an Italian-dominated Trentino in the South and a German-dominated North visible in Figure 4.1a. In the Appendix, I report comparisons between the share of predicted Italian soldiers and the actual Italian population share, aggregated by districts (*Bezirke*) in Table 4.B.3 and by subordinate judicial districts (*Gerichtsbezirke*) in Table 4.B.4. Two observations stand out: first, although the predicted ethnic shares are often significantly different from the population shares, the residual differences are in most cases relatively small and remain in a single-digit range. Second, the districts with the largest divergences, e.g., *Primiero* or *Ampezzo* with a more than 20 percentage point deviation, are those in the Ladin settlement area. As Ladins are not separated from Italian speakers in the census and the language classification might not be able to accurately classify them, this might bias the predictions in these areas. I later show that the main results remain robust to the exclusion of Ladin-speaking areas as well as of those districts with large residual differences between predicted and population share.

I construct an alternative measure of a soldier's ethnicity solely based on surnames.[32] For this approach, I exploit the fact that the ethnic groups of the Habsburg Empire map into its modern-day national successor states. From the genealogical website `forebears.io`, I retrieve lists of the most common surnames and their frequency in Austria and Italy, respectively.[33] I then match soldiers' reported surnames to these lists and classify them as German or Italian if their surname is among the most-common names in Austria or Italy, respectively.[34] Overall, this approach allows me to classify the ethnicity of 26.3% of soldiers in my sample, of which 72.6% are classified as German and 27.4% as Italian.

4.4 Results

4.4.1 Municipal-level Results

I begin my empirical analysis by testing for differences in municipality-level casualty rates between ethnic groups. More precisely, I estimate the following regression equation:

$$CasualtyRate_m = \alpha_0 + \alpha_1 \cdot ShareItalian_m + X'_m\gamma + \epsilon_m \tag{4.1}$$

where $CasualtyRate_m$ measures the number of WWI casualties in municipality m per

[32] As the `textcat` algorithm is designed to predict the language of longer texts than names, I use the full name for this approach to maximize text length.

[33] These lists contain 1,000 surnames for Italy (https://forebears.io/italy/surnames) and 700 surnames for Austria (https://forebears.io/austria/surnames, both last accessed on 10/14/2022).

[34] In the case where the same surname shows up in both lists, I assign it to the country where it is more frequent. This only happens once with the surname *Martin* that is more common in Austria than in Italy.

Table 4.1: Municipal-level results

	Casualty Rate			
	(1) Total	(2) Dead	(3) Wounded	(4) Captured
Share Italian	-23.50***	-11.58***	-20.38***	7.75***
	(2.83)	(0.90)	(1.65)	(1.21)
Constant	96.54***	25.31***	43.66***	25.66**
	(28.97)	(8.48)	(14.91)	(11.87)
Controls	✓	✓	✓	✓
Observations	846	846	846	846
R^2	0.076	0.145	0.152	0.075

Notes: Table reports municipal-level regressions as in Equation (4.1). The dependent variable measures the number of WWI casualties per 1,000 male inhabitants. The explanatory variable captures the share of Italians relative to the total population in a municipality in the 1900 census. Controls include the following municipal-level variables: total population, population density, male population share, ethnic diversity (HHI), share of Catholics, Protestants, Jews, share of areable land, statutory city dummy, presence of a military garrison. Heteroskedasticity-robust standard errors are reported in parentheses: *** p<0.01, ** p<0.05, and * p<0.1.

1,000 male inhabitants in the 1900 census. *ShareItalian* captures the share of Italian speakers relative to the total population in municipality m from the 1900 census. Finally, X_m includes municipality-level control variables.

Table 4.1 reports the results of estimating Equation (4.1) on four different types of casualty rates. Column (1) shows the correlation between the pre-war share of Italians and the relative number of *total* WWI casualties suffered by a municipality. On average, a fully Italian-speaking municipality incurs 23.5 less casualties per 1,000 male inhabitants than a homogeneously German municipality. Relative to an estimated baseline total casualty rate of 96 losses per 1,000 men, this constitutes a sizeable 25% difference. In the remaining columns, I now separately report this differential for the three main types of casualties recorded in the WWI lists. Column (2) shows a likewise negative correlation between the share of the Italian population and the rate of soldiers killed during WWI. Relative to the baseline, this amounts to an even larger difference of close to 50% between fully Italian and German municipalities. A similar result arises in column (3) with respect to severely wounded soldiers, with more Italian localities observing a significantly lower rate. In stark contrast, column (4) reports a significantly positive correlation between the share of Italians and the relative number of soldiers having ended up in captivity during WWI. Compared to a German municipality, more than seven additional soldiers from a majoritarian Italian municipality were captured, amounting to a more than 30% difference.

In sum, these findings seem to support my previously formulated hypotheses: soldiers from German-dominated localities seem to be more willing to risk or even give their life during war, while Italian municipalities observe much higher rates of prisoners of war among their soldiers. Although this might be due to differences in war effort, this aggregate-level analysis is merely correlational and does not allow inferences about individual-level behavior. In particular, one might reasonably be concerned about municipality-level selection into the war or specific army units, as the Habsburg Army drafted soldiers based on territorial principles (Wandruszka and Urbanitsch 1987). If, for example, soldiers from certain municipalities were more likely to serve in army units that were (strategically) exposed to battles of varying intensity, this might drive the observed differences in casualty rates. I will therefore now introduce the main individual-level analysis that can address a number of these potential selection concerns.

4.4.2 Individual-level Results

The main regression framework to estimate differential effects of ethnicity on war outcomes on the level of individual soldiers is as follows:

$$Captured_{imct} = \beta \cdot ItalianSoldier_i + Z_i'\delta + \theta_m + \kappa_c + \psi_t + \epsilon_{imct} \tag{4.2}$$

where $Captured_{imct}$ is a dummy variable equal to 1 if soldier i from municipality m serving in company c was recorded as having been captured at time t in the WWI casualty lists. $ItalianSoldier_i$ is a dummy variable equal to 1 if the soldier was predicted to be Italian based on the language classification approach described in Section 4.3.2. Z_i contains individual-control variables, including a soldier's age as well as dummies for officers and volunteer soldiers. My preferred specification includes three sets of fixed effects: θ_m is a vector of fixed effects for the home municipality m of a soldier, whereas κ_c is a vector of fixed effects for the company c in which the soldier was recorded to have served in. Finally, ψ_t is a vector of time fixed effects that will be either the month or the date of the recorded casualty entry.[35]

My main coefficient of interest β captures the differential effect for an Italian soldier to get captured during World War I *relative* to being killed or wounded. In line with my hypotheses – and suggestively confirmed by the municipal-level results in Table 4.1 – I argue that disloyal minority soldiers should be *more* likely to get captured but *less* likely to get killed or wounded due to differences in exerted war effort; hence, I would expect a significantly positive estimate for β.

[35] As the predicted ethnicity, i.e., the treatment variable, is assigned on the level of an individual soldier, I report heteroskedasticity-robust standard errors. One might, however, be worried that casualty outcomes are correlated within army units or reported in batches in the same casualty lists. Alternatively, clustering standard errors on the company-level or the date-level does not affect my results.

Table 4.2: Individual-level results

	Captured						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Italian Soldier	0.185***	0.192***	0.120***	0.098***	0.078***	0.062***	0.035***
	(0.005)	(0.005)	(0.009)	(0.009)	(0.010)	(0.008)	(0.006)
Individual Controls	-	✓	✓	✓	✓	✓	✓
Municipal Controls	-	-	✓	✓	-	-	-
Municipality FE	-	-	-	✓	✓	✓	✓
Company FE	-	-	-	-	✓	✓	✓
Month FE	-	-	-	-	-	✓	-
Date FE	-	-	-	-	-	-	✓
Observations	35,260	32,141	28,193	28,180	20,917	20,880	20,581
Adjusted R^2	0.038	0.044	0.056	0.073	0.206	0.485	0.708
Dep. Var. Mean	0.315	0.317	0.323	0.323	0.309	0.309	0.312

Notes: Table reports soldier-level regressions as in Equation (4.2). The dependent variable is a dummy variable equal to 1 if soldier *i* was recorded as having been captured in the WWI casualty lists. The explanatory variable is a dummy equal variable equal to 1 if soldier *i* was predicted to be Italian using the `textcat` classification on the full name with *ECIMCI* language profiles. Individual controls include: soldier's age as of 1914, dummies for officers and volunteers. Municipal controls include: total population, population density, male population share, share of Germans, Italians, other ethnicities, ethnic diversity (HHI), share of Catholics, Protestants, Jews, share of areable land, statutory city dummy, presence of a military garrison. Municipality fixed effects are dummies for the hometown of soldier *i*. Company fixed effects are dummies for the company-regiment-combination soldier *i* was recorded to serve in. Month and date fixed effects are dummies for the month and the date of the reported casualty or, if not available, the publication date of the respective casualty list issue. Heteroskedasticity-robust standard errors reported in parentheses: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1.

Table 4.2 reports the results of estimating Equation (4.2) on the full sample of Tyrolean soldiers. Column (1) shows the estimated coefficient from a regression without any additional controls. As hypothesized, the effect is positive and statistically significant: an Italian soldier is 18.5 percentage points more likely to get captured than to get wounded or killed relative to a German soldier. Column (2) adds individual-level controls for the age and rank of a soldier, while column (3) adds an array of municipal-level controls from the pre-war 1900 census, e.g., for total population size, ethnic and religious diversity, or the local presence of a military garrison. While the size of the estimated coefficient is somewhat reduced, it remains positive and significant.

Column (4) addresses the concern that, despite controlling for a range of municipal characteristics, there might still be some unobserved local factors correlated with a soldier's ethnicity and affecting his likelihood of capture. With the inclusion of municipality fixed effects, these factors are held constant and only soldiers of different ethnicity *within* the same municipality are compared with each other.[36] Nevertheless, Italian soldiers are still relatively more likely to end up in captivity than German sol-

[36]Note that due to being perfectly collinear with the municipality fixed effects, municipality-level controls can no longer be estimated.

diers.

Next, I additionally include fixed effects for the company in which a soldier served during World War I. This speaks to an important selection concern: one might reasonably expect that – even when holding constant the municipality of origin that is central in the Habsburg drafting system – soldiers are non-randomly assigned to army units. Italian soldiers might have been more likely sent into regiments, or even companies therein, that were then deployed differently during battle to less critical front sections.[37] If this, in turn, affects the relative capture likelihood of Italian soldiers, this might upwardly bias the estimated coefficient. However, column (5) reports a continued positive and significant differential effect for Italians, despite a sizeable drop in sample size due to missing data on soldiers' companies.

Finally, I include fixed effects accounting for the month (column 6) or date (column 7) when the casualty was reported. This accounts for time-specific factors such as variation in battle intensity over the course of the war that might differentially affect the reported casualty outcome of Italian and German soldiers. Despite a further drop in magnitude, the estimated coefficient remains positive and significant. In my preferred specification, Italian soldiers are 3.5 percentage points more likely to get captured rather than killed or wounded, compared to German soldiers from the same municipality serving in the same company and holding time effects constant.[38] Relative to an average capture probability of 31.2%, this amounts to a non-negligible 10% difference.

When running the same set of regressions on a binary indicator for the other casualty outcomes, I confirm the results from the municipal-level analysis. Table 4.B.6 in the Appendix shows a consistently lower likelihood of getting killed (vs. getting captured or wounded) for Italian soldiers. The same negative relationship is found for the likelihood of incurring serious wounds (vs. getting killed or captured) in Appendix Table 4.B.7.[39] Furthermore, in Appendix Table 4.B.8 I confirm the relatively higher likelihood of getting captured for Italian soldiers in a sample that excludes all wounded soldiers. By solely comparing captured with dead soldiers, I restrict myself to cases where a casualty could arguably be incurred only once. Both dead or captured soldiers could not return to service in the Austro-Hungarian Army, while wounded soldiers might potentially have recovered and sustained an additional casualty later.

[37] For example, Lyall (2020, p. 292) reports how the placement and rotation of army units along the front lines was driven by the deep-seated mistrust against ethnic minorities among the Austro-Hungarian military command.

[38] In Table 4.B.5 in the Appendix, I show that the results remain robust to the inclusion of company fixed effects interacted with time dummies (month or date). This additionally controls for company-specific time effects that might have differentially affected soldiers.

[39] Note that, obviously, these two negative effects are partly just the opposite of the positive effect on capture likelihood reported in Table 4.2.

4.4.3 Robustness Checks

In this section, I discuss two types of robustness checks relating to potential concerns about the accuracy of the ethnicity classification approach. First, in Table 4.3, I re-estimate my preferred specification of Equation (4.2) using alternative individual-level measures of ethnicity as explanatory variable.[40] In column (2), I classify soldiers as German or Italian using the alternative `TC_byte` language profiles provided by the `textcat` package on the full name of soldiers. Although this method performs worse in terms of the representativeness of the prediction (67.3% German vs. 32.7% Italian soldiers), I again find a positive and significant estimate for the captivity differential between Italian and German soldiers. In column (3), I employ an alternative method based on the most common surnames in the modern-day successor states of Austria and Italy. Although this approach can only classify a quarter of soldiers and substantially reduces the sample size, I nevertheless find a positive coefficient estimate significant at the 10% level.

Second, I directly address some weaknesses of my preferred ethnicity measure. As discussed in Section 4.3.2, the `textcat` classification resulted in significant differences between the predicted and actual population share of Italians in Ladin-speaking areas. In column (2) of Appendix Table 4.B.9, I exclude all districts making up the historic settlement area of Ladins; however, I do not find that this affects my main result. In columns (3) and (4), I go one step further and drop all districts (*Bezirke*) or judicial districts (*Gerichtsbezirke*) where the residual difference between predicted and population Italian share is bigger than 10 percentage points. Again, excluding these roughly 30% of observations from my sample, I do not find them to be driving my results.

4.5 Heterogeneities

My results so far have established significant differences in casualty outcomes between Italian and German soldiers, with the former being more likely to get captured and less likely to get killed or wounded. I have previously argued that this might be due to differences in loyalty, resulting in lower war effort exerted by soldiers from non-core ethnic groups. Due to the unobservability of individuals' motivation or loyalty in this historical setting, I cannot directly test for this implied mechanism. However, I will provide suggestive evidence from heterogeneities in the observed differential effects that are in line with such an interpretation.

To this end, I interact the predicted ethnicity of a soldier with a number of potential mediating factors. I began by investigating whether the local composition of the population differentially affects the war effort exerted by Italian soldiers. For example,

[40] Column (1) corresponds to column (7) from Table 4.2 and is included for reference.

Table 4.3: Robustness: Alternative Ethnicity Measures

	Captured		
	(1)	(2)	(3)
Italian Soldier (ECI)	0.035***		
	(0.007)		
Italian Soldier (TCB)		0.020***	
		(0.006)	
Italian Soldier (Surname)			0.041*
			(0.024)
Individual Controls	✓	✓	✓
Municipality FE	✓	✓	✓
Company FE	✓	✓	✓
Date FE	✓	✓	✓
Observations	20,581	20,577	5,076
Adjusted R^2	0.708	0.708	0.703
Dep. Var. Mean	0.312	0.312	0.304

Notes: Table reports soldier-level regressions as in Equation (4.2). The dependent variable is a dummy variable equal to 1 if soldier i was recorded as having been captured in the casualty lists. The explanatory variable is a dummy equal variable equal to 1 if soldier i was predicted to be Italian as follows: in column (1), using `textcat` classification on full name with `ECIMCI` language profiles; in column (2) using `textcat` classification on full name with `TC_byte` language profiles; in column (3) using most common surnames in Austria and Italy from `forebears.io`. See Table 4.2 for notes on controls and fixed effects. Heteroskedasticity-robust standard errors are reported in parentheses: *** p<0.01, ** p<0.05, and * p<0.1.

one might expect that an ethnic Italian living in a German-dominated village might be better integrated into the German majority culture and, hence, have a higher loyalty towards the Empire. I test this claim by augmenting my baseline regression as follows:

$$\begin{aligned} Captured_{imct} =& \beta_1 \cdot ItalianSoldier_i + \beta_2 \cdot ItalianSoldier_i \times ShareItalian_m \\ &+ Z_i'\delta + \theta_m + \kappa_c + \psi_t + \epsilon_{imct} \end{aligned} \tag{4.3}$$

where $ShareItalian_m$ is the share of the Italian-speaking population in the home municipality m of soldier i as of 1900. The remaining variables are defined in the same way as in Equation (4.2).[41] Column (1) of Table 4.4 presents the results from estimating Equation (4.3). While the baseline coefficient β_1 for $ItalianSoldier_i$ is not significant and close to zero, coefficient β_2 for the interaction term with the local Italian population share is positive and statistically significant. This implies that the differential effect in casualty outcomes is entirely driven by Italian soldiers coming from ethni-

[41] Note that due to being perfectly collinear with municipality fixed effects θ_m, the baseline effect for $ShareItalian_m$ cannot be estimated in this framework.

Table 4.4: Heterogeneities

	Captured			
	(1)	(2)	(3)	(4)
Italian Soldier	0.009	-0.058	-0.003	0.007
	(0.009)	(0.048)	(0.008)	(0.006)
Italian Soldier × Share Italian	0.047***			
	(0.014)			
Italian Soldier × HHI		0.102*		
		(0.052)		
Italian Soldier × Post Italy			0.057***	
			(0.009)	
Italian Front				-0.648***
				(0.011)
Italian Soldier × Italian Front				0.035***
				(0.008)
Individual Controls	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓
Company FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Observations	20,581	20,581	20,581	19,394
Adjusted R^2	0.708	0.708	0.709	0.816
Dep. Var. Mean	0.312	0.312	0.312	0.324

Notes: Table reports soldier-level regressions as in Equation (4.3). The dependent variable is a dummy variable equal to 1 if soldier i was recorded as having been captured in the WWI casualty lists. *ItalianSoldier* is a dummy equal variable equal to 1 if soldier i was predicted to be Italian using the `textcat` classification on the full name with *ECIMCI* language profiles. *ShareItalian* is the share of Italian population from the 1900 census. *HHI* is the Herfindahl-Hirschman-Index of ethnic diversity calculated from the population of Italians, Germans, and others in the 1900 census. *PostItaly* is a dummy variable equal to 1 if the casualty was recorded after May 23, 1915. *ItalianFront* is a dummy variable equal to 1 if the casualty was recorded on the Italian front. See Table Table 4.2 for notes on controls and fixed effects. Heteroskedasticity-robust standard errors reported in parentheses: *** p<0.01, ** p<0.05, and * p<0.1.

cally homogeneous Italian municipalities.[42] On the other hand, Italian soldiers living among Germans are not more likely to get captured during World War I, suggesting they might feel greater loyalty to the Empire and their German compatriots.

In a similar vein, it might be suspected that local ethnic diversity affects the loyalty of Austro-Hungarian citizens. On the one hand, its multi-nationality was one of the Empire's central features, potentially making soldiers from ethnically heterogeneous areas more willing to defend it. On the other hand, long-standing research in social psychology has established the importance of intergroup contact for the reduction of

[42]Note that the vast majority of soldiers come from either ethnically homogeneous German or Italian municipalities, as can be seen in Figure 4.A.3 in the Appendix.

prejudices (Allport 1954). I test this hypothesis in column (2) by interacting an individual soldier's ethnicity with a municipality-level measure of ethnic diversity. The Herfindahl-Hirschman Index (HHI) measures the local concentration of ethnic groups, i.e., larger values implying less ethnic diversity, and constitutes a common approach in the literature (see Steele et al. 2022).[43] As can be seen in Appendix Figure 4.A.4, although more than 30% of soldiers come from fully homogeneous municipalities, there still is a certain degree of variation in local ethnic diversity. The estimated coefficients from replacing $ShareItalian_m$ with HHI_m in Equation (4.3) are shown in column (2) of Table 4.4. As hypothesized, I find that Italian soldiers from less diverse communities are relatively more likely to end up in war captivity. This is in line with an interpretation of higher exposure to out-groups reducing prejudice and decreasing the likelihood of disloyal behavior in times of war.

Next, I will exploit the time dimension of my dataset recording casualties throughout the course of the war. In particular, I am interested in exploring how Italy's unexpected entry to World War I changed the effort exerted by Italian soldiers fighting on the side of Austria-Hungary against other Italians. As discussed in Section 4.2.1, the Italian war declaration led to a massive deterioration of the situation for many Italian soldiers, facing distrust and resentment from Austrians and leading to feelings of alienation and increasing distance to Germans (Di Michele 2020, p. 57). In the theoretical framework by Alesina, Reich, and Riboni (2020), this corresponds to an intensification of "negative nation-building" resorting to aggressive propaganda against the enemy. While this can motivate members of the in-group to increase their war effort, out-group members might be alienated and increasingly reluctant to give their life for the state. I will test for such changes in exerted war effort by estimating the following regression:

$$\begin{aligned} Captured_{imct} =& \beta_1 \cdot ItalianSoldier_i + \beta_2 \cdot ItalianSoldier_i \times PostItaly_t \\ &+ Z_i'\delta + \theta_m + \kappa_c + \psi_t + \epsilon_{imct} \end{aligned} \tag{4.4}$$

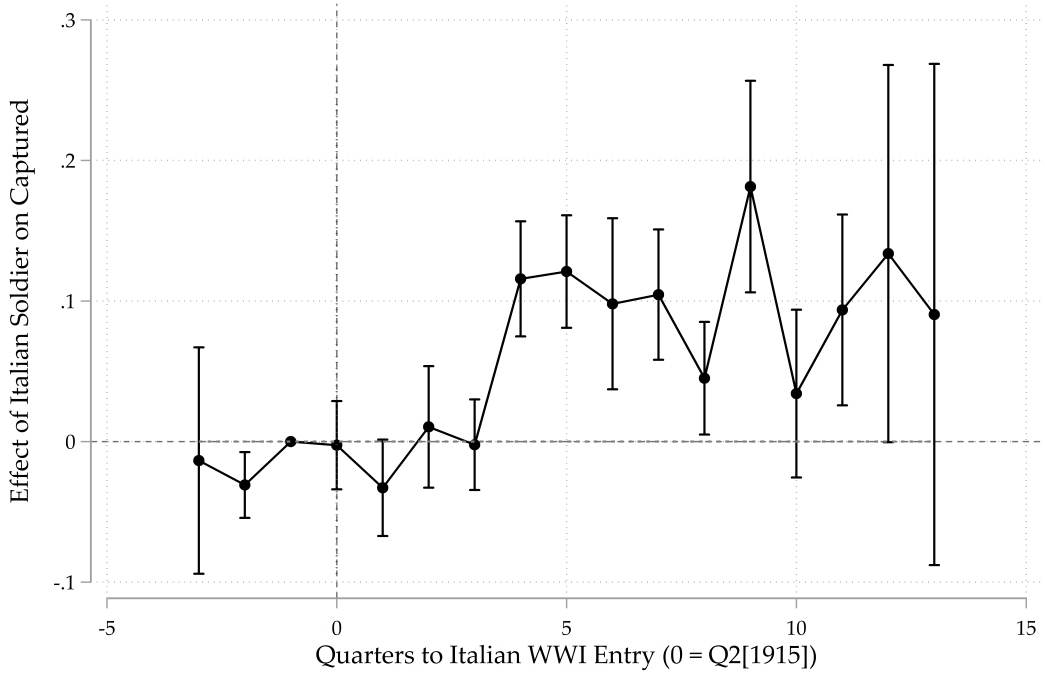
where $PostItaly_t$ is a dummy variable equal to 1 if the casualty was recorded after the Italian war entry on May 23, 1915. All remaining variables are defined as in Equation (4.2).[44]

Column (3) of Table 4.4 reports the ensuing results from estimating Equation (4.4). The estimated coefficient for β_1 is a precisely estimated zero, implying that *before* Italy's declaration of war Italian soldiers were not more likely to get captured than their Ger-

[43] More precisely, the HHI computes the sum of the squared shares p of each ethnic group $j \in G$ in the local population: $HHI = \sum_{j=1}^{G} p_j^2$. Larger values indicate less diversity, with the maximum value 1 implying total homogeneity.

[44] Again, note that due to the inclusion of date fixed effects ψ_t the baseline effect for $PostItaly_t$ cannot be estimated.

Figure 4.2: Effect of Italian Soldier on Captured over Time



Notes: The event study graph plots quarterly coefficients for the effect of soldiers with predicted Italian ethnicity on a binary indicator for the soldier reported as getting captured in the WWI casualty lists. Circles represent coefficients and spikes depict 95% confidence intervals. The vertical line highlights the Italian war entry in Q2[1915] (May 23, 1915). The interaction with Q1[1915] is excluded from the regression.

man comrades. Contrary to this, the positive estimate for β_2 implies that *after* Italy had entered the war, Italian soldiers were significantly more likely to get captured than killed or wounded relative to Germans. Although I cannot fully rule out that this change might be due to more discrimination of Italian soldiers even within companies, this result is also in line with an increasing disloyalty and reduced willingness to fight, as hypothesized by Alesina, Reich, and Riboni (2020).

I further explore this change in casualty outcomes over time by separately estimating the differential effect for Italian soldiers by quarter.[45] The resulting quarterly coefficients are plotted in Figure 4.2. I find that there is no differential effect on a soldier getting captured before the Italian war entry in the second quarter of 1915 which contin-

[45] Formally, I run the following event-study regression framework:

$$Captured_{imct} = \beta_1 \cdot ItalianSoldier_i + \sum_{\tau=Q3[1914]}^{Q3[1918]} \beta_j \cdot ItalianSoldier_i \times \mathbb{1}[Q = \tau] + Z_i'\delta + \theta_m + \kappa_c + \psi_t + \epsilon_{imct}$$

where a sequence of coefficients β_j for the interaction of each quarter Q that was fully part of WWI, i.e., Q3[1914] to Q3[1918], with the ethnicity indicator for a soldier is estimated. All estimated coefficients are relative to the last quarter before the Italian war entry (Q1[1915]) that is excluded from the regression.

ues until the first quarter of 1916. In contrast, from the second quarter of 1916 onward I observe a significantly higher relative capture rate among Italian minority soldiers that persists until the end of WWI. A number of reasons might explain why I observe this effect only a year after the Italian war entry: first, the Austro-Hungarian Army was primarily concerned with holding and stabilizing the newly opened front in the first year of war with Italy. Only in May 1916 – after having secured advances on the Eastern front – did Austria-Hungary launch major counteroffensives against Italy.[46] It is conceivable that disloyal Italian soldiers were less willing to exert life-threatening effort during such risky offensives. Second, historians have pointed to the rapidly collapsing morale among Austro-Hungarian soldiers with mass desertion waves among ethnic minority soldiers in the last years of WWI (Überegger 2003, p. 357; Lyall 2020, p. 315). Last, part of the delay might be mechanically explained by the reporting lag for many observations in my data until a soldier was recorded in the casualty lists (see Footnote 26).

Finally, I explore whether Italian soldiers exerted different war effort depending on which front they were deployed. One might reasonably expect that Italian soldiers where less likely to fight until death when confronted with "their own" on the Southern Front against the Italian Army. I empirically test for this hypothesis by estimating the following regression framework:

$$\begin{aligned} Captured_{imct} =& \beta_1 \cdot ItalianSoldier_i + \beta_2 \cdot ItalianFront_i \\ &+ \beta_3 \cdot ItalianSoldier_i \times ItalianFront_i \\ &+ Z_i'\delta + \theta_m + \kappa_c + \psi_t + \epsilon_{imct} \end{aligned} \tag{4.5}$$

where $ItalianFront_i$ is a dummy variable equal to 1 if the casualty of soldier i was recorded on the Italian front. This variable is constructed in two steps: first, I manually code whether an explicitly mentioned place of the recorded casualty lies on the Italian front. Second, for entries lacking this information, I rely on historical accounts by Glaise von Horstenau (1932) to infer the front to which a soldier's regiment was deployed at the time of the recorded entry date. All remaining variables are defined as in Equation (4.2). Column (4) of Table 4.4 reports the resulting coefficient estimates. As can be seen by the positive and significant estimate for β_3, I find that only Italian soldiers fighting on the Italian front observe a relatively higher likelihood of getting captured.[47]

[46] The so-called *Südtiroloffensive* ("South Tyrol Offensive") was initiated on May 15, 1916 (Glaise von Horstenau 1932, Vol. 4, p. 254).

[47] The large and negative estimate for β_2 can be explained by the specific battlefield conditions at the different fronts: while the Eastern war theater saw a dynamic war with major front shifts, the Italian front quickly developed into a static trench warfare with generally lower rates of captivity (see Di Michele 2020, p. 98).

In this section, I provided suggestive evidence that the established differential capture rate among Italian soldiers is partly driven by two sets of factors. First, local ethnic composition matters, with Italian soldiers that have less exposure to Germans in their daily life being more likely to end up in captivity. Second, the entry of Italy to World War I was crucial: only afterwards and only on the Italian Front do I observe differences in soldiers' behavior. While I cannot fully rule out alternative explanations relating to discrimination of minority soldiers, these results are in line with an increasing disloyalty and reduced willingness to fight among Italians (Alesina, Reich, and Riboni 2020; Lyall 2020).

4.6 Conclusion

Does loyalty and the willingness to fight in a war differ by a soldier's ethnic identity? Using a novel individual-level dataset from the Austro-Hungarian Army during World War I, I have shown that Italian minority soldiers were significantly less likely to die and get seriously wounded, but significantly more likely to end up in enemy captivity than their German brothers-in-arms. This differential is largely driven by soldiers with less exposure to their German out-group and, crucially, materializes only after the Italian declaration of war against Austria-Hungary. Moreover, when fighting against "their own" on the Southern front, Italian soldiers seem to be less willing to engage in potentially mortal combat.

I have claimed that the observed differences reflect disloyal behavior among Italian minority soldiers, in line with theoretical predictions by Alesina, Reich, and Riboni (2020). However, it should be noted that the empirical analysis does not allow me to fully rule out alternative explanations. Despite controlling for potential selection concerns relating to selection into companies, it could still be the case that Italian soldiers are differentially treated even within such small-scale units. For example, company commanders might send minority soldiers to more or less risky front sections, out of either spite or fear of rebellion; in turn, this might result in differential casualty outcomes. Investigating the multi-ethnic interplay between commanders and soldiers within the Habsburg Army constitutes an interesting avenue for future research.

In any case, this work has confirmed the importance of social identities for the behavior of individuals. Feelings of belonging to or exclusion from a common group can leave deep impacts on consequential choices, such as risking one's life for the nation. Understanding these motives is especially relevant in times when the specter of war has returned to Europe.

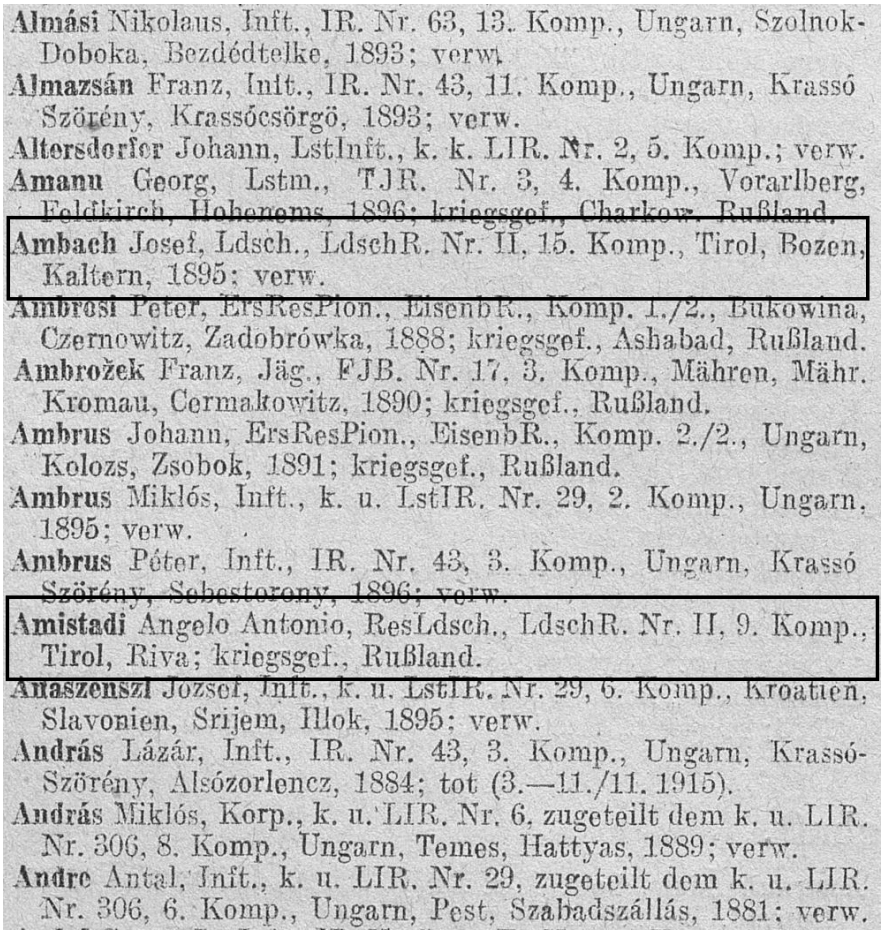
Appendix to Chapter 4

4.A Additional Figures

Figure 4.A.1: Ethnic Groups in the Austro-Hungarian Empire, 1910

Notes: Geographic distribution of ethnic groups in the Austro-Hungarian Empire as recorded in the 1910 census. Source: https://commons.wikimedia.org/wiki/File:Austria_Hungary_ethnic.svg (last accessed on 03/10/2022).

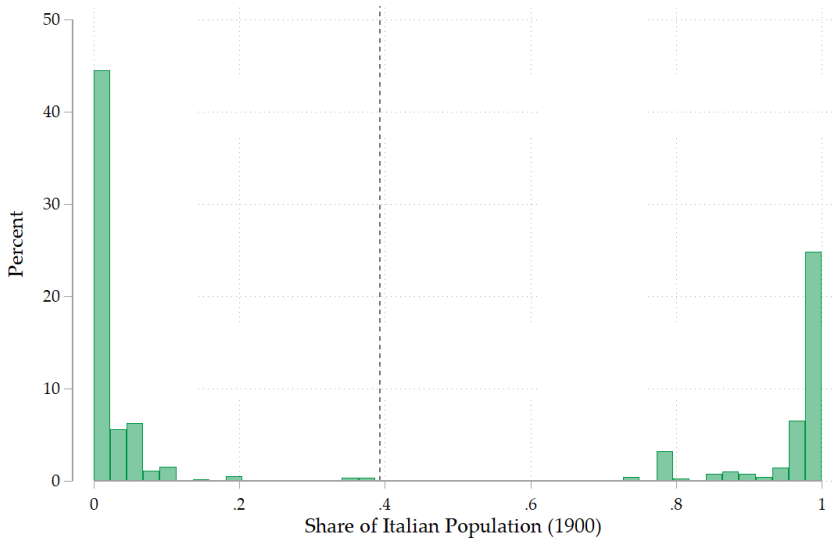
Figure 4.A.2: Exemplary entries in WWI Casualty Lists



Almási Nikolaus, Inft., IR. Nr. 63, 13. Komp., Ungarn, Szolnok-Doboka, Bezdédtelke, 1893; verw.
Almazsán Franz, Inft., IR. Nr. 43, 11. Komp., Ungarn, Krassó Szörény, Krassócsörgö, 1893; verw.
Altersdorfer Johann, LstInft., k. k. LIR. Nr. 2, 5. Komp.; verw.
Amann Georg, Lstm., TJR. Nr. 3, 4. Komp., Vorarlberg, Feldkirch, Hohenems, 1896; kriegsgef., Charkow, Rußland.
Ambach Josef, Ldsch., LdschR. Nr. II, 15. Komp., Tirol, Bozen, Kaltern, 1895; verw.
Ambrosi Peter, ErsResPion., EisenbR., Komp. 1./2., Bukowina, Czernowitz, Zadobrówka, 1888; kriegsgef., Ashabad, Rußland.
Ambrožek Franz, Jäg., FJB. Nr. 17, 3. Komp., Mähren, Mähr. Kromau, Cermakowitz, 1890; kriegsgef., Rußland.
Ambrus Johann, ErsResPion., EisenbR., Komp. 2./2., Ungarn, Kolozs, Zsobok, 1891; kriegsgef., Rußland.
Ambrus Miklós, Inft., k. u. LstIR. Nr. 29, 2. Komp., Ungarn, 1895; verw.
Ambrus Péter, Inft., IR. Nr. 43, 3. Komp., Ungarn, Krassó Szörény, Sebestorony, 1896; verw.
Amistadi Angelo Antonio, ResLdsch., LdschR. Nr. II, 9. Komp., Tirol, Riva; kriegsgef., Rußland.
Anaszenszi Jozsef, Inft., k. u. LstIR. Nr. 29, 6. Komp., Kroatien, Slavonien, Srijem, Illok, 1895; verw.
András Lázár, Inft., IR. Nr. 43, 3. Komp., Ungarn, Krassó-Szörény, Alsózorlencz, 1884; tot (3.—11./11. 1915).
András Miklós, Korp., k. u. LIR. Nr. 6, zugeteilt dem k. u. LIR. Nr. 306, 8. Komp., Ungarn, Temes, Hattyas, 1889; verw.
Andre Antal, Inft., k. u. LIR. Nr. 29, zugeteilt dem k. u. LIR. Nr. 306, 6. Komp., Ungarn, Pest, Szabadszállás, 1881; verw.

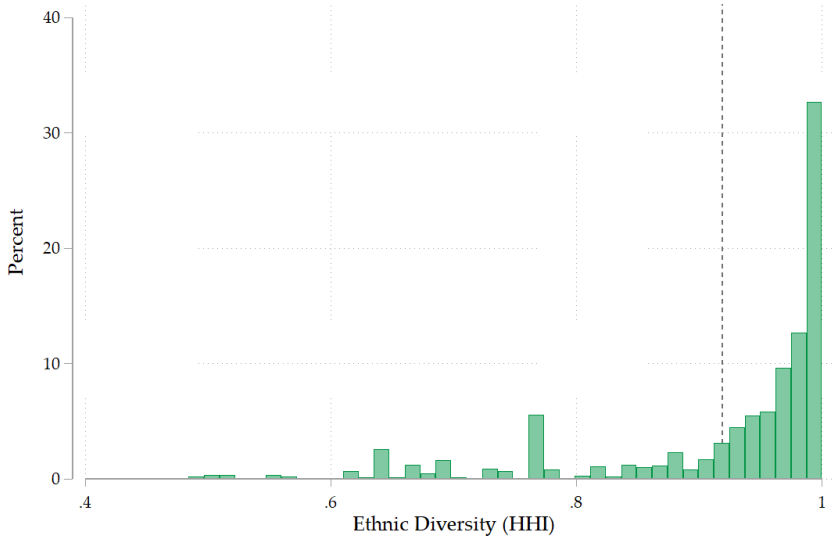
Notes: Figure shows an exemplary page cutout from the Casualy List No. 417 published on May 5, 1916. The first highlighted soldier is *Josef Ambach* from the municipality of *Kaltern* in the district of *Bozen* in Tyrol. He was born in 1895 and served in Company No. 15 of *Landesschützen* Regiment No. II. He is reported to have been wounded. The language classification predicts him to be German. The second highlighted soldier shows *Angelo Antonio Amistadi* from the municipality of *Riva* from the district of the same name. He served in Company No. 9 of *Landesschützen* Regiment No. II. He is reported to have been captured in Russia. The language classification predicts him to be Italian. Source: Verein für Computergenealogie (2021).

Figure 4.A.3: Distribution of Italian Population Share (1900)



Notes: Figure shows the distribution of the municipal-level Italian population share from the 1900 census in the full sample of soldiers. The dashed vertical line highlights the sample mean (0.39).

Figure 4.A.4: Distribution of Herfindahl-Hirschman Index (1900)



Notes: Figure shows the distribution of the municipal-level ethnic diversity as measured by the Herfindahl-Hirschman Index (HHI) from the 1900 census in the full sample of soldiers. Higher values imply less ethnic diversity. The dashed vertical line highlights the sample mean (0.92).

4.B Additional Tables

Table 4.B.1: Summary Statistics

	Mean	Std. Dev.	Min.	Max.	Obs.
PANEL A: *Individual-level Variables*					
Soldier Dead	0.21	0.41	0.00	1.00	35,261
Soldier Wounded	0.49	0.50	0.00	1.00	35,261
Soldier Captured	0.32	0.46	0.00	1.00	35,261
Age	26.22	6.77	13.00	89.00	33,762
Regular	0.97	0.17	0.00	1.00	33,533
Officer	0.02	0.15	0.00	1.00	33,533
Volunteer	0.01	0.08	0.00	1.00	33,533
Post Italy	0.74	0.44	0.00	1.00	35,196
Italian Front	0.44	0.50	0.00	1.00	31,612
PANEL B: *Municipal-level Variables*					
Casualties (per 1,000 male)	86.59	37.91	1.34	552.08	30,706
Dead (per 1,000 male)	18.58	12.68	0.00	177.08	30,706
Wounded (per 1,000 male)	42.65	22.15	0.00	276.47	30,706
Captured (per 1,000 male)	27.09	14.18	0.00	114.58	30,706
Share German	0.58	0.46	0.00	1.00	31,008
Share Italian	0.39	0.46	0.00	1.00	31,008
Share Other Ethnicity	0.00	0.02	0.00	0.19	31,008
Ethnic Diversity (HHI)	0.92	0.11	0.45	1.00	31,008
Share Catholic	0.99	0.04	0.00	1.00	31,008
Share Protestant	0.00	0.03	0.00	0.72	31,008
Share Jewish	0.00	0.00	0.00	0.04	31,008
Share Other Religion	0.00	0.00	0.00	0.01	31,008
Share Male	0.50	0.03	0.21	0.82	31,008
Population Density (per ha)	9.36	32.29	0.01	216.46	30,933
Share Areable Land	0.86	0.16	0.13	2.74	30,933
Military Garrison	0.14	0.35	0.00	1.00	31,008
Statutory City	0.09	0.28	0.00	1.00	31,008
PANEL C: *Ethnicity Measures*					
Italian Soldier (ECI)	0.39	0.49	0.00	1.00	35,260
Italian Soldier (TCB)	0.33	0.47	0.00	1.00	35,257
Italian Soldier (Surname)	0.27	0.45	0.00	1.00	9,281

Table 4.B.2: Most common surnames by predicted ethnicity

#	German	Obs.	Italian	Obs.
1	Mair	173	Tomasi	66
2	Hofer	140	Moser	65
3	Huber	137	Sartori	61
4	Kofler	135	Rossi	58
5	Gruber	112	Pedrotti	57
6	Maier	108	Ferrari	54
7	Pichler	106	Zeni	53
8	Steiner	98	Casagranda	46
9	Egger	97	Degasperi	45
10	Thaler	97	Martinelli	45
11	Auer	92	Bortolotti	43
12	Moser	84	Stefani	40
13	Wieser	81	Agostini	39
14	Gasser	78	Valentini	39
15	Leitner	78	Fontana	37
16	Rainer	76	Girardi	36
17	Lechner	73	Bazzanella	34
18	Mayr	69	Dalpiaz	34
19	Brunner	68	Giuliani	34
20	Mayer	66	Lorenzi	34

Notes: Table shows the twenty most common surnames among the soldiers predicted as German or Italian, respectively. The `textcat` classification was run on the full name of soldiers using the `ECIMCI` language profiles.

Table 4.B.3: Italian Shares by District (*Bezirk*)

	Bezirk	Obs.	Predicted	Census	Difference	p-Value
1	Ampezzo	115	73.04	95.74	-22.70	0.00***
2	Borgo	1680	83.15	95.99	-12.83	0.00***
3	Bozen (Stadt)	447	19.02	10.32	8.69	0.00***
4	Bozen, Bolzano	2517	17.40	11.80	5.60	0.00***
5	Brixen, Bressanone	1146	6.89	2.54	4.35	0.00***
6	Bruneck	1856	7.97	15.95	-7.98	0.00***
7	Cavalese	959	77.06	92.49	-15.43	0.00***
8	Cles	1316	77.36	96.15	-18.80	0.00***
9	Imst	848	6.49	0.21	6.28	0.00***
10	Innsbruck	1879	5.11	1.85	3.26	0.00***
11	Innsbruck (Stadt)	1372	20.85	5.42	15.43	0.00***
12	Kitzbühel	1116	6.63	0.12	6.51	0.00***
13	Kufstein	1382	4.78	0.36	4.42	0.00***
14	Landeck	878	5.47	0.82	4.65	0.00***
15	Lienz	1283	3.74	0.08	3.66	0.00***
16	Meran	2277	8.34	2.17	6.18	0.00***
17	Primiero	468	62.39	97.95	-35.56	0.00***
18	Reutte	767	6.39	0.03	6.36	0.00***
19	Riva	733	90.59	88.43	2.15	0.05**
20	Roveredo (Stadt)	231	79.65	85.06	-5.41	0.04**
21	Roveredo, Rovereto	1864	88.57	97.37	-8.80	0.00***
22	Schwaz	1281	3.67	0.06	3.61	0.00***
23	Tione	971	93.20	98.25	-5.04	0.00***
24	Trient (Stadt)	697	78.77	79.45	-0.68	0.66
25	Trient, Trento	2924	85.33	95.47	-10.14	0.00***

Notes: Table shows data aggregated by district (*Bezirk*). *Obs*. reports how many soldiers in the casualty lists are reported to originate from the respective district. *Predicted* reports the percentage of soldiers predicted to be Italian. *Census* reports the percentage of the Italian-speaking population from the 1900 census (k. k. Statistische Zentralkommission 1907). $p - Value$ reports the p-value from a one-sample two-sided t-test on the difference between *Predicted* and *Census*. Significance levels: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1.

Table 4.B.4: Italian Shares by Judicial District (*Gerichtsbezirk*)

	Gerichtsbezirk	Obs.	Predicted	Census	Difference	p-Value
1	Ala	323	91.33	93.34	-2.00	0.2
2	Ampezzo	103	73.79	94.43	-20.65	0.00***
3	Arco	384	92.71	88.73	3.98	0.00***
4	Borgo	633	83.57	98.52	-14.95	0.00***
5	Bozen (Stadt)	447	19.02	10.32	8.69	0.00***
6	Bozen, Bolzano	760	13.03	5.25	7.77	0.00***
7	Brixen, Bressanone	710	7.18	3.47	3.71	0.00***
8	Bruneck	593	4.22	2.23	1.99	0.02**
9	Buchenstein, Livinallongo	12	66.67	97.15	-30.48	0.06*
10	Cavalese	819	79.49	90.94	-11.45	0.00***
11	Cembra	334	87.72	99.72	-11.99	0.00***
12	Civezzano	313	88.82	97.17	-8.35	0.00***
13	Cles	582	77.49	97.63	-20.14	0.00***
14	Condino	234	92.31	98.15	-5.84	0.00***
15	Enneberg	288	29.86	97.90	-68.04	0.00***
16	Fassa	140	62.86	99.67	-36.81	0.00***
17	Fondo	303	63.04	90.29	-27.25	0.00***
18	Fügen	250	3.60	0.08	3.52	0.00***
19	Glurns	398	10.05	0.09	9.96	0.00***
20	Hall	548	6.02	1.70	4.32	0.00***
21	Hopfgarten	339	5.31	0.24	5.07	0.00***
22	Imst	384	5.73	0.16	5.57	0.00***
23	Innsbruck	490	6.12	2.82	3.31	0.00***
24	Innsbruck (Stadt)	1372	20.85	5.42	15.43	0.00***
25	Kaltern, Caldaro	642	22.59	8.24	14.35	0.00***
26	Kastelruth, Castelrotto	311	13.50	45.45	-31.94	0.00***
27	Kitzbühel	777	7.21	0.07	7.14	0.00***
28	Klausen, Chiusa	312	6.09	0.16	5.93	0.00***
29	Kufstein	819	5.49	0.41	5.08	0.00***
30	Lana	436	6.19	1.34	4.85	0.00***
31	Landeck	545	4.59	1.18	3.41	0.00***
32	Lavis	354	84.46	97.89	-13.43	0.00***
33	Levico	555	78.92	91.60	-12.68	0.00***
34	Lienz	437	3.89	0.11	3.78	0.00***
35	Malè	431	87.24	98.79	-11.55	0.00***
36	Meran	771	11.54	4.10	7.45	0.00***
37	Mezolombardo	452	83.41	97.68	-14.27	0.00***
38	Mieders	164	3.66	0.36	3.30	0.03**
39	Mori	363	92.29	98.08	-5.79	0.00***
40	Nauders	205	9.27	0.41	8.85	0.00***
41	Neumarkt, Egna	351	36.18	22.87	13.32	0.00***
42	Nogaredo	412	91.75	98.99	-7.25	0.00***
43	Passeier	217	5.07	1.24	3.83	0.01**
44	Pergine	478	75.10	84.52	-9.42	0.00***
45	Primiero	468	62.39	97.95	-35.56	0.00***
46	Rattenberg	563	3.73	0.28	3.45	0.00***
47	Reutte	767	6.39	0.03	6.36	0.00***
48	Ried	128	3.12	0.02	3.10	0.05**
49	Riva	275	87.27	84.34	2.94	0.15

(continued on next page)

	Bezirk	Obs.	Predicted	Census	Difference	p-Value
50	Roveredo (Stadt)	231	79.65	85.06	-5.41	0.04**
51	Roveredo, Rovereto	766	83.94	98.26	-14.32	0.00***
52	Sarnthal	141	4.26	1.41	2.84	0.1*
53	Schlanders	455	5.05	0.10	4.95	0.00***
54	Schwaz	619	4.52	0.07	4.45	0.00***
55	Sillian	444	1.80	0.11	1.70	0.01***
56	Silz	464	7.11	0.24	6.87	0.00***
57	Steinach	278	2.52	1.84	0.68	0.47
58	Stenico	198	94.44	97.47	-3.03	0.06*
59	Sterzing	436	6.42	1.03	5.39	0.00***
60	Strigno	492	87.40	98.16	-10.76	0.00***
61	Taufers	520	3.65	0.08	3.57	0.00***
62	Telfs	399	5.01	0.22	4.80	0.00***
63	Tione	539	93.14	98.84	-5.70	0.00***
64	Trient (Stadt)	697	78.77	79.45	-0.68	0.66
65	Trient, Trento	597	89.45	94.50	-5.05	0.00***
66	Val di Ledro	74	91.89	98.60	-6.71	0.04**
67	Vezzano	396	89.65	99.59	-9.94	0.00***
68	Welsberg	455	3.96	0.89	3.06	0.00***
69	Windischmatrei	402	5.72	0.01	5.71	0.00***
70	Zell am Ziller	412	2.43	0.04	2.39	0.00***

Notes: Table shows data aggregated by judicial district (*Gerichtsbezirk*). *Obs*. reports how many soldiers in the casualty lists are reported to originate from the respective district. *Predicted* reports the percentage of soldiers predicted to be Italian. *Census* reports the percentage of the Italian-speaking population from the 1900 census (k. k. Statistische Zentralkommission 1907). $p - Value$ reports the p-value from a one-sample two-sided t-test on the difference between *Predicted* and *Census*. Significance levels: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1.

Table 4.B.5: Interacted Fixed Effects

	Captured	
	(1)	(2)
Italian Soldier	0.041***	0.007**
	(0.006)	(0.004)
Individual Controls	✓	✓
Municipality FE	✓	✓
Company FE	✓	✓
Month FE	✓	-
Date FE	-	✓
Company × Month FE	✓	-
Company × Date FE	-	✓
Observations	19,109	17,269
Adjusted R^2	0.731	0.923
Dep. Var. Mean	0.318	0.330

Notes: Table reports soldier-level regressions as in Equation (4.2). The dependent variable is a dummy variable equal to 1 if soldier *i* was recorded as having been captured in the WWI casualty lists. The explanatory variable is a dummy equal variable equal to 1 if soldier *i* was predicted to be Italian using the `textcat` classification on the full name with `ECIMCI` language profiles. Individual controls include: soldier's age as of 1914, dummies for officers and volunteers. Municipal controls include: total population, population density, male population share, share of Germans, Italians, other ethnicities, ethnic diversity (HHI), share of Catholics, Protestants, Jews, share of areable land, statutory city dummy, presence of a military garrison. Municipality fixed effects are dummies for the hometown of soldier *i*. Company fixed effects are dummies for the company-regiment-combination soldier *i* was recorded to serve in. Month and date fixed effects are dummies for the month and the date of the reported casualty or, if not available, the publication date of the respective casualty list issue. Heteroskedasticity-robust standard errors are reported in parentheses: *** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

Table 4.B.6: Individual-level results: Effect on Dead

	Dead						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Italian Soldier	-0.084***	-0.085***	-0.052***	-0.044***	-0.028***	-0.026***	-0.006
	(0.004)	(0.004)	(0.008)	(0.008)	(0.009)	(0.007)	(0.004)
Individual Controls	-	✓	✓	✓	✓	✓	✓
Municipal Controls	-	-	✓	✓	-	-	-
Municipality FE	-	-	-	✓	✓	✓	✓
Company FE	-	-	-	-	✓	✓	✓
Month FE	-	-	-	-	-	✓	-
Date FE	-	-	-	-	-	-	✓
Observations	35,260	32,141	28,193	28,180	20,917	20,880	20,581
Adjusted R^2	0.010	0.014	0.017	0.026	0.041	0.266	0.748
Dep. Var. Mean	0.211	0.209	0.210	0.210	0.169	0.169	0.160

Notes: Table reports soldier-level regressions as in Equation (4.2). The dependent variable is a dummy variable equal to 1 if soldier *i* was recorded as having died in the WWI casualty lists. The explanatory variable is a dummy equal variable equal to 1 if soldier *i* was predicted to be Italian using the `textcat` classification on the full name with `ECIMCI` language profiles. Individual controls include: soldier's age as of 1914, dummies for officers and volunteers. Municipal controls include: total population, population density, male population share, share of Germans, Italians, other ethnicities, ethnic diversity (HHI), share of Catholics, Protestants, Jews, share of areable land, statutory city dummy, presence of a military garrison. Municipality fixed effects are dummies for the hometown of soldier *i*. Company fixed effects are dummies for the company-regiment-combination soldier *i* was recorded to serve in. Month and date fixed effects are dummies for the month and the date of the reported casualty or, if not available, the publication date of the respective casualty list issue. Heteroskedasticity-robust standard errors are reported in parentheses: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1.

Table 4.B.7: Individual-level results: Effect on Wounded

	Wounded						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Italian Soldier	-0.104***	-0.110***	-0.069***	-0.053***	-0.051***	-0.040***	-0.037***
	(0.005)	(0.006)	(0.009)	(0.010)	(0.011)	(0.010)	(0.008)
Individual Controls	-	✓	✓	✓	✓	✓	✓
Municipal Controls	-	-	✓	✓	-	-	-
Municipality FE	-	-	-	✓	✓	✓	✓
Company FE	-	-	-	-	✓	✓	✓
Month FE	-	-	-	-	-	✓	-
Date FE	-	-	-	-	-	-	✓
Observations	35,260	32,141	28,193	28,180	20,917	20,880	20,581
Adjusted R^2	0.010	0.021	0.025	0.037	0.111	0.339	0.617
Dep. Var. Mean	0.492	0.493	0.487	0.487	0.545	0.544	0.550

Notes: Table reports soldier-level regressions as in Equation (4.2). The dependent variable is a dummy variable equal to 1 if soldier *i* was recorded as having been wounded in the WWI casualty lists. The explanatory variable is a dummy equal variable equal to 1 if soldier *i* was predicted to be Italian using the `textcat` classification on the full name with `ECIMCI` language profiles. Individual controls include: soldier's age as of 1914, dummies for officers and volunteers. Municipal controls include: total population, population density, male population share, share of Germans, Italians, other ethnicities, ethnic diversity (HHI), share of Catholics, Protestants, Jews, share of areable land, statutory city dummy, presence of a military garrison. Municipality fixed effects are dummies for the hometown of soldier *i*. Company fixed effects are dummies for the company-regiment-combination soldier *i* was recorded to serve in. Month and date fixed effects are dummies for the month and the date of the reported casualty or, if not available, the publication date of the respective casualty list issue. Heteroskedasticity-robust standard errors are reported in parentheses: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1.

Table 4.B.8: Individual-level results: Effect on Captured (vs. Dead)

	Captured (vs. Dead)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Italian Soldier	0.227***	0.231***	0.137***	0.116***	0.095***	0.070***	0.015**
	(0.007)	(0.007)	(0.013)	(0.014)	(0.017)	(0.012)	(0.006)
Individual Controls	-	✓	✓	✓	✓	✓	✓
Municipal Controls	-	-	✓	✓	-	-	-
Municipality FE	-	-	-	✓	✓	✓	✓
Company FE	-	-	-	-	✓	✓	✓
Month FE	-	-	-	-	-	✓	-
Date FE	-	-	-	-	-	-	✓
Observations	18,450	16,818	14,954	14,925	9,789	9,777	9,440
Adjusted R^2	0.053	0.058	0.070	0.092	0.219	0.559	0.887
Dep. Var. Mean	0.603	0.605	0.610	0.609	0.653	0.653	0.671

Notes: Table reports soldier-level regressions as in Equation (4.2). The sample only includes soldiers that were reported to have been either captured or killed during World War I. The dependent variable is a dummy variable equal to 1 if soldier *i* was recorded as having been captured in the WWI casualty lists. The explanatory variable is a dummy equal variable equal to 1 if soldier *i* was predicted to be Italian using the `textcat` classification on the full name with `ECIMCI` language profiles. Individual controls include: soldier's age as of 1914, dummies for officers and volunteers. Municipal controls include: total population, population density, male population share, share of Germans, Italians, other ethnicities, ethnic diversity (HHI), share of Catholics, Protestants, Jews, share of areable land, statutory city dummy, presence of a military garrison. Municipality fixed effects are dummies for the hometown of soldier *i*. Company fixed effects are dummies for the company-regiment-combination soldier *i* was recorded to serve in. Month and date fixed effects are dummies for the month and the date of the reported casualty or, if not available, the publication date of the respective casualty list issue. Heteroskedasticity-robust standard errors are reported in parentheses: $***$ p<0.01, $**$ p<0.05, and $*$ p<0.1.

Table 4.B.9: Robustness – Subsamples

	Captured			
	(1) Full Sample	(2) No Ladin Areas	(3) <10 Diff Bezirke	(4) <10 Diff Gerichtsbez.
Italian Soldier	0.035***	0.033***	0.021**	0.017**
	(0.006)	(0.007)	(0.008)	(0.008)
Individual Controls	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓
Company FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Observations	20,581	17,296	14,526	14,822
Adjusted R^2	0.708	0.708	0.696	0.704
Dep. Var. Mean	0.312	0.308	0.277	0.297

Notes: Table reports soldier-level regressions as in Equation (4.2). The dependent variable is a dummy variable equal to 1 if soldier *i* was recorded as having been captured in the WWI casualty lists. The explanatory variable is a dummy equal variable equal to 1 if soldier *i* was predicted to be Italian using the `textcat` classification on the full name with `ECIMCI` language profiles. Individual controls include: soldier's age as of 1914, dummies for officers and volunteers. Municipal controls include: total population, population density, male population share, share of Germans, Italians, other ethnicities, ethnic diversity (HHI), share of Catholics, Protestants, Jews, share of areable land, statutory city dummy, presence of a military garrison. Municipality fixed effects are dummies for the hometown of soldier *i*. Company fixed effects are dummies for the company-regiment-combination soldier *i* was recorded to serve in. Month and date fixed effects are dummies for the month and the date of the reported casualty or, if not available, the publication date of the respective casualty list issue. Column (1) reports the preferred specification from column (7) in Table 4.2. Column (2) excludes all districts belonging to the historic Ladin settlement area (Bihl 1980): Ampezzo, Bruneck, Cavalese, Cles, Primiero. Column (3) excludes all districts (*Bezirke*) with a residual difference between predicted and population Italian share of more than 10 percentage points as reported in Table 4.B.3: Primiero, Ampezzo, Cles, Cavalese, Borgo, Innsbruck (Stadt), Trient/Trento. Column (3) excludes all judicial districts (*Gerichtsbezirke*) with a residual difference between predicted and population Italian share of more than 10 percentage points as reported in Table 4.B.4: Enneberg, Fassa, Primiero, Kastelruth/Castelrotto, Buchenstein/Livinallongo, Fondo, Ampezzo, Cles, Borgo, Roveredo/Rovereto, Mezolombardo, Lavis, Levico, Cembra, Malè, Cavalese, Strigno, Neumarkt/Egna, Kaltern/Caldaro, Innsbruck (Stadt). Heteroskedasticity-robust standard errors are reported in parentheses: *** p<0.01, ** p<0.05, and * p<0.1.

Bibliography

ABDELGADIR, AALA and VASILIKI FOUKA (2020). "Political Secularism and Muslim Integration in the West: Assessing the Effects of the French Headscarf Ban." *American Political Science Review*, 114 (3), 707–723.

ABOU-CHADI, TARIK (2014). "Niche Party Success and Mainstream Party Policy Shifts – How Green and Radical Right Parties Differ in Their Impact." *British Journal of Political Science*, 46 (2), 417–436.

ABOU-CHADI, TARIK and WERNER KRAUSE (2020). "The Causal Effect of Radical Right Success on Mainstream Parties' Policy Positions: A Regression Discontinuity Approach." *British Journal of Political Science*, 50 (3), 829–847.

ABRAMITZKY, RAN, ADELINE DELAVANDE, and LUIS VASCONCELOS (2011). "Marrying Up: The Role of Sex Ratio in Assortative Matching." *American Economic Journal: Applied Economics*, 3 (3), 124–157.

ACEMOGLU, DARON, DAVID H. AUTOR, and DAVID LYLE (2004). "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury." *Journal of Political Economy*, 112 (3), 497–551.

ACEMOGLU, DARON, GIUSEPPE de FEO, GIACOMO de LUCA, and GIANLUCA RUSSO (2022). "War, Socialism, and the Rise of Fascism: an Empirical Exploration." *The Quarterly Journal of Economics*, 137 (2), 1233–1296.

AGER, PHILIPP, LEONARDO BURSZTYN, LUKAS LEUCHT, and HANS-JOACHIM VOTH (2022). "Killer Incentives: Rivalry, Performance and Risk-Taking among German Fighter Pilots, 1939–45." *The Review of Economic Studies*, 89 (5), 2257–2292.

AKERLOF, GEORGE A. and RACHEL E. KRANTON (2000). "Economics and Identity." *The Quarterly Journal of Economics*, 115 (3), 715–753.

— (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*. Princeton, NJ: Princeton University Press.

ALATAS, VIVI, ARUN CHANDRASEKHAR, MARKUS MOBIUS, BENJAMIN OLKEN, and CINDY PALADINES (2022). "Designing Effective Celebrity Messaging: Results from a Nationwide Twitter Experiment Promoting Vaccination in Indonesia." *Working Paper*.

ALBA, BEATRICE, DORIS MCILWAIN, LADD WHEELER, and MICHAEL P. JONES (2014). "Status Consciousness." *Journal of Individual Differences*, 35 (3), 166–176.

ALBORNOZ, FACUNDO, JAKE BRADLEY, and SILVIA SONDEREGGER (2020). "The Brexit referendum and the rise in hate crime: Conforming to the new norm." *CeDEx Discussion Paper*, No. 2020-12.

ALESINA, ALBERTO, ARMANDO MIANO, and STEFANIE STANTCHEVA (2023). "Immigration and Redistribution." *The Review of Economic Studies*, 90 (1), 1–39.

ALESINA, ALBERTO and BRYONY REICH (2015). "Nation Building." *Working Paper*.

ALESINA, ALBERTO, BRYONY REICH, and ALESSANDRO RIBONI (2020). "Nation-building, nationalism, and wars." *Journal of Economic Growth*, 25 (4), 381–430.

ALEXANDER, HELMUT, STEFAN LECHNER, and ADOLF LEIDLMAIR (1993). *Heimatlos: Die Umsiedlung der Südtiroler.* Wien: Deuticke.

ALLPORT, GORDON W. (1954). *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley.

ALSAN, MARCELLA and SARAH EICHMEYER (2021). "Experimental Evidence on the Effectiveness of Non-Experts for Improving Vaccine Demand." *NBER Working Paper*, No. 28593.

ANDERSON, BENEDICT (1983). *Imagined communities: Reflections on the origin and spread of nationalism*. London: Verso.

ANDREONI, JAMES (1989). "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy*, 97 (6), 1447–1458.

ANDREONI, JAMES and RAGAN PETRIE (2004). "Public Goods Experiments without Confidentiality: A Glimpse into Fund-Raising." *Journal of Public Economics*, 88 (7-8), 1605–1623.

ARBAK, EMRAH and MARIE-CLAIRE VILLEVAL (2013). "Voluntary Leadership: Motivation and Influence." *Social Choice and Welfare*, 40 (3), 635–662.

ARENDT, HANNAH (1962). *The Origins of Totalitarianism*. Cleveland, New York: The World Publishing Company.

ARON, ARTHUR, ELAINE N. ARON, and DANNY SMOLLAN (1992). "Inclusion of Other in the Self Scale and the structure of interpersonal closeness." *Journal of Personality and Social Psychology*, 63 (4), 596–612.

ARZHEIMER, KAI (2015). "The AfD: Finally a Successful Right-Wing Populist Eurosceptic Party for Germany?" *West European Politics*, 38 (3), 535–556.

ARZHEIMER, KAI and CARL C. BERNING (2019). "How the Alternative for Germany (AfD) and their voters veered to the radical right, 2013–2017." *Electoral Studies*, 60, 102040.

ATKIN, DAVID, EVE COLSON-SIHRA, and MOSES SHAYO (2021). "How Do We Choose Our Identity? A Revealed Preference Approach Using Food Consumption." *Journal of Political Economy*, 129 (4), 1193–1251.

ATZPODIEN, DANA SIOBHAN (2022). "Party Competition in Migration Debates: The Influence of the AfD on Party Positions in German State Parliaments." *German Politics*, 31 (3), 381–398.

AUTONOME PROVINZ BOZEN (2022). *Geoportal Südtirol*. URL: https://geoportal.buergernetz.bz.it/ (last accessed on 03/05/2023).

AUTONOMOUS PROVINCE OF SOUTH TYROL PROVINCIAL STATISTICS INSTITUTE (2021). *South Tyrol in Figures 2021*. URL: https://astat.provinz.bz.it/downloads/Siz_2021-eng(4).pdf (last accessed on 01/28/2023).

BALE, TIM, CHRISTOFFER GREEN-PEDERSEN, ANDRÉ KROUWEL, KURT RICHARD LUTHER, and NICK SITTER (2009). "If you can't Beat them, Join them? Explaining Social Democratic Responses to the Challenge from the Populist Radical Right in Western Europe." *Political Studies*, 58 (3), 410–426.

BANERJEE, ABHIJIT VINAYAK, ESTHER DUFLO, RACHEL GLENNERSTER, and DHRUVA KOTHARI (2010). "Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives." *BMJ*, 340:c2220.

BARTOŠ, VOJTĚCH, MICHAL BAUER, JANA CAHLÍKOVÁ, and JULIE CHYTILOVÁ (2022). "Communicating doctors' consensus persistently increases COVID-19 vaccinations." *Nature*, 606, 542–549.

BAUER, MICHAL, CHRISTOPHER BLATTMAN, JULIE CHYTILOVÁ, JOSEPH HENRICH, EDWARD MIGUEL, and TAMAR MITTS (2016). "Can War Foster Cooperation?" *Journal of Economic Perspectives*, 30 (3), 249–274.

BAUER, MICHAL, JANA CAHLÍKOVÁ, JULIE CHYTILOVÁ, GÉRARD ROLAND, and TOMÁŠ ŽELINSKÝ (2023). "Shifting Punishment on Minorities: Experimental Evidence of Scapegoating." *The Economic Journal*.

BAYERISCHES LANDESAMT FÜR STATISTIK (2019). *Demographie-Spiegel für bayerische Gemeinden bis zum Jahr 2037 veröffentlicht*. URL: https://www.e-statistik.eu/presse/mitteilungen/2019/pm181/index.html (last accessed on 07/25/2022).

— (2022). *Arbeitslose (Jahresdurchschnitt): Gemeinden, ausgewählte Personengruppen, Jahr (code: 13211-004r)*. URL: https://www.statistikdaten.bayern.de/genesis/online (last accessed on 08/16/2022).

BAZZI, SAMUEL, MASYHUR HILMY, and BENJAMIN MARX (2022). "Religion, Education, and Development." *NBER Working Paper*, No. 27073.

BEATTON, TONY, AHMED SKALI, and BENNO TORGLER (2019). "Protestantism and Effort Expenditure on the Battlefield: Soldier-Level Evidence from World War II." *Working Paper*.

BECK KNUDSEN, ANNE SOFIE (2022). "Those Who Stayed: Selection and Cultural Change in the Age of Mass Migration." *Working Paper*.

BECKER, SASCHA O. (2022). "Forced displacement in history: Some recent research." *Australian Economic History Review*, 62 (1), 2–25.

BECKER, SASCHA O., KATRIN BOECKH, CHRISTA HAINZ, and LUDGER WOESSMANN (2016). "The Empire is Dead, Long Live the Empire! Long–Run Persistence of Trust and Corruption in the Bureaucracy." *The Economic Journal*, 126 (590), 40–74.

BECKER, SASCHA O. and ANDREAS FERRARA (2019). "Consequences of forced migration: A survey of recent findings." *Labour Economics*, 59, 1–16.

BECKER, SASCHA O., IRENA GROSFELD, PAULINE GROSJEAN, NICO VOIGTLÄNDER, and EKATERINA ZHURAVSKAYA (2020). "Forced Migration and Human Capital: Evidence from Post-WWII Population Transfers." *American Economic Review*, 110 (5), 1430–1463.

BECKER, SASCHA O., VOLKER LINDENTHAL, SHARUN MUKAND, and FABIAN WALDINGER (2022). "Persecution and Escape: Professional Networks and High-Skilled Emigration from Nazi Germany." *Working Paper*.

BELMONTE, ALESSANDRO (2022). "Punishing or rallying 'round the flag? Heterogeneous effects of terrorism in South Tyrol." *Kyklos*, 75 (4), 536–563.

BELMONTE, ALESSANDRO and ARMANDO DI LILLO (2021). "Backlash against affirmative action: Evidence from the South Tyrolean package." *European Economic Review*, 137, 103802.

BÉNABOU, ROLAND and JEAN TIROLE (2006). "Incentives and Prosocial Behavior." *American Economic Review*, 96 (5), 1652–1678.

— (2011). "Identity, Morals, and Taboos: Beliefs as Assets." *The Quarterly Journal of Economics*, 126 (2), 805–855.

BENNETT, JOEL B. (1988). "Power and Influence as Distinct Personality Traits: Development and Validation of a Psychometric Measure." *Journal of Research in Personality*, 22 (3), 361–394.

BENNHOLD, KATRIN and MELISSA EDDY (2019). "'Hitler or Höcke?' Germany's Far-Right Party Radicalizes." *The New York Times*. URL: https://www.nytimes.com/2019/10/26/world/europe/afd-election-east-germany-hoecke.html (last accessed on 03/03/2023).

BESLEY, TIMOTHY and TORSTEN PERSSON (2018). "Democratic Values and Institutions." *Working Paper*.

— (2021). "The Rise of Identity Politics: Policy, Political Organization, and Nationalist Dynamics." *Working Paper*.

BETSCH, CORNELIA, LOTHAR H. WIELER, and KATRINE HABERSAAT (2020). "Monitoring behavioural insights related to COVID-19." *The Lancet*, 395 (10232), 1255–1256.

BIHL, WOLFDIETER (1980). "Notizen zu den ethnischen und religiösen Splitter-, Rest- und Sondergruppen in den Habsburgischen Ländern." *Die Völker des Reichs*. Ed. by ADAM WANDRUSZKA and PETER URBANITSCH. Die Habsburgermonarchie, Vol. III. Wien: Verlag der Österreichischen Akademie der Wissenschaften, 949–974.

BLASNIK, MICHAEL (2010). "RECLINK: Stata module to probabilistically match records." *Statistical Software Components*, S456876, Boston College Department of Economics.

BONOMI, GIAMPAOLO, NICOLA GENNAIOLI, and GUIDO TABELLINI (2021). "Identity, Beliefs, and Political Conflict." *The Quarterly Journal of Economics*, 136 (4), 2371–2411.

BREWER, NOEL T., GRETCHEN B. CHAPMAN, ALEXANDER J. ROTHMAN, JULIE LEASK, and ALLISON KEMPE (2017). "Increasing Vaccination: Putting Psychological Science Into Action." *Psychological Science in the Public Interest*, 18 (3), 149–207.

BREYER, MAGDALENA (2022). "Populist positions in party competition: Do parties strategically vary their degree of populism in reaction to vote and office loss?" *Party Politics*, forthcoming.

BUGGLE, JOHANNES, THIERRY MAYER, SEYHUN ORCAN SAKALLI, and MATHIAS THOENIG (2023). "The Refugee's Dilemma: Evidence from Jewish Migration out of Nazi Germany." *The Quarterly Journal of Economics*, forthcoming.

BUNDESAMT FÜR VERFASSUNGSSCHUTZ (2020). *Bundesamt für Verfassungsschutz stuft AfD-Teilorganisation "Der Flügel" als gesichert rechtsextremistische Bestrebung ein*. URL: https://tinyurl.com/bdhppbux (last accessed on 03/03/2023).

— (2023). *Bundesamt für Verfassungsschutz wins lawsuit before the Administrative Court in Cologne against AfD*. URL: https://www.verfassungsschutz.de/SharedDocs/pressemitteilungen/EN/2022/press-release-2022-1-afd-1.html (last accessed on 03/03/2023).

BUNDESWAHLLEITER (2022). *Der Bundeswahlleiter*. URL: https://www.bundeswahlleiter.de/ (last accessed on 03/03/2023).

BURSZTYN, LEONARDO, GEORGY EGOROV, and STEFANO FIORIN (2020). "From Extreme to Mainstream: The Erosion of Social Norms." *American Economic Review*, 110 (11), 3522–3548.

BURSZTYN, LEONARDO, ALESSANDRA L. GONZÁLEZ, and DAVID YANAGIZAWA-DROTT (2020). "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia." *American Economic Review*, 110 (10), 2997–3029.

BURSZTYN, LEONARDO and ROBERT JENSEN (2015). "How Does Peer Pressure Affect Educational Investments?" *The Quarterly Journal of Economics*, 130 (3), 1329–1367.

— (2017). "Social Image and Economic Behavior in the Field: Identifying, Understanding, and Shaping Social Pressure." *Annual Review of Economics*, 9 (1), 131–153.

BURSZTYN, LEONARDO, AAKAASH RAO, CHRISTOPHER ROTH, and DAVID YANAGIZAWA-DROTT (2022). "Opinions as Facts." *The Review of Economic Studies*, forthcoming.

CAMPANTE, FILIPE and DAVID YANAGIZAWA-DROTT (2016). "The Intergenerational Transmission of War." *Working Paper*.

CANTONI, DAVIDE, YUYU CHEN, DAVID Y. YANG, NOAM YUCHTMAN, and Y. JANE ZHANG (2017). "Curriculum and Ideology." *Journal of Political Economy*, 125 (2), 338–392.

CANTONI, DAVIDE, FELIX HAGEMEISTER, and MARK WESTCOTT (2020). "Persistence and Activation of Right-Wing Political Ideology." *Working Paper*.

CANTONI, DAVIDE, LOUIS-JONAS HEIZLSPERGER, DAVID Y. YANG, NOAM YUCHTMAN, and Y. JANE ZHANG (2022). "The Fundamental Determinants of Protest Participation: Evidence from Hong Kong's Antiauthoritarian Movement." *Journal of Public Economics*, 211, 104667.

CANTONI, DAVIDE, DAVID Y. YANG, NOAM YUCHTMAN, and Y. JANE ZHANG (2019). "Protests as Strategic Games: Experimental Evidence from Hong Kong's Antiauthoritarian Movement." *The Quarterly Journal of Economics*, 134 (2), 1021–1077.

CAPPELEN, ALEXANDER W., BJØRN-ATLE REME, ERIK Ø. SØRENSEN, and BERTIL TUNGODDEN (2016). "Leadership and Incentives." *Management Science*, 62 (7), 1944–1953.

CAPRETTINI, BRUNO and HANS-JOACHIM VOTH (2022). "New Deal, New Patriots: How 1930s Government Spending Boosted Patriotism During World War II." *The Quarterly Journal of Economics*, 138 (1), 465–513.

CARVALHO, JEAN-PAUL, MARK KOYAMA, and COLE WILLIAMS (2022). "Resisting Education." *Working Paper*.

CHARNESS, GARY and YAN CHEN (2020). "Social Identity, Group Behavior, and Teams." *Annual Review of Economics*, 12 (1), 691–713.

CHEN, DANIEL L. (2017). "The Deterrent Effect of the Death Penalty: Evidence from British Commutations during World War I." *Working Paper*.

CHEN, DANIEL L., MARTIN SCHONGER, and CHRIS WICKENS (2016). "oTree – An open-source platform for laboratory, online, and field experiments." *Journal of Behavioral and Experimental Finance*, 9 (1), 88–97.

CHOI, JUNG-KYOO and SAMUEL BOWLES (2007). "The Coevolution of Parochial Altruism and War." *Science*, 318 (5850), 636–640.

CIALDINI, ROBERT B., STEPHANIE L. BROWN, BRIAN P. LEWIS, CAROL LUCE, and STEVEN L. NEUBERG (1997). "Reinterpreting the empathy–altruism relationship: When one into one equals oneness." *Journal of Personality and Social Psychology*, 73 (3), 481–494.

CLEMENS, MICHAEL A. (2021). "Violence, development, and migration waves: Evidence from Central American child migrant apprehensions." *Journal of Urban Economics*, 124, 103355.

CLOTS–FIGUERAS, IRMA and PAOLO MASELLA (2013). "Education, Language and Identity." *The Economic Journal*, 123 (570), 332–357.

COLE, LAURENCE (2014). *Military Culture and Popular Patriotism in Late Imperial Austria*. Oxford: Oxford University Press.

— (2018). "Questions of Nationalization in the Habsburg Monarchy." *Nations, Identities and the First World War*. Ed. by NICO WOUTERS and LAURENCE VAN YPERSELE. London: Bloomsbury, 115–134.

CONZO, PIERLUIGI and FRANCESCO SALUSTRI (2015). "A war is forever: The long-run effects of early exposure to World War II on trust." *European Economic Review*, 120, 103313.

CORSINI, UMBERTO (1980). "Die Italiener." *Die Völker des Reichs*. Ed. by ADAM WANDRUSZKA and PETER URBANITSCH. Die Habsburgermonarchie, Vol. III. Wien: Verlag der Österreichischen Akademie der Wissenschaften, 839–879.

COSMO (2021). *COSMO – COVID-19 Snapshot Monitoring – Wave 41, April 2021*. URL: https://projekte.uni-erfurt.de/cosmo2020/web/ (last accessed on 06/25/2022).

COSTA, DORA L. and MATTHEW E. KAHN (2003). "Cowards and Heroes: Group Loyalty in the American Civil War." *The Quarterly Journal of Economics*, 118 (2), 519–548.

— (2008). *Heroes and Cowards: The Social Face of War*. Princeton: Princeton University Press.

DAHIS, RICARDO, EMILY NIX, and NANCY QIAN (2020). "Choosing Racial Identity in the United States, 1880-1940." *Working Paper*.

DAHLSTRÖM, CARL and ANDERS SUNDELL (2012). "A losing gamble. How mainstream parties facilitate anti-immigrant party success." *Electoral Studies*, 31 (2), 353–363.

DAI, HENGCHEN, SILVIA SACCARDO, MARIA A. HAN, LILY ROH, NAVEEN RAJA, SITARAM VANGALA, HARDIKKUMAR MODI, SHITAL PANDYA, MICHAEL SLOYAN, and DANIEL M. CROYMANS (2021). "Behavioural nudges increase COVID-19 vaccinations." *Nature*, 597 (7876), 404–409.

DANNENBERG, ASTRID (2015). "Leading by example versus leading by words in voluntary contribution experiments." *Social Choice and Welfare*, 44 (1), 71–85.

DE JUAN, ALEXANDER, FELIX HAASS, CARLO KOOS, SASCHA RIAZ, and THOMAS TICHELBAECKER (2022). "War and Nationalism: How WW1 battle deaths fueled civilians' support for the Nazi Party." *American Political Science Review*, forthcoming.

DEÁK, ISTVÁN (1990). *Beyond Nationalism: A Social and Political History of the Habsburg Officer Corps, 1848-1918*. New York: Oxford University Press.

DEHDARI, SIRUS H. and KAI GEHRING (2022). "The Origins of Common Identity: Evidence from Alsace-Lorraine." *American Economic Journal: Applied Economics*, 14 (1), 261–292.

DELL, MELISSA and PABLO QUERUBIN (2018). "Nation Building Through Foreign Intervention: Evidence from Discontinuities in Military Strategies." *The Quarterly Journal of Economics*, 133 (2), 701–764.

DELLAVIGNA, STEFANO, JOHN A. LIST, and ULRIKE MALMENDIER (2012). "Testing for Altruism and Social Pressure in Charitable Giving." *The Quarterly Journal of Economics*, 127 (1), 1–56.

DELLAVIGNA, STEFANO, JOHN A. LIST, ULRIKE MALMENDIER, and GAUTAM RAO (2017). "Voting to Tell Others." *The Review of Economic Studies*, 84 (1), 143–181.

DEPETRIS-CHAUVIN, EMILIO, RUBEN DURANTE, and FILIPE CAMPANTE (2020). "Building Nations through Shared Experiences: Evidence from African Football." *American Economic Review*, 110 (5), 1572–1602.

DEUTSCHER BUNDESTAG (2020). *Bundestag ändert Geschäftsordnung wegen Coronavirus*. URL: https://www.bundestag.de/dokumente/textarchiv/2020/kw13-de-corona-geschaeftsordnung-689124 (last accessed on 03/03/2023).

— (2022a). *Deutscher Bundestag - Ausschüsse*. URL: https://www.bundestag.de/webarchiv/Ausschuesse (last accessed on 03/03/2023).

— (2022b). *Parlamentsbegriffe - Sitzungswoche*. URL: https://www.bundestag.de/services/glossar/glossar/S/sitzungswochen-247330 (last accessed on 03/03/2023).

— (1954-2017). *Amtliches Handbuch des Deutschen Bundestages*. Darmstadt: Neue Darmstädter Verlagsanstalt.

DI MICHELE, ANDREA (2008). *Die unvollkommene Italianisierung: Politik und Verwaltung in Südtirol 1918 - 1943*. Innsbruck: Wagner.

— (2018). "'Italiani d'Austria'. Italienischsprachige Soldaten der Habsburgermonarchie im Ersten Weltkrieg." *Minderheiten-Soldaten*. Ed. by OSWALD ÜBEREGGER. Paderborn: Ferdinand Schöningh, 45–68.

— (2020). *Soldaten zwischen zwei Uniformen. Österreichische Italiener im Ersten Weltkrieg.* Wien: Böhlau.

DINAS, ELIAS, VASILIKI FOUKA, and ALAIN SCHLÄPFER (2021a). "Family History and Attitudes toward Out-Groups: Evidence from the European Refugee Crisis." *The Journal of Politics*, 83 (2), 647–661.

— (2021b). "Recognition of Collective Victimhood and Outgroup Prejudice." *Public Opinion Quarterly*, 85 (2), 517–538.

DIPPEL, CHRISTIAN and ANDREAS FERRARA (2022). "Leadership and Team Cohesion under Flying Bullets: Captains and their Companies in the Civil War." *Working Paper*.

DJOURELOVA, MILENA (2023). "Persuasion through Slanted Language: Evidence from the Media Coverage of Immigration." *American Economic Review*, 113 (3), 800–835.

DROUVELIS, MICHALIS and DANIELE NOSENZO (2013). "Group identity and leading-by-example." *Journal of Economic Psychology*, 39, 414–425.

DUPRAZ, YANNICK and ANDREAS FERRARA (2023). "Fatherless: The Long-Term Effects of Losing a Father in the U.S. Civil War." *Journal of Human Resources*, forthcoming.

EHNL, MAXIMILIAN (1934). *Die österreichisch-ungarische Landmacht nach Aufbau, Gliederung, Friedensgarnison, Einteilung und nationaler Zusammensetzung im Sommer*

1914: Ergänzungsheft 9 zum Werke "Österreich-Ungarns letzter Krieg". Wien: Verlag der Militärwissenschaftlichen Mitteilungen.

ENDERSTAM, DÉSIRÉE (2020). "Sprechen Sie rechts? Rhetorik und Sprachgebrauch des Nationalsozialismus in politischen Reden des AfD-Politikers Björn Höcke." Undergraduate Thesis. Dalarna University. URL: http://urn.kb.se/resolve?urn=urn%3Anbn%3Ase%3Adu-32197 (last accessed on 03/03/2023).

ENGLER, SARAH, THERESA GESSLER, TARIK ABOU-CHADI, and LUCAS LEEMANN (2022). "Democracy challenged: how parties politicize different democratic principles." *Journal of European Public Policy*, forthcoming.

ERFURTH, PHILIPP EMANUEL (2022). "Is the European Union More Unequal Than the Habsburg Empire? Examining Regional Inequalities in Habsburg Regions From 1870 to 2018." *Working Paper*.

ESGUERRA, EMILIO, LEONHARD VOLLMER, and JOHANNES WIMMER (2021). "Anticipated Peer Effects: An Application to COVID-19 Vaccinations." *AEA RCT Registry*, April 12. URL: https://doi.org/10.1257/rct.7437 (last accessed on 03/03/2023).

— (2023). "Influence Motives in Social Signaling: Evidence from COVID-19 Vaccinations in Germany." *American Economic Review: Insights*, 5 (2), 275–291.

ESTEBAN, JOAN and DEBRAJ RAY (2008). "On the Salience of Ethnic Conflict." *American Economic Review*, 98 (5), 2185–2202.

— (2011). "A Model of Ethnic Conflict." *Journal of the European Economic Association*, 9 (3), 496–521.

EUROPEANA 1914-1918 (2011). *"Viel Feind, viel Ehr ... und Gott mit Euch". Postkarte an die Front von Therese Holub, 1915.* URL: https://www.europeana.eu/item/2020601/https___1914_1918_europeana_eu_contributions_1227 (last accessed on 02/03/2023).

FALK, ARMIN, ANKE BECKER, THOMAS DOHMEN, BENJAMIN ENKE, DAVID HUFFMAN, and UWE SUNDE (2018). "Global Evidence on Economic Preferences." *The Quarterly Journal of Economics*, 133 (4), 1645–1692.

FELDKAMP, MICHAEL F. (2018). "Deutscher Bundestag 1998 bis 2017/18: Parlaments- und Wahlstatistik für die 14. bis beginnende 19. Wahlperiode." *Zeitschrift für Parlamentsfragen*, 49 (2), 207–222.

FERRARA, ANDREAS (2022). "World War II and Black Economic Progress." *Journal of Labor Economics*, 40 (4), 1053–1091.

FERRARA, ANDREAS and PRICE FISHBACK (2022). "Discrimination, Migration, and Economic Outcomes: Evidence from World War I." *Review of Economics and Statistics*, forthcoming.

FIVA, JON H., ODA NEDREGÅRD, and HENNING ØIEN (2021). "Polarization in Parliamentary Speech." *CESifo Working Paper*, No. 8818.

FOERSTER, MANUEL and JOËL J. VAN DER WEELE (2021). "Casting Doubt: Image Concerns and the Communication of Social Impact." *The Economic Journal*, 131 (639), 2887–2919.

FOUKA, VASILIKI (2019). "How Do Immigrants Respond to Discrimination? The Case of Germans in the US During World War I." *American Political Science Review*, 113 (2), 405–422.

— (2020). "Backlash: The Unintended Effects of Language Prohibition in U.S. Schools after World War I." *The Review of Economic Studies*, 87 (1), 204–239.

FOUKA, VASILIKI, SOUMYAJIT MAZUMDER, and MARCO TABELLINI (2022). "From Immigrants to Americans: Race and Assimilation during the Great Migration." *The Review of Economic Studies*, 89 (2), 811–842.

FOUKA, VASILIKI and HANS-JOACHIM VOTH (2022). "Collective Remembrance and Private Choice: German–Greek Conflict and Behavior in Times of Crisis." *American Political Science Review*, 1–20.

GÄCHTER, SIMON, DANIELE NOSENZO, ELKE RENNER, and MARTIN SEFTON (2012). "Who makes a good leader? Cooperativeness, Optimism, and Leading-by-example." *Economic Inquiry*, 50 (4), 953–967.

GÄCHTER, SIMON and ELKE RENNER (2018). "Leaders as role models and 'belief managers' in social dilemmas." *Journal of Economic Behavior & Organization*, 154, 321–334.

GÄCHTER, SIMON, CHRIS STARMER, and FABIO TUFANO (2015). "Measuring the Closeness of Relationships: A Comprehensive Evaluation of the 'Inclusion of the Other in the Self' Scale." *PLoS One*, 10 (6), e0129478.

GALLETTA, SERGIO and TOMMASO GIOMMONI (2023). "War Violence Exposure and Tax Compliance." *Working Paper*.

GARCÍA-JIMENO, CAMILO, ANGEL IGLESIAS, and PINAR YILDIRIM (2022). "Information Networks and Collective Action: Evidence from the Women's Temperance Crusade." *American Economic Review*, 112 (1), 41–80.

GATTERER, CLAUS (1972). *Erbfeindschaft Italien - Österreich*. München: Europaverlag.

GENNAIOLI, NICOLA and HANS-JOACHIM VOTH (2015). "State Capacity and Military Conflict." *The Review of Economic Studies*, 82 (4), 1409–1448.

GENTZKOW, MATTHEW, BRYAN KELLY, and MATT TADDY (2019). "Text as Data." *Journal of Economic Literature*, 57 (3), 535–574.

GENTZKOW, MATTHEW, JESSE M. SHAPIRO, and MATT TADDY (2019). "Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech." *Econometrica*, 87 (4), 1307–1340.

GESIS – LEIBNIZ-INSTITUT FÜR SOZIALWISSENSCHAFTEN (2019). *ALLBUS/GGSS 2018 (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften/German General Social Survey 2018)*.

GESSLER, THERESA and SOPHIA HUNGER (2021). "How the refugee crisis and radical right parties shape party competition on immigration." *Political Science Research and Methods*, 10 (3), 524–544.

GILES, HOWARD and TANIA OGAY (2007). "Communication Accommodation Theory." *Explaining communication: Contemporary theories and exemplars*. Ed. by BRYAN B. WHALEY and WENDY SAMTER. New York: Routledge, 293–310.

GLAISE VON HORSTENAU, EDMUND, ed. (1932). *Österreich-Ungarns letzter Krieg 1914-1918*. Wien: Verlag der Militärwissenschaftlichen Mitteilungen.

GOETTE, LORENZ and EGON TRIPODI (2021). "Social Influence in Prosocial Behavior: Evidence from a Large-Scale Experiment." *Journal of the European Economic Association*, 19 (4), 2373–2398.

GREENE, DEREK and JAMES P. CROSS (2015). "Unveiling the Political Agenda of the European Parliament Plenary." *Proceedings of the ACM Web Science Conference*. Ed. by DAVID de ROURE. ACM Digital Library. New York: ACM, 1–10.

GROSSMAN, GENE M. and ELHANAN HELPMAN (2021). "Identity Politics and Trade Policy." *The Review of Economic Studies*, 88 (3), 1101–1126.

GROSSMAN, GUY, DEVORAH MANEKIN, and DAN MIODOWNIK (2015). "The Political Legacies of Combat: Attitudes Toward War and Peace Among Israeli Ex-Combatants." *International Organization*, 69 (4), 981–1009.

GROTE, GEORG (2012). *The South Tyrol Question, 1866–2010*. Bern: Peter Lang.

GRÜNDL, JOHANN (2020a). *multidictR: R package*. URL: https://github.com/jogrue/multidictR (last accessed on 03/03/2023).

— (2020b). *popdictR: R package*. URL: https://github.com/jogrue/popdictR (last accessed on 03/03/2023).

— (2020c). *regexhelpeR: R package*. URL: https://github.com/jogrue/regexhelpeR (last accessed on 03/03/2023).

— (2022). "Populist ideas on social media: A dictionary-based measurement of populist communication." *New Media & Society*, 24 (6), 1481–1499.

GURCIULLO, STEFANO et al. (2015). "Complex Politics: A Quantitative Semantic and Topological Analysis of UK House of Commons Debates." *Working Paper*.

GURIEV, SERGEI and ELIAS PAPAIOANNOU (2022). "The Political Economy of Populism." *Journal of Economic Literature*, 60 (3), 753–832.

GÜTH, WERNER, M. VITTORIA LEVATI, MATTHIAS SUTTER, and ELINE VAN DER HEIJDEN (2007). "Leading by example with and without exclusion power in voluntary contribution experiments." *Journal of Public Economics*, 91 (5), 1023–1042.

HAALAND, INGAR, CHRISTOPHER ROTH, and JOHANNES WOHLFART (2023). "Designing Information Provision Experiments." *Journal of Economic Literature*, 61 (1), 3–40.

HAGEMEISTER, FELIX (2022). "Populism and Propagation of Far-Right Extremism." *European Journal of Political Economy*, 72, 102116.

HAGER, ANSELM and HANNO HILBIG (2020). "Does Public Opinion Affect Political Speech?" *American Journal of Political Science*, 64 (4), 921–937.

HAIDT, JONATHAN (2013). *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York: Vintage Books.

HAIGNER, STEFAN D. and FLORIAN WAKOLBINGER (2010). "To lead or not to lead: Endogenous sequencing in public goods games." *Economics Letters*, 108 (1), 93–95.

HAN, KYUNG JOON (2014). "The Impact of Radical Right-Wing Parties on the Positions of Mainstream Parties Regarding Multiculturalism." *West European Politics*, 38 (3), 557–576.

HANSEN, MICHAEL A. and JONATHAN OLSEN (2018). "Flesh of the Same Flesh: A Study of Voters for the Alternative for Germany (AfD) in the 2017 Federal Election." *German Politics*, 28 (1), 1–19.

— (2022). "The Alternative for Germany (AfD) as Populist Issue Entrepreneur: Explaining the Party and its Voters in the 2021 German Federal Election." *German Politics*, forthcoming.

HARRIS, CHARLES R. MILLMAN et al. (2020). "Array programming with NumPy." *Nature*, 585 (7825), 357–362.

HEINZE, ANNA-SOPHIE (2022). "Dealing with the populist radical right in parliament: mainstream party responses toward the Alternative for Germany." *European Political Science Review*, 14 (3), 333–350.

HENRICH, JOSEPH (2020). *The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*. New York: Farrar Straus and Giroux.

HERMALIN, BENJAMIN E. (1998). "Toward an Economic Theory of Leadership: Leading by Example." *American Economic Review*, 88 (5), 1188–1206.

HJORTH, FREDERIK and MARTIN VINÆS LARSEN (2020). "When Does Accommodation Work? Electoral Effects of Mainstream Left Position Taking on Immigration." *British Journal of Political Science*, 52 (2), 949–957.

HOBSBAWM, ERIC J. (1992). *Nations and Nationalism since 1780: Programme, Myth, Reality*. 2nd edition. Cambridge: Cambridge University Press.

HORNIK, KURT, PATRICK MAIR, JOHANNES RAUCH, WILHELM GEIGER, CHRISTIAN BUCHTA, and INGO FEINERER (2013). "The textcat Package for n -Gram Based Text Categorization in R." *Journal of Statistical Software*, 52 (6), 1–17.

HORNIK, KURT, JOHANNES RAUCH, CHRISTIAN BUCHTA, and INGO FEINERER (2020). *R Package 'textcat': Version 1.0-7*. URL: https://cran.r-project.org/web/packages/textcat/index.html (last accessed on 03/08/2023).

HUTCHINSON, JOHN (2017). *Nationalism and War*. Oxford: Oxford University Press.

INDACOCHEA, DANIEL (2019). "A Farewell to Army Segregation: The Effects of Racial Integration during the Korean War." *Working Paper*.

ISMAYR, WOLFGANG (2001). *Der Deutsche Bundestag*. Wiesbaden: VS Verlag für Sozialwissenschaften.

JAGADITS, ANDRÁS (2021). "Emigration and Local Structural Change. Evidence from (Austria-)Hungary in the Age of Mass Migration." *Working Paper*.

JIA, RUIXUE and TORSTEN PERSSON (2021). "Choosing Ethnicity: The Interplay Between Individual and Social Motives." *Journal of the European Economic Association*, 19 (2), 1203–1248.

K. K. STATISTISCHE ZENTRALKOMMISSION (1907). *Gemeindelexikon der im Reichsrate vertretenen Königreiche und Länder: VIII. Tirol und Vorarlberg*. Wien: Druck und Verlag der k. k. Hof- und Staatsdruckerei.

— (1912). "Die Ergebnisse der Volkszählung vom 31. Dezember 1910 in den im Reichsrathe vertretenen Königreichen und Ländern: Die summarischen Ergebnisse der Volkszählung." *Österreichische Statistik, Neue Folge*, 1 (1).

K. U. K. KRIEGSMINISTERIUM (1911). *Militärstatistisches Jahrbuch für das Jahr 1910*. Wien: Druck und Verlag der k. k. Hof- und Staatsdruckerei.

KALLIS, ARISTOTLE (2002). *Fascist Ideology: Territory and Expansionism in Italy and Germany, 1922-1945*. London: Routledge.

KARING, ANNE (2023). "Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone." *Working Paper*.

KARLAN, DEAN and MARGARET A. MCCONNELL (2014). "Hey look at me: The effect of giving circles on giving." *Journal of Economic Behavior & Organization*, 106, 402–412.

KELLY, BRYAN, DIMITRIS PAPANIKOLAOU, AMIT SERU, and MATT TADDY (2021). "Measuring Technological Innovation over the Long Run." *American Economic Review: Insights*, 3 (3), 303–320.

KENNAN, GEORGE F. (1981). *The Decline of Bismarck's European Order: Franco-Russian Relations 1875-1890*. Princeton, NJ: Princeton University Press.

KERSTING, FELIX and NIKOLAUS WOLF (2021). "On the origins of national identity. German nation-building after Napoleon." *CEPR Discussion Paper*, No. 16314.

KIBRIS, ARZU and RESUL CESUR (2022). "War Does not Foster Cooperation when Neoclassical Mechanisms are Absent! Evidence from a Natural Experiment among Turkish Conscripts." *NBER Working Paper*, No. 30674.

KOENIG, CHRISTOPH (2023). "Loose Cannons: War Veterans and the Erosion of Democracy in Weimar Germany." *The Journal of Economic History*, 83 (1), 167–202.

KOMLOS, JOHN (1989). *Nutrition and Economic Development in the Eighteenth-Century Habsburg Monarchy: An Anthropometric History*. Princeton, NJ: Princeton University Press.

KRAUSE, WERNER, DENIS COHEN, and TARIK ABOU-CHADI (2023). "Does accommodation work? Mainstream party strategies and the success of radical right parties." *Political Science Research and Methods*, 11 (1), 1–8.

LEWANDOWSKY, MARCEL, JULIA SCHWANHOLZ, CHRISTOPH LEONHARDT, and ANDREAS BLÄTTE (2022). "New Parties, Populism, and Parliamentary Polarization: Evidence from Plenary Debates in the German Bundestag." *The Palgrave Handbook of Populism*. Ed. by MICHAEL OSWALD. Basingstoke: Palgrave Macmillan Cham, 611–627.

LOCHMANN, ALEXIA (2020). "Culture and Emigration: Evidence from the Hitler-Mussolini Migration Option Agreement." *Working Paper*.

LUTT, ALEXANDER (2016). "Südtiroler Option – Archivbestände und ihre gestörte Überlieferung." Master Thesis. University of Vienna.

LYALL, JASON (2020). *Divided Armies: Inequality and Battlefield Performance in Modern War*. Princeton and Oxford: Princeton University Press.

LYALL, JASON and YURI M. ZHUKOV (2021). "Fratricidal Coercion in Modern War." *Working Paper*.

MAGNUSSON, MÅNS, RICHARD ÖHRVALL, KATARINA BARRLING, and DAVID MIMNO (2018). "Voices from the far right: a text analysis of Swedish parliamentary debates." *Working Paper*.

MARCHAIS, GAUTHIER, CHRISTIAN M. MUGARUKA, SÁNCHEZ DE LA SIERRA, RAÚL, and DAVID QIHANG WU (2022). "The Pro-Social Determinants of Violent Collective Action: Evidence from Participation in Militias in Eastern Congo." *Working Paper*.

MCKINNEY, WES (2010). "Data Structures for Statistical Computing in Python." *Proceedings of the 9th Python in Science Conference*. Ed. by STÉFAN J. VAN DER WALT and JARROD MILLMAN, 56–61.

MEGUID, BONNIE M. (2005). "Competition Between Unequals: The Role of Mainstream Party Strategy in Niche Party Success." *American Political Science Review*, 99 (3), 347–359.

MILKMAN, KATHERINE L., JOHN BESHEARS, JAMES J. CHOI, DAVID LAIBSON, and BRIGITTE C. MADRIAN (2011). "Using implementation intentions prompts to enhance influenza vaccination rates." *Proceedings of the National Academy of Sciences*, 108 (26), 10415–10420.

MILKMAN, KATHERINE L., MITESH S. PATEL, LINNEA GANDHI, HEATHER N. GRACI, DENA M. GROMET, HUNG HO, JOSEPH S. KAY, TIMOTHY W. LEE, MODUPE AKINOLA, JOHN BESHEARS, et al. (2021). "A megastudy of text-based nudges encouraging patients to get vaccinated at an upcoming doctor's appointment." *Proceedings of the National Academy of Sciences*, 118 (20), 10–12.

MUDDE, CAS (2007). *Populist Radical Right Parties in Europe*. Cambridge: Cambridge University Press.

MUDDE, CAS and CRISTÓBAL ROVIRA KALTWASSER (2018). "Studying Populism in Comparative Perspective: Reflections on the Contemporary and Future Research Agenda." *Comparative Political Studies*, 51 (13), 1667–1693.

MÜLLER, KARSTEN and CARLO SCHWARZ (2020). "From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment." *American Economic Journal: Applied Economics*, forthcoming.

— (2021). "Fanning the flames of hate: Social media and hate crime." *Journal of the European Economic Association*, 19 (4), 2131–2167.

MUMELTER, GERHARD (2017). "Tiefes Südtiroler Unbehagen." *Der Standard*. URL: https : / / www . derstandard . at / story / 2000054779267 / tiefes - suedtiroler - unbehagen (last accessed on 02/04/2023).

NEWMAN, BENJAMIN, JENNIFER L. MEROLLA, SONO SHAH, DANIELLE CASAREZ LEMI, LOREN COLLINGWOOD, and S. KARTHICK RAMAKRISHNAN (2021). "The Trump effect: An experimental investigation of the emboldening effect of racially inflammatory elite communication." *British Journal of Political Science*, 51 (3), 1138–1159.

OBERMAIR, HANNES (2021). *"Großdeutschland ruft!": Südtiroler NS-Optionspropaganda und völkische Sozialisation : Katalog zur Ausstellung = "La Grande Germania chiama!" : la propaganda nazionalsocialista sulle Opzioni in Alto Adige e la socializzazione völkisch*. Tirol: Schloss Tirol.

OLSON, MANCUR (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.

PALOMARES, NICHOLAS A., HOWARD GILES, JORDAN SOLIZ, and CINDY GALLOIS (2016). "Intergroup Accommodation, Social Categories, and Identities." *Communication Accommodation Theory*. Ed. by HOWARD GILES. Cambridge: Cambridge University Press, 123–151.

PEDREGOSA, FABIAN et al. (2011). "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research*, 12 (85), 2825–2830.

PEREZ-TRUGLIA, RICARDO and GUILLERMO CRUCES (2017). "Partisan Interactions: Evidence from a Field Experiment in the United States." *Journal of Political Economy*, 125 (4), 1208–1243.

POTTERS, JAN, MARTIN SEFTON, and LISE VESTERLUND (2007). "Leading-by-example and signaling in voluntary contribution games: an experimental study." *Economic Theory*, 33 (1), 169–182.

QIAN, NANCY and MARCO TABELLINI (2022). "Discrimination and State Capacity: Evidence from WWII U.S. Army Enlistment." *Harvard Business School Working Paper*, No. 21-005.

RAABE, TOBIAS (2020). *A Python tool for managing scientific workflows*. URL: https://github.com/pytask-dev/pytask (last accessed on 03/03/2023).

RAUCHENSTEINER, MANFRIED (2014). *The First World War and the End of the Habsburg Monarchy, 1914-1918*. Wien: Böhlau.

REHUREK, RADIM and PETR SOJKA (2011). "Gensim - python framework for vector space modelling." *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic*, 3 (2).

REINSTEIN, DAVID and GERHARD RIENER (2012). "Reputation and influence in charitable giving: an experiment." *Theory and Decision*, 72 (2), 221–243.

RICHTER, FLORIAN, PHILIPP KOCH, OLIVER FRANKE, JAKOB KRAUS, FABRIZIO KURUC, ANJA THIEM, JUDITH HÖGERL, STELLA HEINE, and KONSTANTIN SCHÖPS (2020). *Open Discourse: Harvard Dataverse, V3*. URL: https://doi.org/10.7910/DVN/FIKIBO (last accessed on 03/03/2023).

ROBERT KOCH INSTITUT (2022). *Archiv mit Daten bis einschließlich 10.09.2021: 7-Tage-Inzidenzen nach Bundesländern und Kreisen (fixierte Werte)*. URL: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Daten/Fallzahlen_Kum_Tab.html (last accessed on 08/16/2022).

ROHNER, DOMINIC and EKATERINA ZHURAVSKAYA, eds. (2023). *Nation Building: Big Lessons from Successes and Failures*. London: CEPR Press.

ROMARRI, ALESSIO (2022). "Strongmen in power are not without consequences: far-right mayors and hate crimes." *Working Paper*.

ROODUIJN, MATTHIJS, SARAH L. de LANGE, and WOUTER VAN DER BRUG (2012). "A Populist Zeitgeist? Programmatic Contagion by Populist Parties in Western Europe." *Party Politics*, 20 (4), 563–575.

ROODUIJN, MATTHIJS, STIJN VAN KESSEL, ANDREA PIRRO, SARAH de LANGE, DAPHNE HALIKIOPOULOU, PAUL LEWIS, CAS MUDDE, and PAUL TAGGART (2019). *The PopuList: An Overview of Populist, Far Right, Far Left and Eurosceptic Parties in Europe*. URL: http://www.popu-list.org/ (last accessed on 03/03/2023).

ROSHWALD, AVIEL (2001). *Ethnic Nationalism and the Fall of Empires: Central Europe, the Middle East and Russia, 1914-23*. London: Routledge.

ROTH, CHRISTOPHER, SONJA SETTELE, and JOHANNES WOHLFART (2022). "Risk Exposure and Acquisition of Macroeconomic Information." *American Economic Review: Insights*, 4 (1), 34–53.

ROZENAS, ARTURAS, ROYA TALIBOVA, and YURI M. ZHUKOV (2022). "Fighting for Tyranny: State Repression and Combat Motivation." *Working Paper*.

SALEM, ARIANE and AWA AMBRA SECK (2023). "En Route: The French Colonial Army, Emigration, and Development in Morocco." *Working Paper*.

SARVIMÄKI, MATTI, ROOPE UUSITALO, and MARKUS JÄNTTI (2022). "Habit Formation and the Misallocation of Labor: Evidence from Forced Migrations." *Journal of the European Economic Association*, 20 (6), 2497–2539.

SASAKI, SHUSAKU, TOMOYA SAITO, and FUMIO OHTAKE (2022). "Nudges for COVID-19 voluntary vaccination: How to explain peer information?" *Social Science & Medicine*, 292, 114561.

SCHEER, TAMARA (2020). "Language Diversity and Loyalty in the Habsburg Army, 1868-1918." Habilitation Thesis. University of Vienna.

SCHILTER, CLAUDIO (2018). "Hate Crime after the Brexit Vote: Heterogeneity Analysis based on a Universal Treatment." *Working Paper*.

SCHMIDT, MANFRED G. (2021). *Das politische System Deutschlands*. 4th ed. München: C.H. Beck.

SCHULZE, MAX-STEPHAN and NIKOLAUS WOLF (2008). "On the origins of border effects: insights from the Habsburg Empire." *Journal of Economic Geography*, 9 (1), 117–136.

— (2012). "Economic nationalism and economic integration: the Austro-Hungarian Empire in the late nineteenth century." *The Economic History Review*, 65 (2), 652–673.

SCHWAABE, CHRISTIAN (2018). *Politische Theorie: von Platon bis zur Postmoderne*. Paderborn: Wilhelm Fink.

SHAYO, MOSES (2009). "A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution." *American Political Science Review*, 103 (2), 147–174.

— (2020). "Social Identity and Economic Policy." *Annual Review of Economics*, 12 (1), 355–389.

SPOON, JAE-JAE and HEIKE KLÜVER (2020). "Responding to Far Right Challengers: Does Accommodation Pay Off?" *Journal of European Public Policy*, 27 (2), 273–291.

STATISTISCHE ÄMTER DES BUNDES UND DER LÄNDER (2022). *Bundestagswahlen: Wahlberechtigte, Wahlbeteiligung, gültige Zweitstimmen nach ausgewählten Parteien - Wahltag - regionale Tiefe: Gemeinden (bis 2017) (code: 14111-01-03-5)*. URL: https://www.regionalstatistik.de/genesis/online (last accessed on 08/16/2022).

STEELE, LIZA G., AMIE BOSTIC, SCOTT M. LYNCH, and LAMIS ABDELAATY (2022). "Measuring Ethnic Diversity." *Annual Review of Sociology*, 48 (1), 43–63.

STEININGER, ROLF (1997a). "Juni 1939: Gehen oder Bleiben? Die Option in Südtirol." *Österreich im 20. Jahrhundert.* Ed. by ROLF STEININGER and MICHAEL GEHLER. Wien: Böhlau, 217–257.

— (1997b). *Südtirol im 20. Jahrhundert: Vom Leben und Überleben einer Minderheit*. Innsbruck, Wien: Studienverlag.

— (2003). *South Tyrol: A Minority Conflict of the Twentieth Century*. New Brunswick: Transaction Publishers.

— (2006). *Autonomie oder Selbstbestimmung? Die Südtirolfrage 1945/46 und das Gruber-De Gasperi-Abkommen*. Innsbruck: Studienverlag.

STERGAR, ROK and TAMARA SCHEER (2018). "Ethnic boxes: the unintended consequences of Habsburg bureaucratic classification." *Nationalities Papers*, 46 (4), 575–591.

SUCHE-POSTLEITZAHL.ORG (2022). *Liste aller Orte mit aktueller PLZ-Zuordnung*. URL: https://www.suche-postleitzahl.org/downloads (last accessed on 08/16/2022).

TAJFEL, HENRI and JOHN C. TURNER (1979). "An Integrative Theory of Intergroup Conflict." *The Social Psychology of Intergroup Relations*. Ed. by STEPHEN WORCHEL and WILLIAM G. AUSTIN. Monterey, CA: Brooks/Cole, 94–109.

TALIBOVA, ROYA (2021). "Repression, Military Service and Insurrection." *Working Paper*.

TAYLOR, ALAN J. P. (1990). *The Habsburg Monarchy 1809–1918. A History of the Austrian Empire and Austria-Hungary*. London: Penguin Books.

THOMPSON, MARK (2008). *The White War: Life and Death on the Italian front, 1915-1919*. London: Faber and Faber.

TILLY, CHARLES (1975). "Reflections on the History of European State-Making." *The Formation of National States in Western Europe*. Ed. by CHARLES TILLY. Princeton, NJ: Princeton University Press, 3–83.

— (1978). *From Mobilization to Revolution*. New York: Random House.

— (1993). *Coercion, Capital and European States: AD 990 - 1992*. Cambridge, MA: Wiley-Blackwell.

— (1994). "States and Nationalism in Europe 1492-1992." *Theory and Society*, 23 (1), 131–146.

TIROLER LANDESMUSEEN (2014). *Ehrenbücher des Landes Tirol - 1. Weltkrieg*. URL: https://www.tiroler-landesmuseen.at/forschung/tiroler-ehrenbuch-digital/ (last accessed on 06/22/2022).

ÜBEREGGER, OSWALD (2003). "Auf der Flucht vor dem Krieg. Trentiner und Tiroler Deserteure im Ersten Weltkrieg." *Militärgeschichtliche Zeitschrift*, 62 (2), 355–394.

— (2018). "Minderheiten-Soldaten. Staat, Militär und Minderheiten im Ersten Weltkrieg - Eine Einführung." *Minderheiten-Soldaten*. Ed. by OSWALD ÜBEREGGER. Paderborn: Ferdinand Schöningh, 9–23.

VALENTIM, VICENTE and TOBIAS WIDMANN (2021). "Does Radical-Right Success Make the Political Debate More Negative? Evidence from Emotional Rhetoric in German State Parliaments." *Political Behavior*, forthcoming.

VAN SPANJE, JOOST (2010). "Contagious Parties." *Party Politics*, 16 (5), 563–586.

VEREIN FÜR COMPUTERGENEALOGIE (2021). *Verlustlisten Österreich-Ungarns 1914-1919*. URL: https://wiki.genealogy.net/Verlustlisten_%C3%96sterreich-Ungarns_1914-1919 (last accessed on 06/07/2021).

VLACHOS, STEPHANOS (2022). "On war and political radicalization: Evidence from forced conscription into the Wehrmacht." *European Economic Review*, 144, 104086.

VON HIRSCHHAUSEN, ULRIKE (2009). "From imperial inclusion to national exclusion: citizenship in the Habsburg monarchy and in Austria 1867–1923." *European Review of History - Revue européenne d'histoire*, 16 (4), 551–573.

WALDEN, JACOB and YURI M. ZHUKOV (2020). "Historical Legacies of Political Violence." *Oxford Research Encyclopedia of Politics*. Ed. by JACOB WALDEN and YURI M. ZHUKOV. Oxford University Press.

WANDRUSZKA, ADAM and PETER URBANITSCH (1980). "'Notwendiger Völkerverein' oder 'Völkerkerker'?" *Die Völker des Reichs*. Ed. by ADAM WANDRUSZKA and PETER URBANITSCH. Die Habsburgermonarchie, Vol. III. Wien: Verlag der Österreichischen Akademie der Wissenschaften, XIII–XVIII.

— (1987). *Die Bewaffnete Macht*. Die Habsburgermonarchie, Vol. V. Wien: Verlag der Österreichischen Akademie der Wissenschaften.

WARTENA, CHRISTIAN (2019). "A Probabilistic Morphology Model for German Lemmatization." *Proceedings of the 15th Conference on Natural Language Processing (KONVENS 2019)*. Proceedings of the 15th Conference on Natural Language Processing (KONVENS 2019), 40–49.

WEDEKIND, MICHAEL (2003). *Nationalsozialistische Besatzungs- und Annexionspolitik in Norditalien 1943 bis 1945*. München: Oldenbourg.

WIDMER, PHILINE, SERGIO GALLETTA, and ELLIOTT ASH (2022). "Media Slant is Contagious." *Working Paper*.

WILKERSON, JOHN and ANDREU CASAS (2017). "Large-Scale Computerized Text Analysis in Political Science: Opportunities and Challenges." *Annual Review of Political Science*, 20 (1), 529–544.

WIMMER, ANDREAS (2013). *Waves of War: Nationalism, State Formation, and Ethnic Exclusion in the Modern World*. Cambridge: Cambridge University Press.

WINKLER, WILHELM (1919). *Die Totenverluste der öst.-ung. Monarchie nach Nationalitäten*. Wien: Seidel.

ZWEIG, STEFAN (1948). *Die Welt von Gestern. Erinnerungen eines Europäers.* Wien: Bermann-Fischer.

Eidesstattliche Versicherung

Ich versichere hiermit eidesstattlich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie mir gegebene Anregungen sind als solche kenntlich gemacht. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Sofern ein Teil der Arbeit aus bereits veröffentlichten Papers besteht, habe ich dies ausdrücklich angegeben.

München, 17. März 2023

EMILIO ESGUERRA