transeuroworks

Working Paper 4/2023

Artificial Intelligence and Political Behavior – Experience with ChatGPT

Siegfried Manschein

transeuroworks

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. N. 101061198

Abstract: This paper seeks to explore the relationship between exposure to artificial intelligence (AI) technology and attitudes towards populism and demand for redistribution. Theoretically, it is hypothesized that exposure to AI may result in individuals experiencing anxiety about their future status and, as a result, increasingly favor populist attitudes and are more likely to reject redistribution. To test this, a pre-registered online survey experiment was conducted, involving 752 participants who were split into three groups: those who had a free interaction with ChatGPT, those who watched an informational video about GPT and generative AI, and a control group. I find evidence regarding populist attitudes which partly increase after being exposed to AI, but no effect on redistribution support. Furthermore, heterogeneity analysis provides evidence concerning different characteristics such as personality, age, and occupational risk.

Keywords: Artificial Intelligence, Technological Change, Political Behavior, Populism, Redistribution

Introduction

In recent months, the introduction of several Large Language Models (LLMs) has sparked a vigorous debate over the implications of Artificial Intelligence (AI) on labor displacement, workplace productivity, and ethical concerns[1]. This surge of investments and new AI-powered products (see Figure 1[2]) has created a complex landscape for companies and governments to navigate. The increasing presence of AI in the workplace has raised questions of whose interests are best served, and has challenged traditional conceptions of the human-machine divide. With the potential for both great rewards and damaging consequences, it is important to take into account the social, ethical, and economic ramifications of these developments when evaluating the potential impact of AI on the workplace. While companies are announcing new developments, they are also reporting layoffs[3], highlighting the possible ramifications of this technology. At the same time, governments are struggling to implement effective regulations that balance the potential benefits and harms of AI (Acemoglu 2021).

Generative AI has the potential to transform the way we interact with the world, reshaping entire industries and sectors beyond what previous technological advancements have achieved. As AI-induced labor markets are becoming commonplace, people are still spending a significant portion of their live at work. Previous iterations of technological change (e.g. robotization or digitalization) are associated with political behavior. This development usually evolves in the background and over decades rather than days and is one of the caveats to identify effects regarding political behavior in the short-term. However, the introduction of ChatGPT in late fall November 2022 marked a significant technological advancement in the field of AI (see Figure 2[4]). This sudden shift in public focus towards AI highlights the impact of technological change on society and how it can bring about new developments and opportunities. An open question remains if some of the long-term effects scholars observed in the past are now immediately recognized or not. As such, it is crucial to examine how the potential benefits and risks of these technologies are perceived.

I provide experimental evidence of the effect of AI on the demand for redistribution and the prevalence of populist attitudes. To do this, a novel online survey experiment was designed: a GPT-

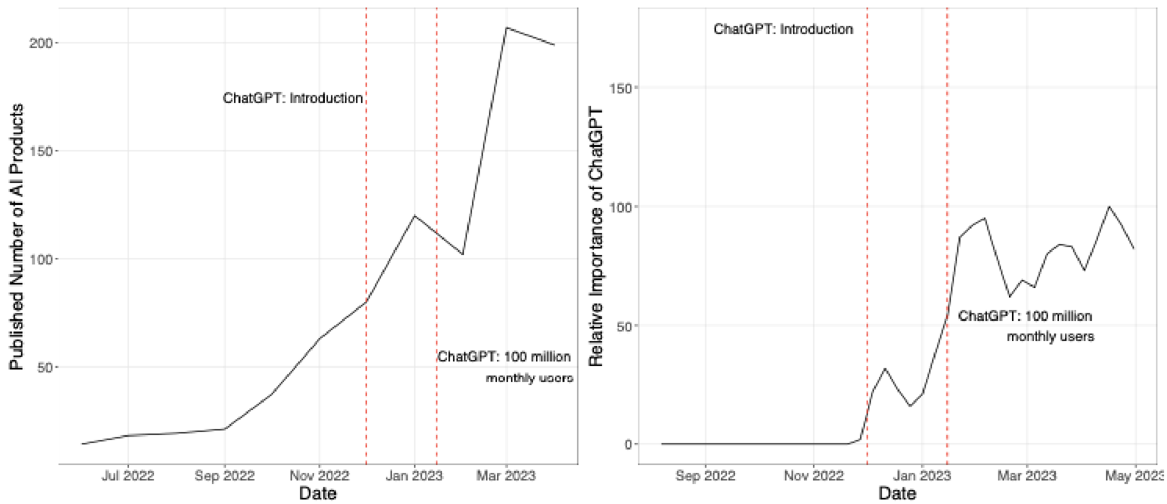
1 Substitution and Productivity

2 source: https://trends.google.es/trends/

3 Microsoft Layoffs, Microsoft & OpenAI Investment, Google AI Investment, IBM Layoffs

4 source: https://theresanaiforthat.com/

3.5 powered chat was deployed to compare the responses of individuals who freely interacted with it, with those of a group that watched an informational video about GPT, and with a control group. The treatments were complemented by a series of questions that served as possible mediators and outcomes, and a causal mediation design was implemented with a sensitivity analysis (Imai, Keele, and Tingley 2010; Imai, Tingley, and Yamamoto 2013). I find support that the exposure to AI affects populist attitudes positively. Respondents in treatment groups agree on average approximately 10-12%-points (Cohen's d = 0.25) more to the statement that the general will of the people should prevail compared to the control group. However, mediation analysis suggests economic interest as a likely channel counter to expected status. Furthermore, there is no evidence that exposure to AI is related to support for redistribution.



(a) Number of Published AI Tools (b) Google Trends for ChatGPT (Germany)

Figure 1: The Rapid Development of AI and ChatGPT

In broad terms, this paper contributes to the growing literature of the de-industrialization of Western European economies and a shift to knowledge and/or information regimes (Esping-Andersen 1990; Pierson 1996). In particular, how labor markets evolve as consequences of underlying technological and demographic changes. Primarily, a divide of labor markets into losers and winners, occupational polarization (D. H. Autor, Levy, and Murnane 2003) or the dualization of labor markets into insiders and outsiders (Häusermann and Schwander 2012; Oesch and Rennwald 2018; Schwander and Häusermann 2013). Tied to these developments is a shift of welfare states towards social investments policies (Hemerijck 2018) as well as the decline of social democratic parties (Berman and Snegovaya 2019; Gingrich and Häusermann 2015) over the last 20-30 years and a shift to populist right-wing parties (Benedetto, Hix, and Mastrorocco 2020).

Therefore, I provide evidence of how AI as a potentially far reaching step of the ongoing digitalization poses new questions about possible winners and losers and their political reactions.

In narrow terms, this paper contributes to the literature of technology induced substitution risks and its relationship to political behavior with a novel causal identification through a survey experiment. Moreover, it is one of the first studies that directly includes AI as treatment, in particular in connection to political attitudes. Noy and Zhang (2023) and Brynjolfsson, Li, and Raymond (2023) are one of the first ones to observe the consequences of AI in terms of workplace productivity. Regarding the theoretical links of expected status, Im et al. (2023) comes closest to this approach by exploring the mechanism through observational data for Finland. Additionally, while technological change usually develops slowly over decades this paper tries to explore if AI has immediate effects not only on susceptible individuals but also on bigger segments of the population.

Furthermore, I try to provide demand side evidence of how technology itself shapes attitudes of individuals, while previous literature focuses overwhelmingly on supply side factors of how political parties appeal to voters (Im et al. 2019; Kurer 2020; Kurer and Palier 2019). This line of research usually links occupational risk and already experienced status decline to voting behavior. While mainstream parties lost their credibility, mostly populist right-wing alternatives offer a nostalgic perspective to the left-behind. Recent electoral success in the Brexit-vote for the UK but also for France's Le Pen in the presidential elections or the gilet jaune protests as well as Trumps win in the US in 2016 (Frey, Berger, and Chen 2018) are only some of the many examples.

Finally, the literature on technological change and political behavior is still theoretically divided on how these concepts should be related. Some researchers argue that the economic interests of the losers of automation should lead to increasing demands for redistribution and voting for the left (Thewissen and Rueda 2019). Others argue that relative status decline has exactly the opposite effect, leading to populist far-right voting and authoritarianism (Im et al. 2023; Kurer 2020; Kurer and Palier 2019). Additionally, it is unclear how and if individuals trade off these two channels. With this paper, I try to help disentangle these mechanisms by providing the direct effect of individuals exposed to AI.

I will continue with an overview of the state of the art regarding the political economy of technological change, as well as propose my own theoretical expectations and hypotheses. Following that, I will provide an overview of the experimental design and an in-depth look into the operationalization of the main variables. After that, I will present the main analysis and results, followed by a discussion of my findings. Lastly, I will provide a roadmap for future research.

The Political Economy of Technological Change

The rise of industrialization in the 18th and 19th centuries brought about concerns over the replacement of human labor by machines, an issue that has persisted over time (Mokyr 2018; Mokyr, Vickers, and Ziebarth 2015). The introduction of spinning jennies in the textile industry in the UK during the late 18th century elicited worker dissatisfaction (Schneider 2023), which ultimately culminated in the Luddite movements that actively and violently opposed the negative impact of machines. The Luddites believed that machines were inherently bad and that their use would lead to the degradation of the quality of work and the loss of skilled labor.

Despite concerns over technological unemployment, recent and historical evidence suggests that the long-term net gains of technological employment have outweighed unemployment (D. Autor and Salomons 2017, 2018; D. Autor, Salomons, and Seegmiller 2021; Mokyr, Vickers, and Ziebarth 2015). The impact of new technologies on the labor market has been complex, with some traditional occupations and tasks being replaced by automation, while at the same time, new job opportunities have emerged, and costs have been lowered, leading to increased demand for labor (Acemoglu and Restrepo 2018, 2020). For example, the rise of online shopping has led to the decline of brick-and-mortar retail stores, but it has also created new job opportunities in fields such as e-commerce management, digital marketing, and logistics.

Yet, the impact of new technologies on the labor market has been significant, with short-term adaptation pressures leading to the distortion of entire industries and the creation of winners and losers. The recent decades were marked by the introduction of personal computers, industrial robotization, and digitalization. As D. H. Autor, Levy, and Murnane (2002) and D. H. Autor, Levy, and Murnane (2003) noticed, the adoption of software and automation has led to the replacement of routine tasks, which has changed the skill structure demanded by the labor market, leading to skill- and routine-biased upgrading. Consequently, high-skilled workers in non-routine jobs have benefited from these changes, while lower-skilled workers in routine jobs have ended up with lower-paid service jobs. This has led to occupational polarization, with a diminishing middle class and a concentration of employment in high- and low-paid jobs (D. H. Autor 2019; Kurer 2020).

The introduction of artificial intelligence as a next step poses new questions to this literature as it is still unclear who the possible winners and losers will be. Webb (2019) doubts that education will be correlated with susceptibility to automation and if, then positively. Similarly, Eloundou et al. (2023) find that GPTs abilities are not necessarily associated to routine-intensity but to programming and writing tasks. Brynjolfsson, Li, and Raymond (2023) and Noy and Zhang (2023) on the other hand provide first evidence that generative AI like ChatGPT mainly helps low skilled workers in terms of productivity to close the skill gap between them and high skilled workers.

Technological Change and Material Interest

The intersection of economics and political science literature lies in the concept of occupational risk, which refers to the susceptibility of certain occupations to substitution by technology. While both fields acknowledge the importance of occupational risk, political science scholars tend to treat it as the main explanatory variable for political behavior, as individuals who are susceptible to substitution or replacement by technology may have different attitudes and preferences. For instance, those who are at risk of losing their jobs due to automation may be more likely to support policies that protect their future economic losses.

Indeed, Thewissen and Rueda (2019) propose a mechanism whereby individuals who are vulnerable to automation are more likely to prefer insurance policies. In particular, this mechanism suggests that those who are at risk of automation may be more likely to prefer policies as a form of protection in the event of future economic loss. This is because they are aware that there is an increased probability of unemployment in their occupation, which could be caused by automation. This means that people who are at higher risk of losing their jobs to machines and algorithms are more likely to demand redistribution and insurance policies (Thewissen and Rueda 2019)[5]. Moreover, the demand for these policies is likely to increase with income, as insurance against future losses is considered a normal good (Moene and Wallerstein 2001).

Evidence for the proposed mechanism linking vulnerability to automation and demand for insurance policies is mixed, with some studies providing support for the mechanism, while others do not. Thewissen and Rueda (2019) provide evidence in favor of the mechanism in an observational study using ESS data, indicating that individuals who are more vulnerable to automation demand more social insurance policies. However, Gallego et al. (2022) show no relationship between vulnerability to automation and demand for social insurance in an experimental study in Spain, suggesting that the relationship may be context-dependent and influenced by factors such as political institutions. Conversely, Gallego, Kurer, and Schöll (2021) find that winners of digitalization support incumbent candidates and the conservative party in the UK, indicating that the winners of technological change may stabilize the political system. Additionally, Sacchi, Guarascio, and Vannutelli (2020) find increasing demand for a minimum wage in Italy by more susceptible individuals, however Busemeyer and Sahm (2022) find no connection between vulnerability to automation and support for social investment policies.

5 The authors specify three additional variables which build the function for redistribution: likelihood of re-gaining employment, degree of risk aversion, and presence of some policy that redistributes resources.

This mixed evidence highlights the complexity of the relationship between automation and social policy preferences. As AI gathered massive salience in the media over the last months it is crucial to understand if the average population has also become aware of its potential impact on the labor market. If this is the case, it may lead to changes in social policy preferences which demand government intervention to mitigate potential negative effects. Furthermore, it is important to note that using conventional measurements like routine task intensity (RTI) as a way to operationalize vulnerability to automation may not be as useful with the introduction of AI (Webb 2019). This is because AI has the potential to automate both routine and non-routine tasks, which makes it difficult to accurately measure the extent of vulnerability using RTI. Therefore, new measures and approaches may be necessary to effectively assess the impact of AI on the labor market and inform policy decisions.

Technological Change and Social Status

Contrary to the above mentioned literature, the socio-psychological approach links occupational risk and political behavior through a different argument. The core concept is social status as an individuals perceived relative position in society's hierarchy (Rosenberg 1953; Jackman and Jackman 1973). Occupation, income, and education are some of the main identifying characteristics shaping one's social status. While these factors are usually relatively stable over time and interrelated, demographic and technological developments can affect the perception of them. Past experiences, both positive and negative, can play a significant role in how individual's perceive their current status. For example, the failure of the European Union and member states to react to immigration in the aftermath of 2015 led to a relative perceived status decline in parts of the existing population.

Importantly, as status is relative and not always directly observed, misattribution to other factors could bias perception of individuals. For example, Wu (2022) explores in a survey experiment in the U.S. how individuals misattribute technological substitution. Confronting individuals with vignettes about technological dismissal in a company, subjects react with demanding lower numbers of immigrants and higher tariffs on trade. Former, are mainly supported by Republicans, while latter are favored by democrats. In similar manner, Wu (2021) shows with electoral data from the U.S. that workers susceptible to automation uniformly oppose free trade agreements and favor immigration restriction.

An important element of status is the comparison to others. Individuals engage in social comparison to comprehend their position in their environment and perceive societal arrangements (Festinger 1954; Tajfel and Turner 1979). This benchmarking helps to assess someones own abilities, preferences, and characteristics in relation to others as well as it categorizes them and oneself into social identities (Turner and Onorato 1999). One problem with previous iterations of technology was its relative low salience and visibility. Consequently, there was no reason to compare oneself with computers. However, the rise of AI possibly overturned this paradigm not only regarding specific risk groups but in general.

Social Status and Populist Attitudes

Contrary to the economic interest mechanism mentioned above, the theoretical predictions would differ vastly. When social status is threatened or decline actually experienced it can lead to a feeling of losing out and a loss of control (Im et al. 2019; Kurer and Palier 2019). This feeling of losing out and uncertainty is related to concepts like nostalgia, societal pessimism, and authoritarian values (Ballard-Rosa et al. 2021; Ballard-Rosa, Jensen, and Scheve 2022; Gest, Reny, and Mayer 2018; Steenvoorden and Harteveld 2018). Nostalgia as the feeling of missing out or that things used to be better in the past, can be particularly strong when the current and future society fails to meet one's own expectations. In addition, some people may view the current and future state of society as a regression from past progress, leading to feelings of disappointment and anxiety. These feelings can further contribute to a sense of losing out and increase uncertainty. This emotions may be connected to an increase in authoritarian values, such as an emphasis on conformity and obedience (Ballard-Rosa et al. 2021; Ballard-Rosa, Jensen, and Scheve 2022).

Frey, Berger, and Chen (2018) provide evidence for how experienced status decline shaped the 2016 US presidential elections by the introduction of robots into the workplace. They find that areas like the old manufacturing centers in the rust belt with a greater exposure to robots had a higher share of votes for Donald Trump. In a similar manner, Anelli, Colantone, and Stanig (2019) show for Western European countries that higher exposure to automation increased support for nationalist and radical-right parties, both at the regional and individual level. Kurer (2020) and Kurer and Staalduinen (2022), with the most comprehensive approach so far, show with individual-level data how occupational change and status discordance[6] lead to support of populist right-wing parties in several Western European countries.

However, these studies mainly take experienced status decline into account. Contrary, Im et al. (2019) and Im et al. (2023) focuses on the prospects of individuals and how they expect their status to be. Im et al. (2019) provides cross-country evidence that individuals threatened by automation as well as barely economically managing are more in favor of populist-right wing parties. Based on this approach, Im et al. (2023) directly measures if individuals expect status decline for Finland in

6 Measured as the distance between the individuals achieved occupation and expected occupational outcome based on the parental background.

an observational study. Similarly, individuals that reported to expect a lower status turn to populist-right wing parties.

As the above mentioned studies mainly focus on voting behavior and only partially observe the links between technological substitution and behavior, my goal is to test if populist attitudes are actually increasing when individuals are exposed to possible threatening new technologies. Following the theoretical arguments provided above I expect the following:

- H1: Individuals exposed to AI are more likely to express support for populist attitudes.

Social Status and Redistribution

The notion of expected status decline can have far-reaching implications for how individuals interact with society. This side of the argument builds on Thal (2020) and how status anxiety can lead to a higher degree of competition for one's own position in society as individuals strive to maintain and even improve their standing. This fear can manifest in various ways, such as trying to outdo others or simply avoiding failure at all costs. For example, Velandia-Morales, Rodrıguez-Bailón, and Mart´ınez (2022) and Sivanathan and Pettit (2010) have provided empirical evidence which suggests that those who are fearful of losing their social status are more likely to buy luxury goods in order to re-establish self-esteem. Furthermore, Lungu (2022) found that both, low- and high income individuals prefer lower taxes compared to higher taxes in order to sustain their spending power and hold onto their social standing.

In a similar manner Kim et al. (2017) argue that the feeling of undeserved deprivation can lead to dissatisfaction and resentment. The authors show in an experimental setting that respondents are more likely to spend money on what they want instead of what they need experiencing relative deprivation. As relative status decline is psychologically painful individuals could be inclined to a myopic worldview preferring to keep up their status (Sivanathan and Pettit 2010; Velandia-Morales, Rodrıguez-Bailón, and Mart´ınez 2022). As a consequence, demand for redistribution could be lower for individuals fearing to lose out in the future because as they want to at least maintain the current status. In line with this literature I propose the following:

- H2: Individuals exposed to AI have lower demands for redistribution.

Methodology Experimental Design

To test the proposed theory I implemented a simple experimental design with three different groups (see Figure 2). All respondents are explained the purpose of the study without mentioning ChatGPT or artificial intelligence directly. After agreeing to the terms of the study individuals continue to answer a battery of demographic questions about age, gender, residency, as well as personality predispositions according to Rammstedt and John (2007). The personality battery consists of ten different items to measure extraversion, agreeableness, conscientiousness, neuroticism, and openness. Questionnaires are designed for both, mobile and computer users to make the experience as easy as possible.

Randomization & Manipulation

Randomization of survey participants into three distinct groups should allow for a clear causal identification. The first group serves as a pure control; the second group freely interacts with ChatGPT; and the third is exposed to an informational video about GPT. The pure control group follows a straightforward path, moving directly from mediators to outcomes without any extra steps.

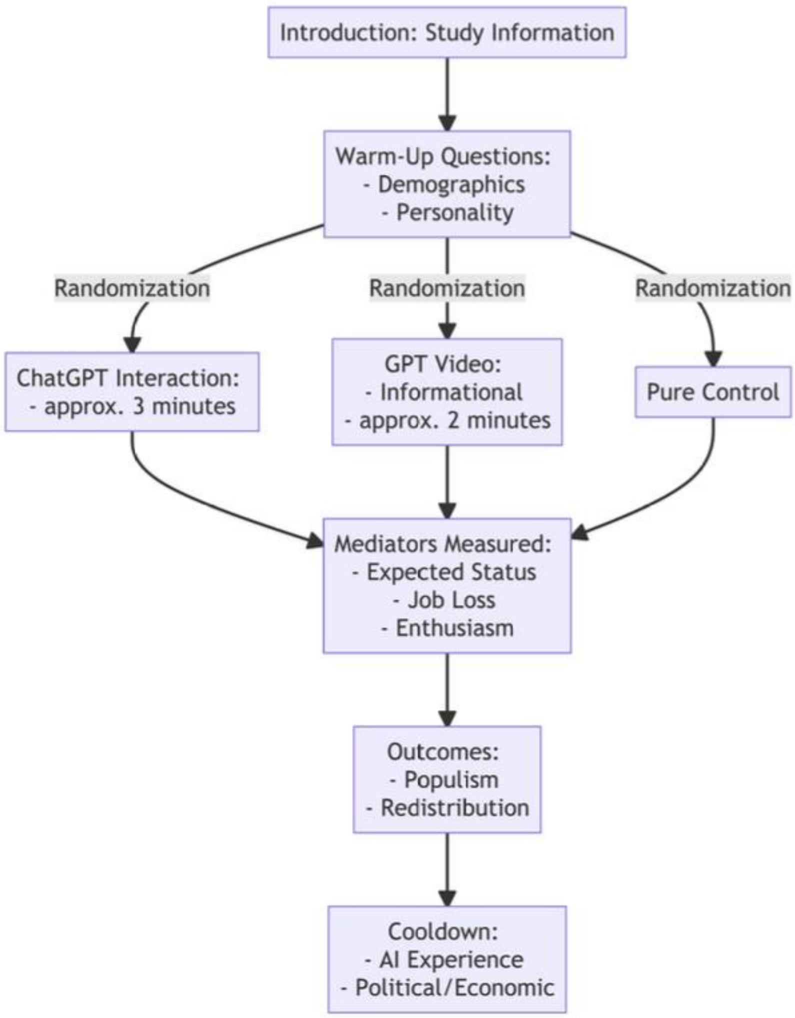


Figure 2: Experimental Flow

In contrast, the first treatment group receives an additional warm-up question before continuing on with an informational screen about ChatGPT. This screen outlines that the respondent will have three minutes to interact with a custom-made chatbot based on GPT-3.5. During this time, respondents are free to interact in any way they want without any restrictions — except for the time limit.

Furthermore, to have more control over treatment conditions a second treatment group is included that watches an informational video about GPT and how several occupational tasks can change through its implementation. The professions shown in the video include teachers, lawyers, journalists, authors, and programmers. This video lasts for two minutes and is presented in a neutral way without any music, in order to avoid invoking any particular emotions. Both treatment groups are asked with an open question after the interaction/video with GPT about their feelings towards GPT and AI in general to reinforce treatment[7].

Mediators & Outcomes

To ensure that there are no question order effects, all respondents were randomly presented with three potential mediators that have been linked to the two main outcomes mentioned in the literature. These mediators were framed in a prospective way, looking towards the next 5-10 years, asking individuals about their expected status, the probability of them losing their job, and how they feel about changes in their working environment. For the main mediator – expected status – I follow Anderson et al. (2012) which includes an image of a ladder involving 10 steps from low to high status with a slight adaptation and respondents are asked the following:

"Imagine that the ladder at the bottom shows where people in Germany stand. On the rung at the bottom are the people who have the least money, the least education and the least respected jobs. On the rung at the top are the people who have the most money, the best education and the most respected jobs. Where would you place yourself on this ladder in the next 5-10 years?"

The other mediators are phrased similarly forward-thinking, aiming to explore how individuals envision their futures in the next 5 to 10 years – the probability of losing their job or their enthusiasm about changes in their working environment. All of the mediators are designed without mentioning AI or similar terms.

7 Especially in early stages of experimental evidence Bullock and Green (2021) argue that varying treatments or treatment intensities can be helpful to establish a relationship between the discussed theoretical concepts.

Following the mediators, participants are guided through a randomized sequence of the main outcomes. They are asked to answer five questions that explore their attitudes towards populism and redistribution. The questions regarding populism are:

The will of the people should be the most important principle in politics. (1)

Traditions should be challenged in order to move society forward. (2)

Germany needs a strong leader who is above the law. (3)

Questions one and three try to capture a reclaiming of control but in distinct ways. The statement about the will of the people can be understood as having a more direct control over politicians aligned with the "general will". Otherwise, the demand for a strong leader can be interpreted gaining control in more authoritarian way. Additionally, the second question should capture norm conformity and the desire for nostalgia which represents the desire for a back to the "good old times". In terms of redistribution demand I ask two questions:

It is the task of the state to reduce the income gap between rich and poor. (4)

Social benefits in Germany cost companies too much in taxes and duties. (5)

Questions four and five should capture the possible implications of AI – the shift from labor to capital – in a general and specific manner. Former question is one of the most commonly applied measurements asking about redistribution (e.g. ESS, WVS, etc.) with a high measurement validity. Latter focuses directly on companies which are the relevant actors regarding capital accumulation as a consequence of AI. Finally, respondents finish the survey with a cool-down that includes questions about their AI experience, ideology, trust (general, business, and science), occupation, and income.

Data

The data has been collected through an online survey experiment by Qualtrics[8]. It has been carried out in late May and beginning of June 2023 and was sampled from the German working age population. I chose Germany as a good starting point for further analysis as it is a country with a relatively wide application of AI compared to other European countries (Eurostat 2023). Furthermore, subjective concern about past iterations of technological change are high (Busemeyer and Sahm 2022) and populist alternatives exist on both sides of the ideological left-right dimension (e.g., Alternative für Deutschland or Die Linke).

8 https://www.qualtrics.com/

This study is specifically designed to examine the ongoing transformation of labor markets and the population directly affected by it. For this reason, the unit of analysis targets working-age individuals between the age of 18 and 65. Retirees and other non-labor market participants have been excluded from the scope of the study. To ensure that sufficient data can be collected for analysis, the target sample size is N = 900. Both, the exclusion of retirees and the sample size are stated in the pre-registration plan including a power analysis.

The actually collected sample includes N = 952 reducing to N = 902 after removing non-respondents. Non-respondents are on average younger, male, less likely to be born in Germany, have lower levels of formal education, and live in more urban areas compared to the rest of the sample. Furthermore, as mentioned above, I exclude retirees and other non-labor market participants as well as low quality respondents leading to a final sample of N = 752. Low quality respondents include the 5%-fastest respondents per experimental group as well as individuals that purposely failed answering some of the qualitative questions[9]. Respondents are randomized into three treatment groups with about a similar size: Control (n = 327), GPT-Interaction (n = 199), and GPT-Video (n = 226). Individuals in the control group take on average 4 minutes and 42 seconds to finish, in the interaction group 10 minutes and 18 seconds, and in the video treatment 8 minutes and 37 seconds.

Table 1 provides an overview of the summary statistics. The first group of variables includes treatment, mediators and outcomes. The average and median respondent has a moderately positive expectation of their own future status with around 6 on a scale between 1 and 10. The other two mediators, the probability of losing one's job and the positive views about workplace changes have means around 3.6 and 4.2 on scales from 1 to 7, indicating a slightly negative view about future job chances but a positive view about workplace changes. The main outcome variables have mean responses between 3 and 3.8 with standard deviations around 1 as hypothesized in the pre-registration plan and in line with previous studies (Fawzi 2019).

The second batteries of variables provides detailed information about demographic characteristics of respondents. The average respondents age is around 30 years which represents a relatively young sample. This has implications for the generalization of the study results which I will discuss further in later sections of the paper. There are more women in the sample (~ 60%) and the clear majority of individuals is born in Germany (~ 90%). The average respondents lives at least in a small city and finished high school with a degree. In line with the age of the sample is the

9 The analysis has been carried out with several thresholds regarding the speed of finishing the survey. Throughout all of the specifications results remain stable and robust.

previous experience with AI as around 50% of respondents already used GPT or similar tools. Furthermore, individuals position themselves at the center of the political left-right dimension with individuals being within one standard deviation from 3 to 7. In terms of trust, respondent's average is between 2.9-3.4 which indicates moderate levels of trust for this sample.

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
Main Variables							
Treatment	3	0	0.9	0.8	0.0	1.0	2.0
Expected Status	10	0	6.1	1.9	1.0	6.0	10.0
Job Loss	8	0	3.6	1.8	1.0	4.0	7.0
Enthusiasm	8	0	4.2	1.3	1.0	4.0	7.0
Will of the People	6	0	3.8	1.0	1.0	4.0	5.0
Challenge Norms	6	1	3.2	1.1	1.0	3.0	5.0
Strong Leader	5	0	3.0	1.3	1.0	3.0	5.0
General Redist.	6	0	3.7	1.0	1.0	4.0	5.0
Company Redist.	5	0	3.0	1.0	1.0	3.0	5.0
Demographics							
Age	4	0	1.9	1.1	1.0	1.0	4.0
Female	3	0	0.6	0.5	0.0	1.0	2.0
Born in Germany	3	1	0.9	0.3	0.0	1.0	1.0
Urban	2	0	0.5	0.5	0.0	0.0	1.0
Education	4	0	1.9	0.9	1.0	2.0	3.0
AI Experience	2	0	0.5	0.5	0.0	0.0	1.0
Ideology	12	0	5.1	2.0	0.0	5.0	10.0
General Trust	5	0	3.0	0.9	1.0	3.0	5.0
Business Trust	6	0	2.9	0.9	1.0	3.0	5.0
Science Trust	6	0	3.4	1.0	1.0	4.0	5.0
Personality							
Extraversion	10	1	3.1	0.9	1.0	3.0	5.0
Agreeableness	10	1	3.4	0.8	1.0	3.5	5.0
Openness	10	1	3.5	0.9	1.0	3.5	5.0
Neuroticism	10	1	2.9	0.9	1.0	3.0	5.0
Conscientiousness	10	1	3.5	0.8	1.0	3.5	5.0
Economic							
Task Risk	16	0	2.3	0.4	1.0	2.2	3.8
Occupational Risk	262	2	0.2	0.9	−1.8	0.2	1.8
Income	11	9	5.1	2.7	1.0	5.0	10.0
Unemployment Hist.	4	0	0.9	1.1	0.0	1.0	3.0
Union Member	3	1	0.2	0.4	0.0	0.0	1.0

Note: **N = 752**, *Expected Status* (low to high: 1-10), *Job Loss* (not likely to certain: 1-7), *Enthusiasm* (not at all to very: 1-7), *Will of the people/Challange Norms'/Strong Leader/General/Specific' Redistribution* (1:5, 'reverse coded), *Age* (18-29, 30-39, 40-49, 50-65), *Female* (0: male, 1: female), *Born in Germany* (0: no, 1: yes), *Urban* (0: rural, 1: at least a small city), *Education* (1: no degree, 2: degree, 3: higher degree), *AI Experience* (0: no, 1: yes), *Ideology* (very left to very right: 0:10), *General/Business/Science Trust* (none at all to high: 1-5), *Personality Types* (low to high: 1:5, .5 steps), *Task Risk* (low to high: 1-3.8, continuous), *Occupational Risk* (-1.8 to 1.8, continuous), *Income* (10 brackets), *Unemployment History* (0: not unemployed, 1: up to 3 months, 2: between 3 and 12 months, 3: more than 12 months), *Union Member* (0: no, 1: yes)

Table 1: Summary Statistics

The third bunch of variables shows details about the big-five personality traits: 1) extraversion, 2) agreeableness, 3) openness, 4) neuroticism, and 5) conscientiousness. The average respondents scores between 2.9 and 3.5 on a scale from 1-5 (.5 steps between) showing moderate levels of the respective personality traits.

Finally, the fourth group of variables includes economic and labor market characteristics. The task-related measure of labor market risk shows a relatively high task risk with an average of 2.3 but a relatively small standard deviation of 0.4 placing most individuals between 1.9 and 2.7. The average household income of respondents is between €2400 and €2800 which is below the German average and around 9% of respondents did not answer. The unemployment history of individuals shows that 50% of respondents have experienced unemployment of less than 3 months and only around 20% are labor union members.

Table 2 illustrates the balance test to check if randomization of individuals into treatment worked. Table 2 presents means and p-values for the between-group comparisons, which indicate whether there are any statistically significant differences between the control group and the two treatment groups on each variable.

Overall, Table 2 suggests that the groups are well balanced on most variables, with small differences observed between the groups for the age variable. For example, there is a difference in mean age between the control group and the GPT-video group ($\Delta_{C-Video}$ = 0.3, p < 0.05). Furthermore, there is a difference in labor union participation for both treatment groups (Δ_{C-GPT} = 0.08, $\Delta_{C-Video}$ = 0.06, p < 0.1). Additionally, there are differences between the treatment groups which are not included in the table in terms of experience with AI and living in a city.

I include unbalanced pre-treatment covariates and experience with AI which is likely unaffected by treatment as control variables in all my specifications in the analysis. In sum, the balance check provides evidence that the groups are similar on most key demographic and psychological variables, which is important for ensuring that any differences observed between the groups after treatment can be attributed to the treatment itself rather than to pre-existing confounding variables between the groups.

Variables	Control Avg.	GPT Avg.	Video Avg.	C - GPT	C - Video
Demographics					
Age	1.82	1.75	2.07	0.44	0.01
Female	0.60	0.57	0.53	0.44	0.11
Born in Germany	0.88	0.88	0.91	0.91	0.32
Urban	0.53	0.46	0.46	0.12	0.12
Education	1.87	1.95	1.88	0.33	0.96
AI Experience	0.48	0.53	0.42	0.29	0.17
Ideology	5.05	5.13	5.07	0.67	0.90
General Trust	3.03	3.03	3.04	0.97	0.87
Business Trust	2.93	2.93	2.98	0.92	0.47
Science Trust	3.43	3.40	3.50	0.74	0.38
Personality					
Extraversion	3.13	3.13	3.12	0.94	0.84
Agreeableness	3.35	3.42	3.39	0.32	0.63
Openness	3.43	3.50	3.47	0.37	0.63
Neuroticism	2.87	2.84	2.86	0.79	0.95
Conscientiousness	3.47	3.54	3.59	0.33	0.11
Economic					
Income	5.13	5.01	5.17	0.64	0.87
Unemployment History	0.89	0.98	0.94	0.34	0.61
Union Member	0.25	0.17	0.19	0.02	0.10

Note: **N = 752**, for detailed variable description see Summary Statistics table

Table 2: Balance Check

Operationalization & Model Specifications

Main Analysis

The treatment variable is straightforwardly coded as a categorical variable indicating the different experimental conditions (0 = control, 1 = interaction, 2 = video). Regarding the outcomes – populist attitudes and demand for redistribution – which are measured either on a 5 point scale or are recoded into binary variables indicating if a respondent agrees or disagrees with the statement. For the continuous operationalization higher values indicate stronger agreement with the statement.

Control variables included as mentioned above are coded the following. Age indicates approximately 10 year intervals from 18-29, 30-39, 40-49, 50-65. The variable is treated as continuous from 1 to 4 where higher values indicate older respondents. Experience with AI is a binary variable indicating if a person has ever used GPT or similar products (= 1). Membership in a union follows the same logic where 1 indicates the membership and 0 otherwise while living in an urban area includes all individuals at least in a small city and otherwise. The baseline specification is a simple OLS model like the following:

$$y_i = \beta treat_i + X_i + \varepsilon_i$$

, where y is the respective outcome for populist attitudes and demand for redistribution, *treat* indicates if the individual is one of the treatment groups or not, X are the included unbalanced covariates, while i indicates the index for individuals, and ε is the residual term.

Mediation Analysis

The mediation analysis will be carried out according to Imai, Tingley, and Yamamoto (2013) which introduce the possibility of causally dependent mechanisms or mediators. Figure 3 shows how this could be the case in the proposed theoretical framework. It is plausible to think that exposure to AI affects the outcomes both through expected status and worries about losing one's job. But at the same time the probability of being out of work can be correlated with expected status which introduces post-treatment bias (Imai, Keele, and Tingley 2010; Imai, Tingley, and Yamamoto 2013). Additionally, the proposed method is robust regarding non-linear models, e.g., using a binary outcome variable as in my case. However, an additional assumption is no-interaction between treatment and mediator. Because of this I will complement the results of the mediation procedure with a sensitivity analysis.

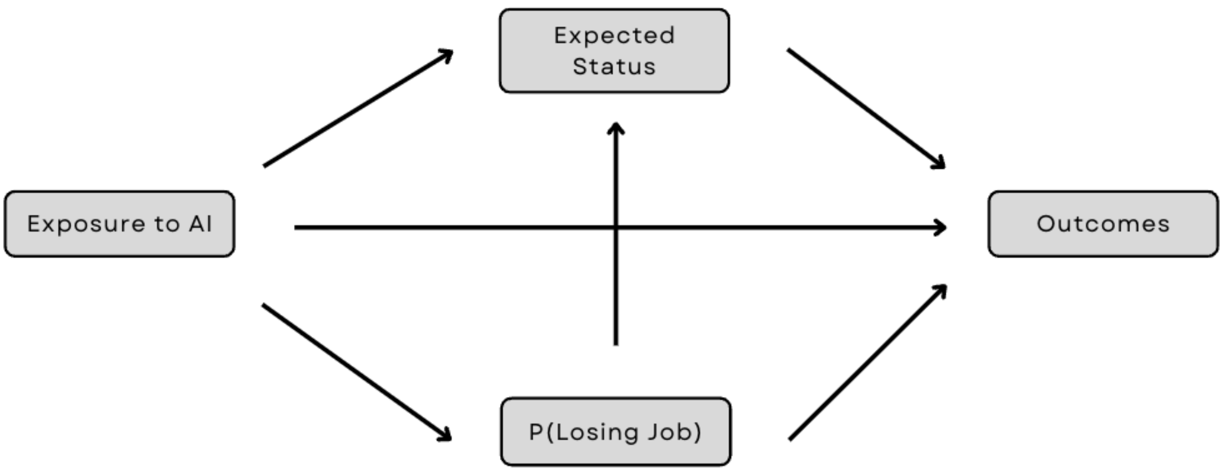


Figure 3: Causal Dependence and Mediation

Mediator variables are measured on different scales. Expected status follows the social ladder approach by Anderson et al. (2012) from low (= 1) to high (=10). Otherwise, the probability of losing the job and the enthusiasm about future changes in the workplace are measured on a 7 point scale. In the former case, higher values (= 7) indicate a higher probability of losing the job while in the latter higher values (= 7) translate to more excitement about the future workplace.

Heterogeneity Analysis

As mentioned in the pre-registration plan I want to explore sub-populations for possible heterogeneous treatment effects. I focus on three different types of variables: 1) personality, 2) age & education, and 3) occupational risk. The 10 personality items are measured on a 5 point scale and include a standard and reverse coded item for each personality type. They are recoded into the specific traits of: 1) extraversion, 2) agreeableness, 3) conscientiousness, 4) neuroticism, and 5) openness. The variables can take values from 1 to 5 with 0.5 steps. Higher values indicate that respondents personality is more likely to be represented by a given trait.

Age will be coded as mentioned above while education is coded in three categories: 1) no degree, 2) degree, and 3) higher degree. Regarding occupational risk different measurements are applied. First, a task-level measurement where respondents where asked how important different tasks in there daily work are (e.g., programming or writing). They are added so that higher values indicate a higher risk of substitution by GPT and similar language models according to Eloundou et al. (2023). Furthermore, a measurement developed by Felten, Raj, and Seamans (2023) which predicts the propensity of how likely different occupations are threatened by generative AI (as GPT):

$$AIOE_k = \frac{\sum_{j=1}^{52} A_{ij} \times L_{jk} \times I_{jk}}{\sum_{j=1}^{52} A_{ij} L_{jk} \times I_{jk}}$$

, where Aij is the ability-level exposure score which is weighted by the prevalence of an ability Ljk and importance Ijk within each occupation. This measurement was recently updated to take newest developments of generative AI into account. Heterogeneity analysis will follow a similarly straightforward approach as before, including an interaction term to estimate the conditional average treatment effect:

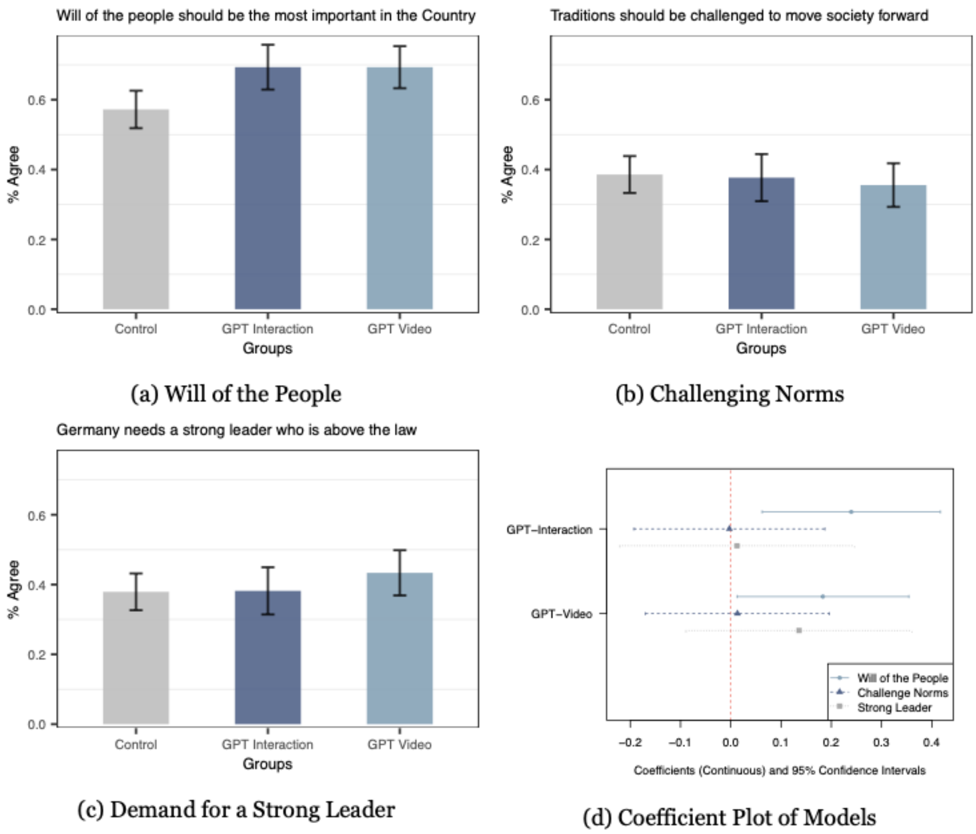
$$y_i = \alpha + \beta treat_i + \gamma Z_i + \delta(treat \times Z)_i + \bar{X} + \varepsilon_i$$

, where in addition to the to above, Z represents the respective moderator (personality, age, education, occupational risk).

Main Analysis

Exposure to AI and Populist Attitudes

Figure 4 presents the results regarding populist attitudes, comparing each treatment to the control group. Figure 4a, Figure 4b, and Figure 4c represent the binary outcomes showing the percentage of respondents agreeing to a specific statement with 95% confidence intervals. Figure 4d illustrates the continuous outcomes as coefficients, also including 95% confidence intervals.



(a) Will of the People

(b) Challenging Norms

(c) Demand for a Strong Leader

(d) Coefficient Plot of Models

Figure 4: Exposure to AI and Populist Attitudes

Figure 4a depicts how individuals respond to the statement that the will of the people should be the most important role in a country. Both treatment groups are significantly different from the control group indicating that after being exposed to GPT 10-12%-points more respondents agree that the general will should prevail ($\beta_{interaction}$ = 0.12, $p_{interaction} < 0.01$, β_{video} = 0.10, $p_{video} < 0.05$). While respondents in the control group agree to 52% to this statement, 62%-64% agree in the treatment group.

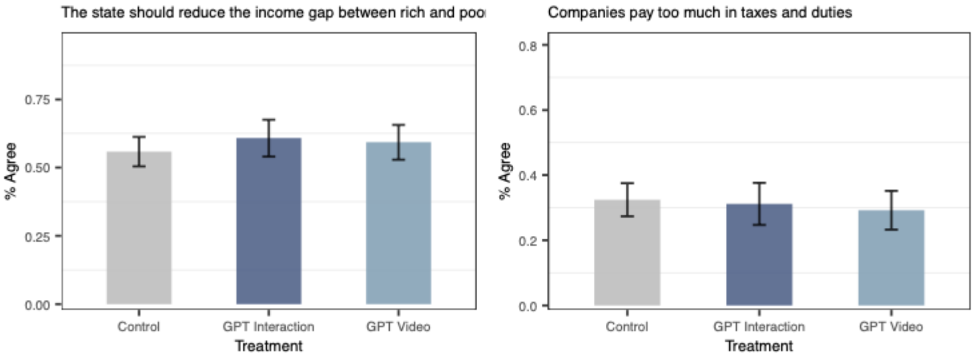
The results indicate an effect size of 1/4 (cohen's d = 0.25) of a standard deviation which is not particularly strong but common for the literature. Figure 4d using the continuous outcomes shows similar results in terms of significance but to a lesser magnitude. This is an indication that respondents that are actually affected by the treatment are the ones that are either already having higher values (in this case a 4 in the continuous outcome) to start with or are in the middle of the distribution (= 3).

Figure 4b and Figure 4c reveal a pattern in line with the hypothesis but without evidence for statistical significance. Both treatments represent lower means of individuals agreeing to challenging traditional norms but difference to the control group is indistinguishable from 0 with only small magnitudes ($\beta_{interaction} = -0.004$, $p_{interaction} > 0.1$, $\beta_{video} = -0.02$, $p_{video} > 0.1$). Similar, in terms of demanding a strong leader means are pointing in the hypothesized direction but not statistically different from the control groups. However, it should be mentioned that this difference is close to significance in the video treatment groups ($\beta video = 0.06$, $pvideo = 0.15$). Figure 4d again indicates that the effect is mainly driven by individuals already having relatively high values of the outcomes as continuous coefficients are smaller and less clear about the direction of the effect.

Regarding the first hypothesis about the effects of exposure to AI on populist attitudes, I conclude support in favor of it, but with further implication to be discussed in the later sections of this paper. Support for the general will clearly increased after individuals are exposed to AI while the other outcomes move in the hypothesized direction. Furthermore, ensuring that this result is not an artifact of multiple hypothesis testing I also applied a Bonferroni test that corrects for this case and the results regarding general will stay robust and significant.

Exposure to AI and Redistribution

Figure 5 presents the results regarding the second hypothesis, that exposure to AI decreases support for redistribution. As above, the plots are presented again as a percentage of respondents agreeing to each of the statements including the 95% confidence intervals. Figure 5a illustrates that both, respondents that interacted with GPT or watched a video about it do not differ in their response about general redistribution compared to the control group ($\beta_{interaction} = 0.05$, $p_{interaction} > 0.1$, $\beta_{video} = 0.03$, $p_{video} > 0.1$). The mean of the control group is around 50% indicating only small effects close to 0. However, it should be noted that the direction of the coefficients is pointing into the opposite direction as hypothesized, thus increasing demand for redistribution.



(a) General Demand for Redistribution (b) Company Taxes too High

Figure 5: Exposure to AI and Demand for Redistribution

Similarly, Figure 5b shows that there is no significant difference between control and treatment about the acceptance that companies pay too much in taxes ($\beta_{interaction} = -0.008$, $p_{interaction} > 0.1$, $\beta_{video} = -0.03$, $p_{video} > 0.1$). Around 30% in each group agree with this statement while both coefficients are negative indicating a similar pattern as above counter the hypothesis. Regarding the second hypothesis about demand for redistribution there seems to be no evidence in favor and if, then in opposing direction than expected. Both differences in means, general redistribution and company specific redistribution, are insignificant across treatment and control.

Exploring the Channels

The first step before testing the proposed mechanisms is to analyze if there is subjective concern about AI. Respondents were asked about their feelings and emotions after being exposed in both of the treatment groups. I categorized the responses into four categories: 1) negative, 2) skeptical, 3) neutral, 4) positive. Negative statements are the ones that clearly state fears and no positive trade-offs about AI, while skeptical answers include the positive and negative impacts. I coded neutral statements as the ones where respondents show indifference (e.g. a response like "no feelings") and positive answers include statements like the ones that speak about future possibilities and how AI increases productivity.

Table 3 shows that approximately 54% of individuals are at least skeptical about artificial intelligence while 36% of the respondents have positive feelings and 10% are more or less indifferent. Both, the relatively high proportion of skeptical respondents and indifferent answers indicate the uncertainty connected to AI. In a more fine grained analysis I focused on the most frequent terms used by respondents. The results reflect the above mentioned emotions as fear and good are the most used words. Other important terms are helpful, interesting, scary, and creepy. Terms about substitution are less commonly used and are in balance with words like support.

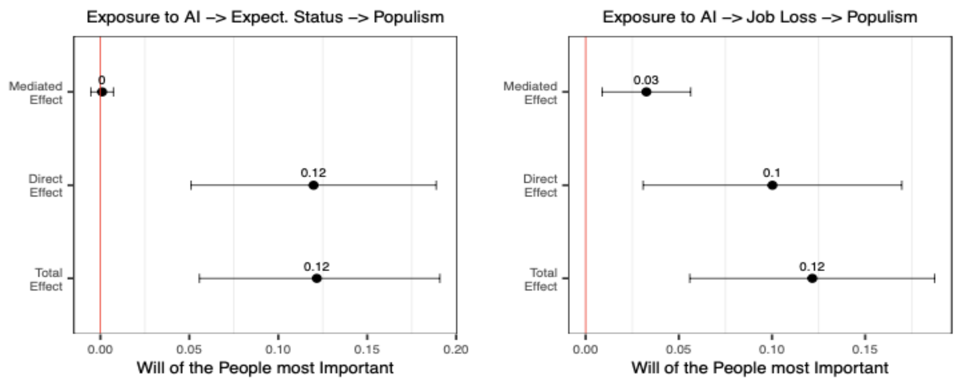
Emotions & Feelings	Proportion of Respondents
Negative	25 %
Skeptical	29 %
Neutral	10 %
Positive	36 %

Note: **N = 425**, only respondents in treatment groups have been asked questions about their feelings and emotions towards AI

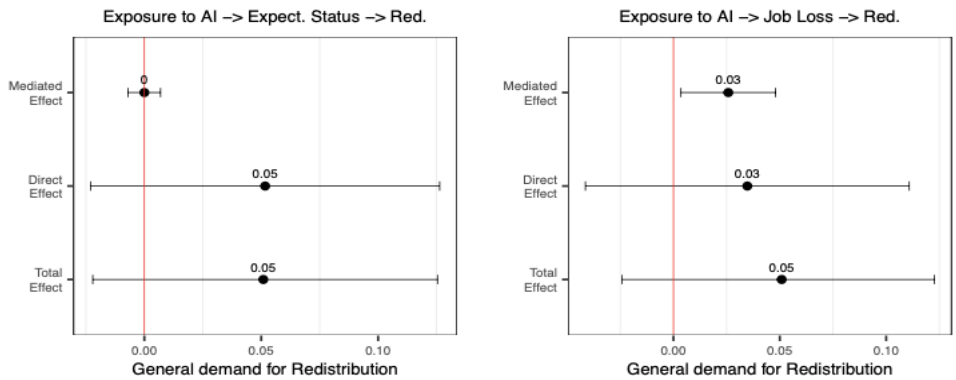
Table 3: Subjective Concern about AI

Mediation Analysis

While these exploratory results indicate the worries and skepticism about AI are indeed important it does not necessarily mean these channel through status fears or job loss. Exploiting the single experimental design I carried out a mediation analysis according to Imai, Tingley, and Yamamoto (2013).



(a) Exposure to AI, Channels & Populism



(b) Exposure to AI, Channels & Redistribution

Figure 6: Mediation Analysis (90%-CI)

Mediation analysis enables to estimate direct and indirect effects. The aim is to estimate to what extent the treatment effect is going through the proposed mediators. In other words, how much of the effect established in the main analysis is going through the expected status, worries of losing the job, or being excited about changes in the workplace. The advantage of Imai, Tingley, and Yamamoto (2013) is the possibility to include dependent or alternative mediators as in my case.

Figure 6 illustrates the results of the analysis including the binary outcomes and two proposed mechanisms from the literature[10]: expected status and probability of job loss, in relation to exposure to AI. The top panel emphasizes the relationship between exposure to AI, expected status/job loss, and attitudes towards the general will. 90%-confidence intervals are obtained through bootstrapping (number of samples = 1000). My theoretical expectations suggest that the individuals exposed to AI should be worried about their future status and thus demand more populist governance of democracy.

However in the left-top panel, both the mediated effect on the treated and the average mediated effect are statistically insignificant and 0. The direct and total effects are similar as in the main analysis indicating an average treatment effect of a 12%-point increase in supporting the statement about the general will of the people. In other words, expected status does not have any explanatory power as a possible mechanism.

Interestingly, the right-top panel provides evidence that worries about losing one's job could be a possible channel connecting exposure to AI and the general will. The mediated effect is significant on a 90% confidence level indicating that around 25% of the total affect (0.03 = 0.25) I channeled through job loss which indicates a significant proportion. On 0.12 the other hand, there are around 75% not explained any of the proposed mechanisms.

The lower panel depicts the same mechanisms as before but focusing on general redistribution as the main outcome. I hypothesized that expected status loss should decrease demand for redistribution. Similarly as for the upper panel, job loss but not expected status seems to be a possible mechanism connecting the two phenomena. Again the mediated effect of status is close to 0 while job loss mediated effect is significant and theoretically explaining most of the total effect (which is insignificant).

As mentioned above I included a sensitivity analysis about how fragile my design is to the interaction of treatment and mediator. As previous literature is scant on values to compare, I use Imai, Tingley, and Yamamoto (2013) paper which provide some evidence of past scholars. The

10 For the sake of space scarcity I am only including mediation analysis for the GPT interaction treatment and two outcome variables (will of the people/general redistribution) as well as I exclude the enthusiasm media- tor. Results follow a similar pattern in the video group/alternative outcomes and the enthusiasm mediator as explained below.

sensitive analysis shows that the results are robust to some extend but that there is relative high uncertainty about it to draw a strong conclusion about this.

Heterogeneity Analysis

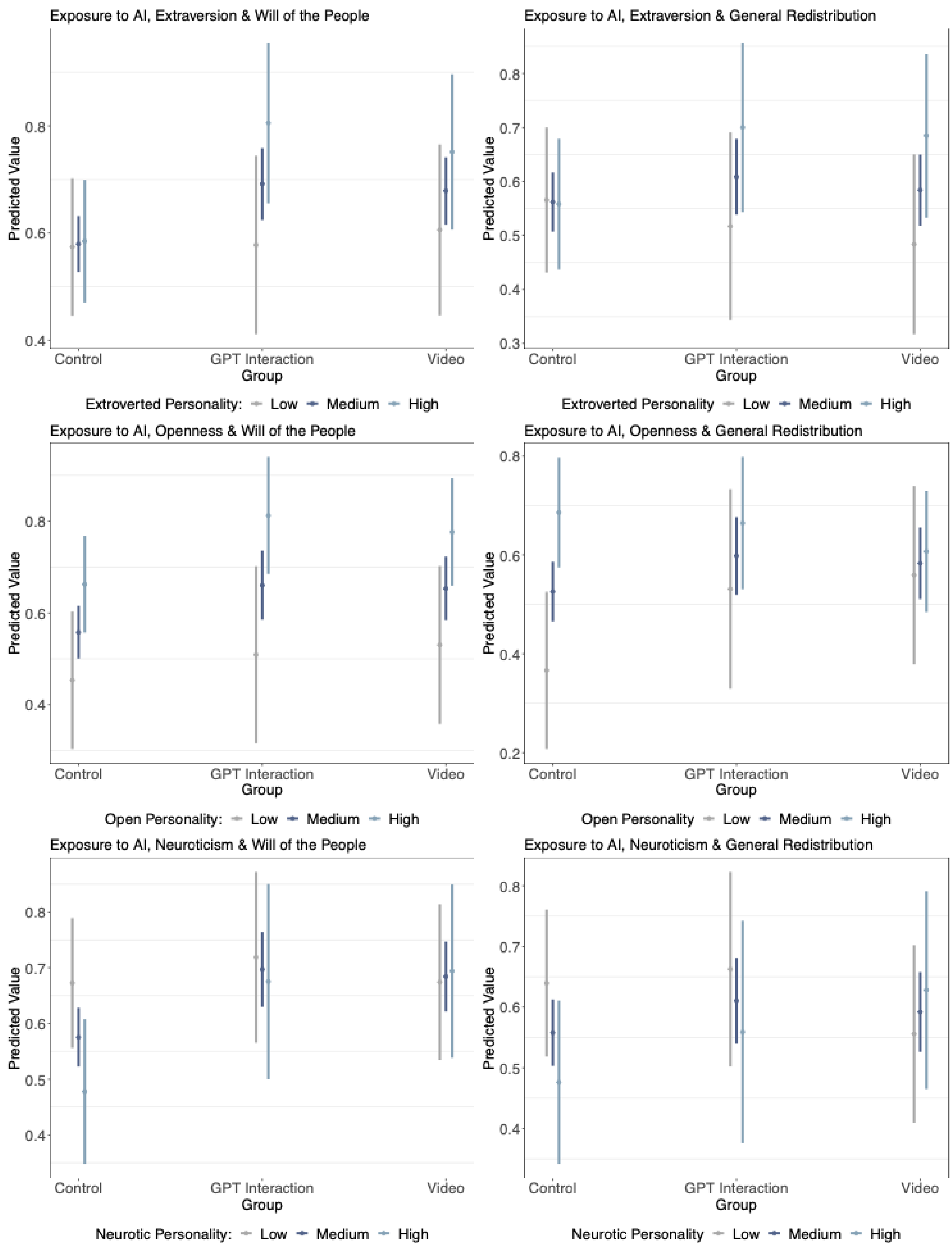
As the main analysis provided some evidence regarding populist attitudes an open question remains about sub-populations. A heterogeneity analysis enables to uncover patterns within these groups for which the hypothesized effects could be relevant. As mentioned in the pre-registration plan I focus on three groups in terms of heterogeneity: 1) personality, 2) age/education, 3) occupational risk. Regarding the first point, whenever individuals experience new technologies the role of personality should be important as it guides individuals of how to perceive them.

Furthermore, age is relevant as worries and enthusiasm about new technologies should be more or less prevalent depending on the temporal position in the labor market. Education on the other hand, provides theoretically the skills for the labor market but could also provide evidence about how informed individuals are. Finally, previous research usually discusses occupational risk direct focusing on the skills used at the workplace (D. H. Autor, Levy, and Murnane 2002; Gallego et al. 2022; Thewissen and Rueda 2019). This part of the analysis should be seen in a more exploratory way without any strong priors. As before I will focus on binary outcomes, in particular the will of the people and general demand for redistribution.

Personality

Personality is treated according to the big-five characteristics that describe individuals traits. It should be relevant for experiencing new technologies as different traits like openness or neuroticism can inherently lead to different perceptions. For example, neurotic individuals should be more likely to have negative feelings about AI and its possible consequences. These individuals are more anxious and are more sensitive to new or unexpected experiences. Figure 7 shows the moderation of three personality traits: 1) extraversion, 2) openness, and 3) neuroticism. The left panel shows moderation regarding the will of the people while the right panel focuses on general redistribution. The predicted values are shown for respondents one standard deviation below the mean, at the mean value and one standard deviation above including the 95%-confidence intervals.

Figure 7: Personality and Exposure to Artificial Intelligence



The first row of Figure 7 shows that, holding introversion constant, that exposure to AI has no effect on these personalities for both outcomes. On the other hand, the more extroverted a person gets the stronger the treatment effect even if the difference between the mean respondent and highly extroverted individuals is not significant. A possible explanation for this finding is that extroverted respondents are more concerned about the impacts on social life of AI as it could mean less interaction with other humans in general and at the workplace. On the other side, introverts are maybe more comfortable communicating with computers or AI in general. As is visible through the strongly varying confidence intervals is that most of the respondents are around the medium level of a specific personality trait which makes it hard to identify significant differences.

Moving to the second row of the Figure 7 a similar pattern as above appears connected to openness, especially regarding the will of the people. Closed-minded respondents are not affected by the treatment while individuals which are open to new experiences are the main drivers of the effect. Interestingly, individuals are already different on a base level (see control group). While closed minded personality traits are connected to indifference and ignorance, open minded respondents are possibly more interested in the trade-offs between the benefits and risks of AI which could drive this result. Regarding redistribution, there is again a wide difference of attitudes among closed and open minded respondents. However, treatment appears to have an equalizing effect moving all groups but the highly open minded to be more supportive of redistribution.

Finally, the third row of Figure 7 illustrates neuroticism as a moderator between exposure to AI and the will of the people. More relaxed individuals on baseline level agree more with the will of the people compared to more neurotic ones. However, treatment seems to have an equalizing effect again moving the leas to most neurotic individuals on a similar level. Regarding redistribution a less clear image appears as there are clearly no difference among different treatment and personality groups (only slightly on baseline level). Overall, there seems to be relevant heterogeneity regarding personality but weak statistical power limits the interpretation to some extend.

Age & Education

In Figure 8 both interactions, age and education are shown in the upper and lower panel. Age is grouped by approximately 10 year intervals while education is shown for individuals without degree, with degree, and with a higher education degree. From a theoretical perspective both an increasing, a decreasing as well as a reversed U-shape effect could make sense. First one is about the fact that older individuals have harder times adapting to new technology and thus rejecting it. Secondly, older individuals are closer to retirement and thus worry less about the future impact of AI. Thirdly, the middle-aged white-collar workers worry the most as they have the most to lose being relatively far away from retirement and also earn close to their labor market peak income.

Looking at the first row of Figure 8 shows a relatively high degree of homogeneity in terms of treatment effect. Only in the video interaction group older respondents (40-65) are stronger affected by treatment compared to the youngest group (18-29). However, holding treatment constant reveals that age itself is a driver of populist attitudes. It seems that older individuals in general are more populist which is in line with previous findings (Rovira Kaltwasser and Van Hauwaert 2020). In terms of redistribution there is no clear pattern as treatments seem to have different effects albeit not statistical significant. The graphs are more in line with the first expectation above that older respondents are more worried about new technologies as adaptation is harder. Figure 8 lower panel provides evidence about the moderating effect of education.

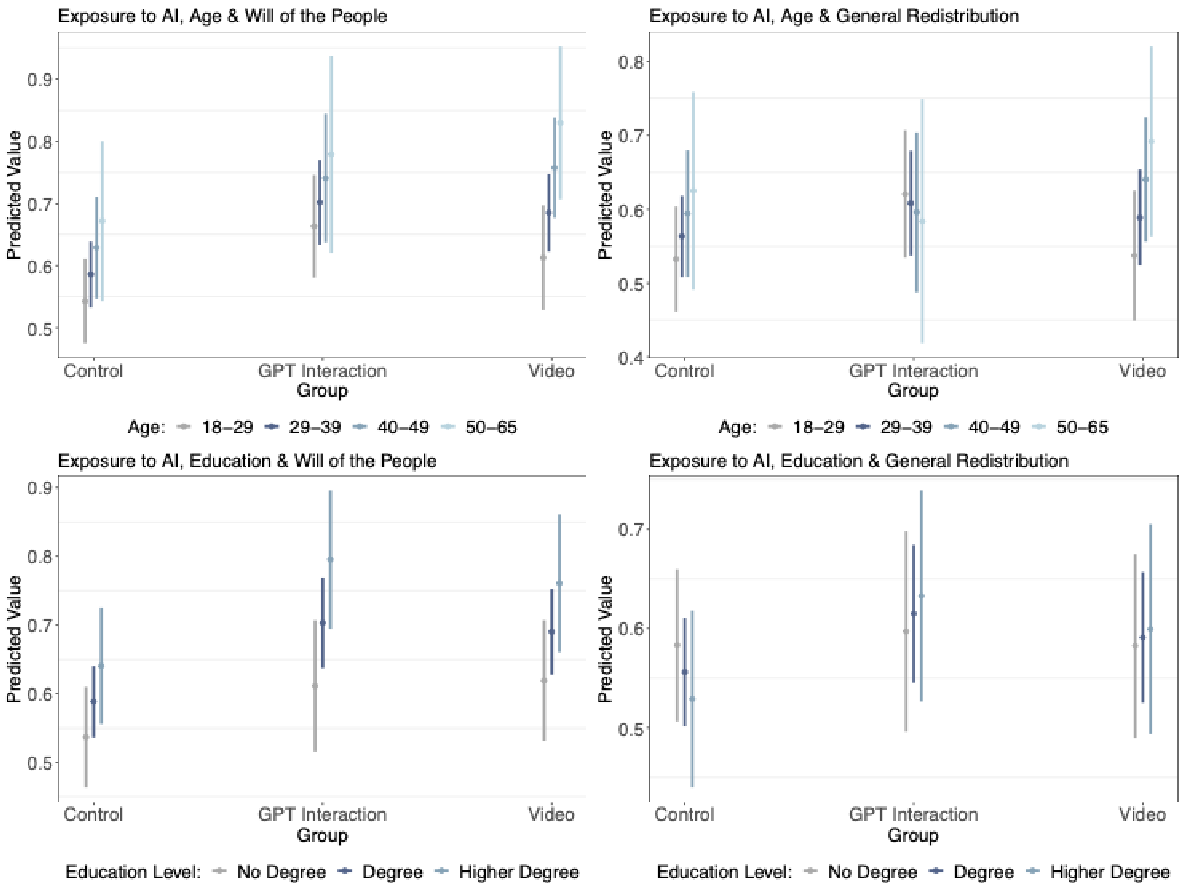


Figure 8: Age & Education and Exposure to Artificial Intelligence

Interestingly, highly educated respondents seem to be more affected by treatment and agree to a higher degree with the populist will. There is already a small difference on baseline level which seems to widen significantly with treatment. This pattern is consistent for both interaction and video group (even if not significant for the video group). In terms of redistribution there is no clear evidence for heterogeneity as demand slightly increases for all groups but indistinguishable from zero.

From a theoretical point of view several explanation could make sense. From the labor market perspective, respondents without a degree are probably less worried about ChatGPT as it only marginally affects through substitution risk. These individuals may have specific and manual skills which are not strongly connected to AI, and in particular, GPT. Specifically, generative AI, such as GPT, could be perceived as less threatening to this group because it has the potential to automate tasks that require less manual skills, such as writing. While manual work is mainly threatened by robotization, LLMs are especially fitted to classic white-collar jobs where educated to highly educated respondents are usually situated (Eloundou et al. 2023). However, it could also be that higher educated individuals are better informed about the debate of AI and its possible consequences which in turn triggers stronger responses. Furthermore, both explanations can be true at the same time which makes it hard to nail down a mechanism.

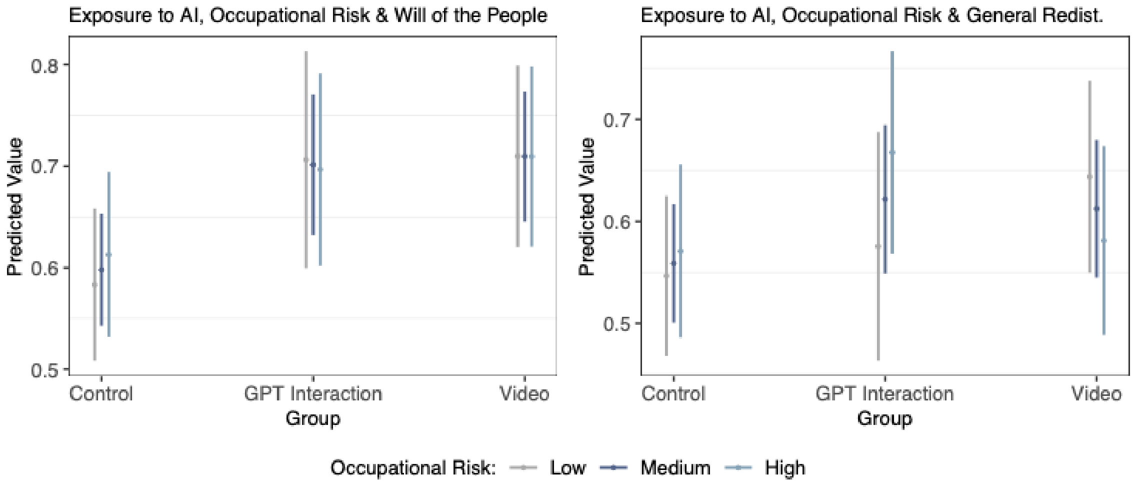
Occupational Risk

Objective vulnerability to automation is one of the main drivers connecting exposure to AI and political behavior. Theoretically higher technological substitution risk should magnify the treatment effect experienced in the main analysis, for both populist attitudes and demand for redistribution.

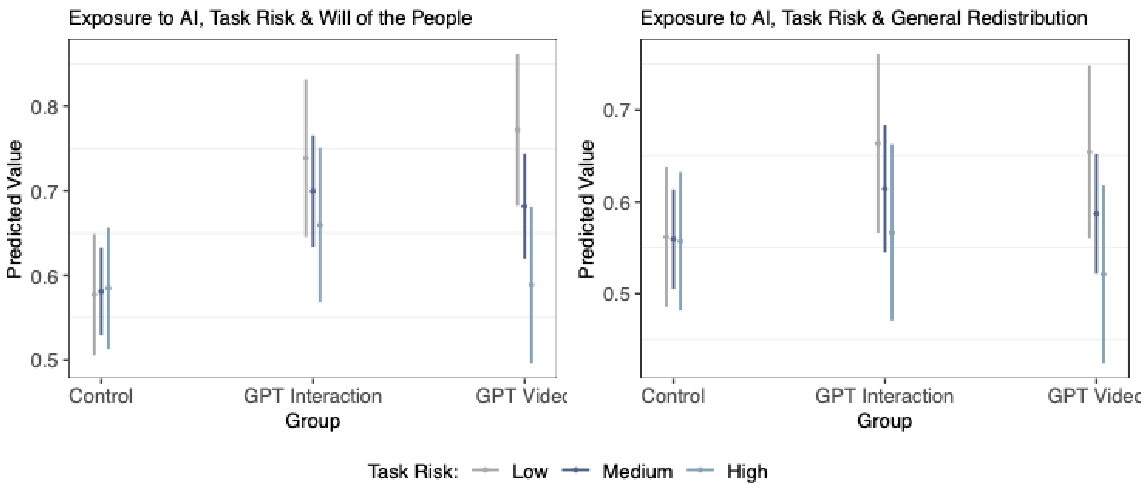
Figure 9 depicts in the upper panel the moderation using an occupation-level measurement of susceptibility developed by Felten, Raj, and Seamans (2023). The measurement indicates the levels of susceptibility to generative AI as GPT is on a continuous dimension between -1.8 and 1.8 where higher levels indicate higher substitution risks. The lower panel illustrates a different measure based on task importance at work adapted to Eloundou et al. (2023). Tasks related to science and logic are negatively correlated to GPTs abilities while writing and coding skills are positively correlated with GPTs abilities. Again higher values indicate a higher substitution risk by AI.

However, Figure 9 does not provide strongly heterogeneous effects. As with personality, the groups are split into one standard deviation below, the average risk, and one standard deviation above the mean. Taking the occupation level into account there is no distinguishable difference among treatment and controls regarding the will of the people. A similar pattern appears for redistribution as confidence intervals are too wide to make any conclusive statements about differences among the groups. One caveat to note is that the regression model for occupation-level risk includes fewer individuals (N = 672) due to non-response. Otherwise, measuring occupational risk at a task level some more heterogeneity appears.

Interestingly, holding treatment constant respondents at higher risks are less supportive of the statement about the general will of the people. The pattern is similar regarding redistribution albeit not statistically significant. One explanation for these results could be on the one side that individuals using programming in their occupations are less worried about GPT as it can also boost productivity. Regarding the importance of writing it could indicate that these individuals are more likely to work in liberal occupations like scriptwriters or journalists. However, there could also be another explanation that individuals just do not know or do not have a sense of the possible impacts of AI. This would be in line with evidence above about education. Finally, there could also the reason of measurement error. There is still no consensus of how to measure the susceptibility of automation by AI. Measurements used here are highly experimental and change nearly daily.



(a) Occupation Level Risk and Exposure to Artificial Intelligence



(b) Task Level Risk and Exposure to Artificial Intelligence

Figure 9: Figure 9: Exposure to AI and Demand for Redistribution

Discussion

The results of this paper point into several directions about the relationship of AI and political behavior. Firstly, regarding the main analysis there seems to be support but only in a very specific dimension. While the will of the people is one of the classical populism measurements in the literature there is no compelling evidence that it is a retro-back to the good old times populism. Otherwise, there would have been stronger evidence regarding norm conformity and a strong leader effect. This is also in line with results about redistribution which pointed in the opposite direction as expected as well as the qualitative evidence about the emotions towards AI. It appears that while people may have an awareness of AI, there is an underlying uncertainty surrounding the technology. In light of these results, the more likely interpretation is that individuals want to take back control in an ever-faster changing world – in a more direct democratic way. This is in line with other findings that showed that individuals actually prefer to be protected from technological change, e.g. through slowing it down (Gallego et al. 2022). Ultimately, this could indicate a growing demand for regulating and managing more actively AI in the future.

The evidence I have provided indicates that there are short-term effects and supply-side factors impacting individuals' preferences in relation to technological change. It is likely that the introduction of new technologies alters preferences quickly. In the case of this survey experiment it is likely that especially individuals with strong priors were affected which could point towards a sort of intensification or polarization of attitudes. In line with this implication is the past status of respondents which could be a possible explanation for modest effects. I also consider the past status of the respondents as a possible explanation for modest effects. For instance, previous winners may be the most likely to be affected by AI (Eloundou et al. 2023), yet they may also be less likely to worry about the transformation due to their relative success in the past.

It is currently unclear why there is no connection between the hypothesized relationship between AI and expected status. This could be attributed to a myriad of reasons, such as different mechanisms, measurement error, time, and the design of the GPT as a product. The first points could mean that AI takes a different direction than previous technological iterations with a multitude of possible channels. Analyzing the qualitative content reveals that there are several possibilities that shape attitudes of individuals regarding AI. Among others, ethical concerns (e.g. deep fakes), surveillance, worries about human knowledge and decision-making, and fears about AI singularity. Regarding the second point, it is possible that the operationalization of expected status is not capturing the theoretical concept or is cognitively too demanding. Furthermore, related the previous point is that worries about expected status need more time to unfold to fully materialize while job loss is relatively easy to imagine. The fourth point is about the design of GPT or AI in general right now. Lastly, the design of AI products means that, even if the salience of the topic is high, the technological iteration still requires human input for most tasks, making it difficult to anticipate any imminent substitution.

Regarding the fear of losing one's job in the future, there seems to be evidence that this channel could actually be relevant, which is in line with Thewissen and Rueda (2019). But for now, this channel is not strong enough for shaping demands for redistribution. This could point towards the possibility that individuals are actually worried about losing their job but not about status decline. An implication could be that individuals when thinking about the future impact of AI are confident that jobs will be substituted but a) new ones will be created or b) welfare states will dampen the negative effects (e.g. universal basic income).

The fifth point is regarding the homogeneous effects of treatment. Previous iterations of technology have demonstrated varying impacts on various groups, but it seems that this difference is either absent or not very pronounced right now. Again, this could be connected with the uncertainty of the actual outcomes of AI on the labor market. There is a strong correlation between treatment and the fear of losing one's job, yet it does not appear to affect any high-risk groups in particular. This could also be, similarly as above, because many high risk individuals nowadays are socialized to be winners and losing out is not really an option for them. At the same time, low risk individuals are not so much worried about losing out because they are not threatened or they already have downgraded – so there is not so much more to lose. However, it is important to note the difficulty in measuring occupational risk, given that there are no established operationalizations that have proven reliable across studies.

Limitations

One important restriction to acknowledge when interpreting the results of this study is the sample composition. The sample doesn't represent the German working age population as a whole; in particular, two major characteristics – age and income – differ substantially. The sample used in this study is relatively young compared to the average working age individual in Germany (30-35 vs. 42; see Bode, Dohse, and Stolzenburg (2023)). Furthermore, the household income is below the average which makes sense looking at the average age (€2400 and €2800 vs. ~ €3800; see Federal Statistic Office Germany (2023)). This limits the interpretation to some extent and overall, treatment effect should be interpreted in a cautious way reflecting young German working individuals. Thus, a population that is less worried about new technologies and has no bigger problems with adopting them. Additionally, as shown in the summary statistics this is also the segment of the population that is more likely to used GPT in the first place.

Two points regarding the treatment should be mentioned. On one hand, I had limited possibilities of testing if individuals truly paid attention to the videos and interacted with the GPT application. Despite this, I was still able to deduce from the aggregate data how much time, on average, respondents spent watching the video (1:30 out of 2 minutes) and how much they used the GPT application, which suggests they paid moderate levels of attention to the task. To further add to this, the fading treatment effects could be problematic, as we included 8 post-treatment variables that could potentially explain the relatively small effect size. In addition, as online survey's respondents answer quality declines at around 7 minutes, treatments could only be designed at a minimum intensity of 2-3 minutes which also limits treatment effects to some extent.

Conclusion

This article explored the effects of artificial intelligence (AI) on political behavior, with a particular focus on populist attitudes and the demand for redistribution. With the increasing presence of AI in the workplace I provided evidence how individuals perceive the possible benefits and risks. The theoretical argument followed that expected status decline affects individuals political behavior. I presented novel experimental evidence including ChatGPT as a treatment to estimate the effect of AI on the demand for redistribution and the prevalence of populist attitudes. The results suggest that exposure to AI indeed affects populist attitudes, especially that the general will of the people should prevail. However, the study did not find evidence that exposure to AI is related to support for redistribution.

Exploring the potential mechanisms behind individuals' preference for retro-politics reveals that it is not necessarily a result of increased demand but rather a skepticism and uncertainty about the future impacts of AI and thus, individuals want to have some sort of control in an ever-faster changing world. While expected status does not seem to be a possible explanation currently, the possibility of job loss does appear to be a plausible mechanism. Furthermore, the treatment effect is quite homogeneous showing only limited evidence that occupational risk by technology is a driver. On the other hand, there is some evidence that personality (extraversion, openness, and neuroticism) explains differences in treatment effects as well as the finding that respondents with higher formal education are more strongly affected by exposure to AI in terms of the general will.

Going forward, there are several avenues of research that can be explored to better understand the complexities of AI and political behavior. Firstly, it is important to delve further into the regulatory framework, both in individual countries and at the European Union level, to test that actual regulations are in line with proposed legislation such as the AI Act. Secondly, this paper provided evidence that AI may shape political behavior through several mechanisms as explored in the qualitative responses. This includes concerns about ethics, surveillance, human decision-making, and fears about AI singularity. Finally, in order to better comprehend the impacts of AI on different sub-populations, it is essential to investigate the lack of pronounced heterogeneity in these impacts through the measuring of occupational risk and the possibility of re-skilling.

This study further emphasizes the need to explore the intricate connection between AI and political behavior in greater depth, as well as to consider the mechanisms, sub-population effects, and generalizability of the findings across different countries and experimental designs. To facilitate this, the inclusion of other countries into the research would allow for an appreciation of the various effects of labor market regulations at different levels, such as the industry or country level. Furthermore, different experimental designs that take occupational risk directly into account can help to identify potential sub-population effects. Moreover, to better distinguish mechanisms, various designs such as conjoint experiments can be used to evaluate how individuals value the proposed mechanisms. Ultimately, this paper highlights the importance of further research on the relationship between AI and political behavior, and its implications on various factors.

Funding Information

This research has received funding from the research project “How Technological Chance Reshapes Politics: Technology, Elections and Policies” (TECHNO), funded by the NOR- FACE program of the European Union (PCI2020-112198) and the Spanish Agency of Research.

References

Acemoglu, Daron. 2021. “Harms of AI.” National Bureau of Economic Research.: 1–57.

Acemoglu, Daron, and Pascual Restrepo. 2018. “The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment.” American Economic Review 108(6): 1488–1542.

———. 2020. “Unpacking Skill Bias: Automation and New Tasks.” AEA Papers and Proceedings 110: 356–61.

Anderson, Cameron, Michael W Kraus, Adam D Galinsky, and Dacher Keltner. 2012. “The Local-Ladder Effect: Social Status and Subjective Well-Being.” Psychological science 23(7): 764771.

Anelli, Massimo, Italo Colantone, and Piero Stanig. 2019. “We Were the Robots: Automation and Voting Behavior in Western Europe.”: 49.

Autor, David H. 2019. “Work of the Past, Work of the Future.” AEA Papers and Proceedings 109: 1–32.

Autor, David H., Frank Levy, and Richard J. Murnane. 2002. “Upstairs, Downstairs: Computers and Skills on Two Floors of a Large Bank.” ILR Review 55(3): 432–47.

———. 2003. “The Skill Content of Recent Technological Change: An Empirical Exploration*.” The Quarterly Journal of Economics 118(4): 1279–1333.

Autor, David, and Anna Salomons. 2017. “Robocalypse Now: Does Productivity Growth Threaten Employment.” In, 45118.

———. 2018. “Is Automation Labor-Displacing? Productivity Growth, Employment, and the Labor Share.” Brookings Papers on Economic Activity 2018(1): 1–87.

Autor, David, Anna Salomons, and Bryan Seegmiller. 2021. “New Frontiers: The Origins and Content of New Work, 1940–2018.” NBER Working Paper: 1–100.

Ballard-Rosa, Cameron, Amalie Jensen, and Kenneth Scheve. 2022. “Economic Decline, Social Identity, and Authoritarian Values in the United States.” International Studies Quarterly 66(1): sqab027.

Ballard-Rosa, Cameron, Mashail A Malik, Stephanie J Rickard, and Kenneth Scheve. 2021. “The Economic Origins of Authoritarian Values: Evidence from Local Trade Shocks in the United Kingdom.” Comparative political studies 54(13): 23212353.

Benedetto, Giacomo, Simon Hix, and Nicola Mastrorocco. 2020. “The Rise and Fall of Social Democracy, 1918–2017.” American Political Science Review 114(3): 928–39. Berman, Sheri, and Maria Snegovaya. 2019. “Populism and the Decline of Social Democracy.” Journal of Democracy 30(3): 5–19.

Bode, Eckhardt, Dirk Dohse, and Ulrich Stolzenburg. 2023. “Aging and Regional Productivity Growth in Germany. Alterung Und Regionales Produktivitätswachstum in Deutschland.” Review of Regional Research.

Brynjolfsson, Erik, Danielle Li, and Lindsey R. Raymond. 2023. “Generative AI at Work.” NBER WORKING PAPER SERIES: 1–58.

Bullock, John G, and Donald P Green. 2021. “The Failings of Conventional Mediation Analysis and a Design-Based Alternative.” Advances in Methods and Practices in Psychological Science 4(4): 25152459211047227.

Busemeyer, Marius R., and Alexander H. J. Sahm. 2022. “Social Investment, Redistribution or Basic Income? Exploring the Association Between Automation Risk and Welfare State Attitudes in Europe.” Journal of Social Policy 51(4): 751–70.

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock. 2023. “Gpts Are Gpts: An Early Look at the Labor Market Impact Potential of Large Language Models.” arXiv preprint arXiv:2303.10130.

Esping-Andersen, Gøsta. 1990. The Three Worlds of Welfare Capitalism. Cambridge, UK: Polity Press.

Eurostat. 2023. “Use of Artificial Intelligence in Enterprises.” https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Use_of_artificial_intelligence_in_ enterprises.

Fawzi, Nayla. 2019. “Untrustworthy News and the Media as “Enemy of the People?” How a Populist Worldview Shapes Recipients’ Attitudes Toward the Media.” The International Journal of Press/Politics 24(2): 146164.

Federal Statistic Office Germany. 2023. “Federal Statistic Office Germany.” https://www.destatis.de/EN/Themes/Society-Environment/Income-Consumption-Living-Conditions/Income-Receipts-Expenditure/Tables/income-expenditure-d-lwr.html. Felten, Edward

W., Manav Raj, and Robert Seamans. 2023. “The Occupational Impact of Artificial Intelligence: Labor, Skills, and Polarization.”

Festinger, Leon. 1954. “A Theory of Social Comparison Processes.” Human relations 7(2): 117140.

Frey, Carl Benedikt, Thor Berger, and Chinchih Chen. 2018. "Political Machinery: Did Robots Swing the 2016 US Presidential Election?" Oxford Review of Economic Policy 34(3): 418–42.

Gallego, Aina, Alexander Kuo, Dulce Manzano, and José Fernández-Albertos. 2022. "Technological Risk and Policy Preferences." Comparative Political Studies 55(1): 60–92.

Gallego, Aina, Thomas Kurer, and Nikolas Schöll. 2021. "Neither Left Behind nor Super- star: Ordinary Winners of Digitalization at the Ballot Box." The Journal of Politics.

Gest, Justin, Tyler Reny, and Jeremy Mayer. 2018. "Roots of the Radical Right: Nostalgic Deprivation in the United States and Britain." Comparative Political Studies 51(13): 16941719.

Gingrich, Jane, and Silja Häusermann. 2015. "The Decline of the Working-Class Vote, the Reconfiguration of the Welfare Support Coalition and Consequences for the Welfare State." Journal of European Social Policy 25(1): 50–75.

Häusermann, Silja, and Hanna Schwander. 2012. "Varieties of Dualization? Labor Market Segmentation and Insider-Outsider Divides Across Regimes." The age of dualization: The changing face of inequality in deindustrializing societies: 27–51.

Hemerijck, Anton. 2018. "Social Investment as a Policy Paradigm." Journal of European Public Policy 25(6): 810–27.

Im, Zhen Jie, Nonna Mayer, Bruno Palier, and Jan Rovny. 2019. "The "Losers of Automation": A Reservoir of Votes for the Radical Right?" Research & Politics 6(1): 16.

Im, Zhen Jie, Hanna Wass, Anu Kantola, and Timo M Kauppinen. 2023. "With Status Decline in Sight, Voters Turn Radical Right: How Do Experience and Expectation of Status Decline Shape Electoral Behavior?" European Political Science Review 15(1): 116135.

Imai, Kosuke, Luke Keele, and Dustin Tingley. 2010. "A General Approach to Causal Mediation Analysis." Psychological Methods 15(4): 309–34.

Imai, Kosuke, Dustin Tingley, and Teppei Yamamoto. 2013. "Experimental Designs for Identifying Causal Mechanisms." Journal of the Royal Statistical Society: Series A (Statistics in Society) 176(1): 551.

Jackman, Mary R., and Robert W. Jackman. 1973. "An Interpretation of the Relation Between Objective and Subjective Social Status." American Sociological Review 38(5): 569–82.

Kim, Hyunji, Mitchell J. Callan, Ana I. Gheorghiu, and William J. Matthews. 2017. "Social Comparison, Personal Relative Deprivation, and Materialism." British Journal of Social Psychology 56(2): 373–92.

Kurer, Thomas. 2020. "The Declining Middle: Occupational Change, Social Status, and the Populist Right." Comparative Political Studies 53(10-11): 1798–1835.

Kurer, Thomas, and Bruno Palier. 2019. "Shrinking and Shouting: The Political Revolt of the Declining Middle in Times of Employment Polarization." Research & Politics 6(1).

Kurer, Thomas, and Briitta Van Staalduinen. 2022. "Disappointed Expectations: Down- ward Mobility and Electoral Change." American Political Science Review 116(4): 1340–56.

Lungu, Laura Silvia. 2022. "Bling-Bling Politics: Exposure to Status-Goods Consumption

Shapes the Social Policy Preferences of the Less Affluent." Socio-Economic Review. Moene, Karl Ove, and Michael Wallerstein. 2001. "Inequality, Social Insurance, and Redistribution." American political science review 95(4): 859874.

Mokyr, Joel. 2018. "The Past and the Future of Innovation: Some Lessons from Economic History." Explorations in Economic History 69: 13–26.

Mokyr, Joel, Chris Vickers, and Nicolas L. Ziebarth. 2015. "The History of Technological Anxiety and the Future of Economic Growth: Is This Time Different?" Journal of Economic Perspectives 29(3): 31–50.

Noy, Shakked, and Whitney Zhang. 2023. "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence." Available at SSRN 4375283.

Oesch, Daniel, and Line Rennwald. 2018. "Electoral Competition in Europe's New Tripolar Political Space: Class Voting for the Left, Centre-Right and Radical Right." European Journal of Political Research 57(4): 783–807.

Pierson, Paul. 1996. "The New Politics of the Welfare State." World Politics 48(2): 143–79.

Rammstedt, Beatrice, and Oliver P. John. 2007. "Measuring Personality in One Minute or Less: A 10-Item Short Version of the Big Five Inventory in English and German." Journal of Research in Personality 41(1): 203–12.

Rosenberg, Morris. 1953. "Perceptual Obstacles to Class Consciousness." Soc. F. 32: 22. Rovira Kaltwasser, Cristóbal, and Steven M. Van Hauwaert. 2020. "The Populist Citizen: Empirical Evidence from Europe and Latin America." European Political Science Review 12(1): 1–18.

Sacchi, Stefano, Dario Guarascio, and Silvia Vannutelli. 2020. "Risk of Technological Unemployment and Support for Redistributive Policies." In eds. Romana Careja, Patrick

Emmenegger, and Nathalie Giger. Wiesbaden: Springer Fachmedien, 277–95. Schneider, B. 2023. "Technological Unemployment in the British Industrial Revolution: The Destruction of Hand Spinning."

Schwander, Hanna, and Silja Häusermann. 2013. "Who Is in and Who Is Out? A Risk- Based Conceptualization of Insiders and Outsiders." Journal of European Social Policy 23(3): 248–69.

Sivanathan, Niro, and Nathan C. Pettit. 2010. "Protecting the Self Through Consumption: Status Goods as Affirmational Commodities." Journal of Experimental Social Psychology 46(3): 564–70.

Steenvoorden, Eefje, and Eelco Harteveld. 2018. "The Appeal of Nostalgia: The Influence of Societal Pessimism on Support for Populist Radical Right Parties." West European Politics 41(1): 2852.

Tajfel, Henri, and John Turner. 1979. "An Integrative Theory of Intergroup Conflict." In OUP Oxford, 33–47.

Thal, Adam. 2020. "The Desire for Social Status and Economic Conservatism Among Af- fluent Americans." American Political Science Review 114(2): 426442.

Thewissen, Stefan, and David Rueda. 2019. "Automation and the Welfare State: Techno- logical Change as a Determinant of Redistribution Preferences." Comparative Political Studies 52(2): 171–208.

Turner, John C, and Rina S Onorato. 1999. "Social Identity, Personality, and the Self- Concept: A Self-Categorizing Perspective."

Velandia-Morales, Andrea, Rosa Rodrıguez-Bailón, and Roćıo Mart´ınez. 2022. "Economic Inequality Increases the Preference for Status Consumption." Frontiers in Psychology 12.

Webb, Michael. 2019. "The Impact of Artificial Intelligence on the Labor Market." SSRN Scholarly Paper: 1–61.

Wu, Nicole. 2021. "Misattributed Blame? Attitudes Toward Globalization in the Age of Automation." Political Science Research and Methods: 1–18.

———. 2022. "Restrict Foreigners, Not Robots: Partisan Responses to Automation Threat." Economics & Politics: 1–24.