Rage Against the Machine? Generative AI Use, Subjective Risk, and Policy Preferences

Matthias Haslberger
University of St. Gallen

Jane Gingrich
University of Oxford

Jasmine Bhatia
Birkbeck, University of London

July 16, 2024

Abstract

Scholars interested in the effect of automation on policy preferences have commonly argued that (subjective) risk predicts demand for insurance. Generative AI potentially challenges this dynamic. Based on a pre-registered online experiment with a near-representative sample of 1,041 UK working-age adults we show that direct exposure to generative AI does not increase subjective risk and leads to more positive attitudes towards the technology. Despite this, treated respondents show greater support for progressive social policy and place themselves politically further left, indicating that sociotropic preferences dominate self-interest. Text analysis of an open-ended question shows thoughtful engagement with the implications of AI. This article provides a first big-picture investigation of the political implications of generative AI and outlines avenues for further research.

We are grateful for the feedback and support we received from Noah Bacine, Patrick Emmenegger, and Lukas Paleckis on the design of the survey, as well as to Simone Tonelli and participants of the GOVPET Advisory Board Meeting 2024, the Zurich Political Economy Seminar Series, the Freiburg Economics Research Seminar, the VPET-ECON Conference 2024, the SASE Conference 2024, and the CES Conference 2024 for feedback on various versions of the paper. All remaining errors are our own. This project has been funded in part by the Canadian Institute for Advanced Research (CIFAR) and the Schweizerisches Staatssekretariat für Bildung, Forschung und Innovation (SERI) in the framework of the GOVPET research project. The authors made use of ChatGPT (GPT-3.5 and GPT-4) in developing the code for various figures and to create example answers to the open-ended question to use for comparison.

Corresponding author: Matthias Haslberger (matthias.haslberger@unisg.ch)

Word count: ca. 8,530
Competing interests: The authors declare none.
Data availability: The data will be made accessible on a suitable repository following acceptance of the manuscript.

Introduction

Does using generative AI make people or more less concerned about disruptive effects — either for their own jobs or those of others? How do people think the use of AI technology ought to be regulated by the government? As more people use generative AI at home and at work, the answers to these questions will have profound implications for the politics of advanced economies — shaping in turn the demand for social and labor market policies. While the public and academic discourse often highlights the possible dangers of generative AI — mass unemployment and political upheaval — little rigorous evidence currently exists as to how users respond to it, and even less is known about how AI use and socio-demographic characteristics interact to shape individuals' perceptions, expectations, and preferences.

Our study addresses these questions using an experimental survey of a near-representative sample of 1,041 working-age adults in the UK. After randomizing respondents into a treatment and control group that perform three realistic work tasks either with or without the help of generative AI, we ask them about their risk perceptions, general attitudes towards generative AI, and social policy preferences. This study, fielded in July 2023, before the mass adoption of generative AI in most workplaces, allows us to observe the causal effects of exposure to technology in complex but realistic tasks on political preferences for a near-representative sample of the general population. As such, it complements existing studies that either examine occupation-specific samples (Dell'Acqua et al., 2023; Noy and Zhang, 2023; Brynjolfsson, Li and Raymond, 2023), do not consider political preferences (Haslberger, Gingrich and Bhatia, 2023; Dell'Acqua et al., 2023; Brynjolfsson, Li and Raymond, 2023), or look at other forms of AI (Raviv, 2023; Margalit and Raviv, 2023).[1] We supplement this experimental design with a quantitative analysis of open-ended survey responses to shed light on how people reason about appropriate policy responses to the labor market effects of generative AI. This study thus goes beyond existing research by drawing on a cross-section of the UK working age population and directly investigating the impact of first-hand experience of transformative new AI technology on political preferences.

Despite the emphasis on possible dangers in the public debate, we find that individuals who use generative AI (ChatGPT) do not perceive their jobs to be at greater risk of replacement by technology or a more tech-savvy person. Furthermore, AI users develop significantly more positive attitudes towards AI on a number of issues: they are more likely to see it as beneficial to themselves and to society as a whole, and less likely to see it as a threat to humanity or to support government limitations on the use of AI. While use of AI does not increase perceived personal economic risks, we do find that treated respondents express more progressive social policy preferences and identify as more politically left-wing. This constellation of

[1]Generally speaking, existing studies in economics focus on the productivity effects of AI, while studies in political science have looked at AI use in decisions that affect citizens, but not AI use by the respondents themselves.

preferences contradicts established work on the relationship between technology and policy preferences which suggests that the relationship runs largely through changing risk perceptions (see, e.g., Gallego and Kurer, 2022). To make sense of these findings, we argue that AI use may activate sociotropic concerns, whereby individuals who expect to benefit from the technology see it as providing them with personal gains but collective disruption, leading them to support government intervention for those who — through no fault of their own — are negatively affected by it.

The paper is structured as follows. Section 2 draws on the existing literature on technological exposure and risk to develop a set of hypotheses about its potential effects. It then speculates as to how generative AI might have different effects on preferences than other forms of technology. Section 3 outlines the design of the survey and our analytical approach. In section 4, we present the results and a final section concludes by discussing the implications of our findings and ways forward for research on the political implications of generative AI.

Argument

The Risk-Insurance Framework in Political Economy Research

Much existing research on the relationship between technological change and social policy preferences conceptualizes it in two-steps: how technological exposure shapes individual material risk and how material risks shape preferences.[2] First, technological exposure affects individuals' perceptions of risk, changing either their perceptions of job security, quality, or income prospects. Analysts look to measure risks in varying ways. Some studies focus on subjective risk measures such as self-assessed unemployment risk, whereas others employ "objective" measures of exposure such as routine task intensity (RTI) scores at the occupation or sector level. Second, this work looks at the relationship between exposure and preferences as mediated by this risk. Some authors focus on compensatory policies (such as unemployment benefits or a universal basic income) while others study activation measures (for example, retraining opportunities and active labour market policies). The underlying logic, however, that people demand insurance against labor market risks posed by technology in line with their degree of exposure, remains the same. We refer to this framework as the "risk-insurance framework".

While the risk-insurance framework builds on well developed theoretical models of both technological shocks and demand for insurance (Rehm, 2011), recent wide-ranging scholarly reviews by both Gallego and Kurer (2022) and Weisstanner (2023) show less empirical consensus on the explanatory power of the

[2] An alternative approach highlights the misattribution of negative consequences of technological change to international trade and immigration (Gallego and Kurer, 2022; Wu, 2022, 2023). However, this perspective is more relevant for explaining populist voting than social policy preferences.

risk-insurance framework. This body of work shows that people tend to favour compensatory policies over activating ones. Only a handful of studies find a positive association between technological risk exposure and support for activating policies (Im, 2021), while others report a null or even negative relationship (Kurer and Häusermann, 2022; Busemeyer and Tober, 2022; Busemeyer and Sahm, 2022). By contrast, a majority of studies looking at passive labor market policy or support for redistribution do find a positive association between technological risk and support (Kurer and Häusermann, 2022; Busemeyer and Tober, 2022; Sacchi, Guarascio and Vannutelli, 2020; Dermont and Weisstanner, 2020; Thewissen and Rueda, 2019), although some also report null findings. There are no clear differences in patterns based on whether studies use subjective or objective risk measures — Kurer and Häusermann (2022), Busemeyer and Tober (2022), and Gallego et al. (2022) find similar effects across these types of measures, while Sacchi, Guarascio and Vannutelli (2020) find an effect of RTI but not subjective risk on support for passive labor market policy. On balance, existing research suggests that technological exposure leads to greater support for passive, compensatory policies than activating ones, and that contextual factors such existing welfare state institutions or education systems may play a moderating role (Busemeyer and Sahm, 2022; Gingrich and Ansell, 2012; Haslberger, Emmenegger and Durazzi, 2023).

However, only a small number of studies have looked to experimentally identify the effects of technological exposure on preferences. In these designs, where the analyst manipulates exposure to technology or information about the effects of technology, the linkage to social policy preferences is more modest. Gallego et al. (2022) use a priming experiment and find exposure to information about technological risks did not increase support for compensatory policies but instead for measures slowing down technological change. Thus, they find an effect on regulatory preferences, a margin of adaptation that has been largely neglected in the literature so far. Zhang (2022) shows in a series of studies that information about automation risks — either general or job-specific — did not affect policy preferences, even as it changed factual beliefs about automation risks. Overall, the limited evidence from experimental studies shows at best little support for the risk-insurance framework. However, these weak findings may be due to the increasingly recognized limitations of information provision experiments, which often fail to induce behavioural responses even as real technological exposure does (Ciani, Freget and Manfredi, 2021).

We are thus left with two key questions: does actual exposure to generative AI increase respondents' perceptions of individual risk? And do these risk perceptions matter for respondents' preferences towards compensatory or regulatory policies? Our experimental design, involving direct exposure, and dis-aggregate risk and preference items, aims to overcome these limitations and provide a clear test of the risk-insurance framework as it relates to generative AI.

Application to Generative AI

Substantial academic debate about the effects of generative AI tools builds on the risk-insurance approach. This work focuses on identifying the potentially large risks of AI for employment — particularly amongst groups of workers insulated from past rounds of automation (Future of Life Institute, 2023; Frey and Osborne, 2023). In particular, there is speculation that people in white collar jobs, who have until now benefited from technological change, might be at increased risk from generative AI (Felten, Raj and Seamans, 2023a). These debates, however, remain relatively removed from the wider public — existing research shows that AI governance is not yet highly politicized (Margalit and Raviv, 2023; Schiff, Schiff and Jacobson, 2024). Citizens have low levels of information about the effects of AI, and their preferences remain uncertain.

In this "pre-political" environment, we expect direct exposure to powerful new AI technology in a work-like setting to provide substantial novel information to survey respondents about its effects. If this novel information increases respondents' perception that AI tools could threaten their job stability or long-run income prospects — i.e. increases their subjective perception of risk — then we would expect rational, self-interested individuals to also expand their support for the state and for policies to alleviate these risks. Possible measures include compensatory policies such as unemployment benefits, activating ones such as retraining and continuing education, and hybrid approaches such as a job guarantee. While economists typically advocate skill-focused policies (Felten, Raj and Seamans, 2023b), compensatory measures are likely to be more popular with the public (see Weisstanner 2023). Gallego et al. (2022) show that policies to regulate and slow down the adoption of new technologies might also be popular with electorates. By contrast, if respondents see AI as enhancing their skills and long-run job prospects it could reduce their demand for these same programs, as they would be less likely to benefit from them.

As outlined above, the bulk of the existing literature suggests that the former effect is likely to hold. Increased experience of the capacities of generative AI should increase subjective unemployment risk and thus demand for social programs. While these policy preferences might be expected to align with higher support for left-wing political parties, several studies from the United States and Europe find the opposite: those most negatively affected by technological change are more likely to support the populist right or other anti-status quo parties (Gallego and Kurer, 2022; Kurer, 2020). Economic winners, by contrast, tend to support the political status quo (Gallego, Kurer and Schöll, 2022). As this group generally holds progressive values on non-economic issues, their overall political preferences may be more likely to lean left (Kitschelt and Rehm, 2014). Although the literature is mixed on how technological change affects political orientation, given that AI has not yet acquired a "second-dimension" partisan framing (Adler and Ansell, 2020), our core expectation is that exposure should shape preferences on the "first-dimension", i.e. making people who support more redistribution also more politically left-wing.

Based on this reading of the extant literature, we designed an experiment that looked to assess the effects of generative AI on preferences. In the experiment, respondents perform realistic work tasks, and are randomized into a group that is encouraged to perform these tasks with the help of generative AI and a group that is encouraged to refrain from using it. Via exposure to AI, the treatment group thus receives tangible information about its capabilities in relation to particular forms of work. We hypothesize that this exposure should provide new knowledge about the risks that the technology entails for their livelihoods and consequently shape their preferences regarding policies that mitigate economic risks. Accordingly, we pre-registered the following hypotheses:[3]

H1.1: Respondents who use generative AI are *more likely* to fear job loss due to technology.

H1.2: Respondents who use generative AI express *more negative* attitudes towards AI.

H1.3: Respondents who use generative AI exhibit *more progressive* social policy preferences.

An Alternative Perspective Based on Sociotropic Preferences

The preceding discussion assumed that people's responses to generative AI should fundamentally mirror previous instances of automation - meaning that on average, it will increase risk perceptions. However, generative AI may be different from previous technological change for two linked reasons.

First, to early adopters of generative AI, usage may reveal information that reduces, rather than increases, anxiety. Busemeyer et al. (2023) find, using cross-national survey data on risks related to automation, a weak tendency for workers who use ICT more frequently to express a lower subjective technology risk. Using generative AI might similarly decrease our participants' perceived employment risk, and instead increase their confidence that they will personally benefit from generative AI. As participants in our study, they are by definition early adopters and may not consider themselves to be at risk — after all, they are ahead of the curve with this groundbreaking technology.[4] Thus, our treatment might lead to more positive attitudes, as participants think only about how the technology might benefit them personally, rather than how it might threaten their job.

Second, treated respondents receive information not just about their own performance, but the capacity of generative AI more broadly. Even where respondents observe their own performance improve with AI, it is not clear that they will interpret these processes as socially beneficial. They could see a) risks that other,

[3] The pre-analysis plan is available under https://osf.io/dgh2u/?view_only=f30331b363c64ab883d9f76ae0565c68

[4] This is even true of people who used generative AI for the first time during our study, as the majority of the general population had not used the technology at the time of the survey.

less tech-savvy workers will be negatively affected by the technology while at the same time b) technological changes will bring productivity gains large enough to financially compensate losers. Alternatively, treated respondents might simply perceive AI as useful and not give much thought to the possible implications for others, although this appears unlikely. In this scenario, we argue that people who anticipate negative consequences for others have good reasons to transcend narrow self-interest and support progressive social policies. We call this the sociotropic model of the politics of generative AI.

Why, precisely, might people increase their support for progressive policies, if not out of material self-interest? As Compton and Lipsmeyer (2019, 540) note, "understanding only the role of self-interest in shaping preferences for social insurance leaves scholars with a limited picture of preferences." Indeed, there is substantial evidence that redistributive preferences are not purely material; people support social policies that they do not expect to benefit from themselves where they view recipients as deserving (Cavaillé, 2023). Generally speaking, when people fall on hard times through no fault of their own, others are more willing to help them. Several studies find that people exhibit greater support for social insurance when macroeconomic conditions are difficult — even if they are not themselves unemployed (Kam and Nam, 2008; Funk, 2000). Compton and Lipsmeyer (2019) qualify that individual insecurity is a stronger motivator, but that people who do not consider themselves at risk may nevertheless support redistributive policies if they expect that others may — through no fault of their own — experience economic difficulties.

Moreover, inequality has negative externalities (Støstad and Cowell, 2024). Widespread unemployment and poverty have negative consequences even for those who are not unemployed (Hakim, 1982), which can make it rational to support redistribution even for people who are not personally at risk of needing the benefits. For example, Rueda and Stegmueller (2016) find that fear of crime makes rich people in highly unequal regions more supportive of redistribution than their peers in more equal areas. Similar motivations may induce people who consider generative AI a good thing for themselves and society to increase their support for redistributive social policies. Taken together, this set of arguments leads to a conflicting set of theoretical predictions which were, however, not pre-registered:

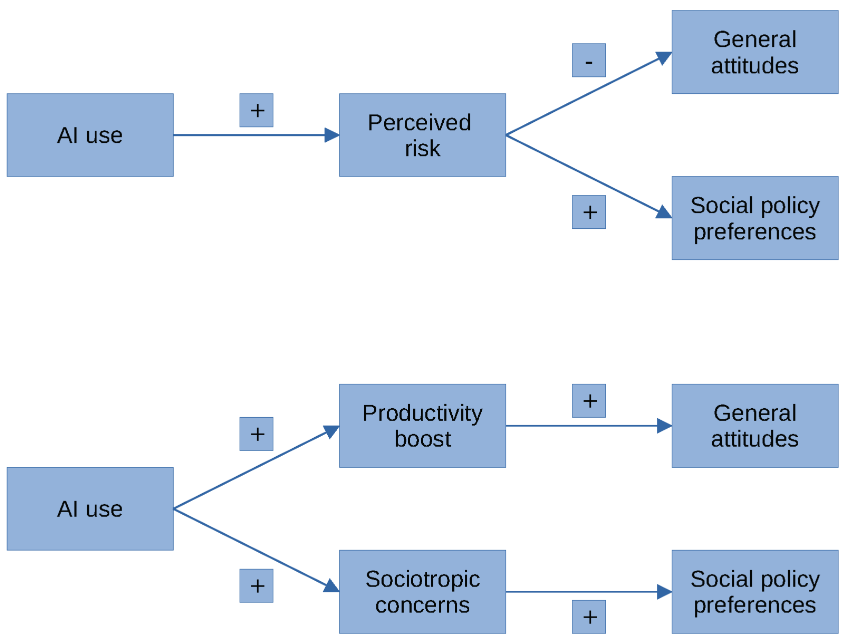
H2.1: Respondents who use generative AI are *no more likely* to fear job loss due to technology.

H2.2: Respondents who use generative AI express *more positive* attitudes towards AI.

H2.3: Respondents who use generative AI exhibit *more progressive* social policy preferences.

Note that H2.3 is identical to H1.3, as the prediction for social policy preferences is the same; merely the causal path is different: sociotropic concerns rather than material self-interest. We schematize these two contrasting sets of expectations in Figure 1. In the first scenario, which corresponds to the standard narrative in political economy, AI use increases people's perceived individual risk, based on which they adopt more negative attitudes towards AI technology and become more supportive of social policies. In the second scenario, AI use does not affect perceptions of individual risk, but leads to more positive attitudes as a consequence of people experiencing the technology as useful. At the same time, AI use activates sociotropic concerns for others who might be at risk, which then positively affects support for social policy.

Figure 1: Two sets of theoretical expectations



Note: The top panel shows the standard political economy view focusing on individual risk. The bottom panel schematizes an alternative perspective where AI use affects general attitudes towards AI through a productivity boost and social policy preferences due to sociotropic concerns.

Methods

Survey Design

Prior to data collection, our study received ethics clearance from the University of Oxford Department of Social Policy and Intervention's Ethics Committee. It adheres to APSA's Principles and Guidance for Human Subjects Research.[5] Through the survey company YouGov, we recruited a sample of 1041 respondents who

[5] For details, see Appendix A.

Table 1: Overview of the tasks

Task	Description	Skills tested
Email	Improve email addressing a hypothetical workplace dispute	Professional communication, diligence, knowledge of spelling and grammar
Assessment	Evaluate persuasiveness of two texts presenting opposing views	Appraisal of conflicting information, persuasive reasoning
Comprehension	Answer questions about a complex text	Dealing with complex information, precision

had ChatGPT accounts and who are largely representative of the UK working-age population.[6] This makes our study the first that combines direct experimental exposure to generative AI and a full-population sample, making it suitable for inferences about attitudes and policy preferences. The sample was split into a treatment group (N = 504) which was encouraged to use ChatGPT and a control group (N = 537) which was instructed to complete the survey without using ChatGPT. Fieldwork for the survey took place from 19th – 28th July 2023, following a pre-test in early July 2023. The median time for completion of the survey was approximately 28 minutes. Participants were informed before undertaking the survey that they would be compensated at twice the normal YouGov rate for a survey of this length, owing to the higher cognitive demands placed on respondents. To further increase participant effort, we offered a performance-based reward amounting to 50% of the participation reward to the top 10% of respondents in each group.

After some self-assessment questions, participants were asked to complete three short text-based tasks of increasing complexity. Table 1 provides an overview and Appendix C describes them in more detail. This intervention was designed to showcase the capabilities of generative AI in tasks that many people are familiar with and that are part — at least in structurally similar form — of many work routines, such as communicating, appraising arguments, and making sense of complex information.[7] By directly exposing treated participants to the technology in a work-like setting, the treatment provides more tangible information about the risks generative AI may pose to people's jobs than information provision alone could. With this design we address some of the limitations to the external validity of information experiments (Barabas and Jerit, 2010). All tasks were designed so that it was possible to get a perfect answer both with and without using AI. The tasks were followed by a set of questions about people's attitudes and expectations regarding AI and redistributive policy. YouGov provided demographic information, including age, sex, education, occupation, income, and other variables.

[6] Additional details about the survey, including summary statistics, are provided in Appendix B. Despite their best efforts, YouGov could not at the time provide a sample of ChatGPT users that is fully representative of the general population. For example, men are over-represented in the sample, reflecting differences in ChatGPT uptake (see also Haslberger, Gingrich and Bhatia 2023). Yet, when we use weights to account for such imbalances, our results remain unaffected.

[7] Figure C1 shows that most respondents perform similar tasks at work at least sometimes.

Analytical Procedure

Post-treatment, respondents encountered a range of attitudinal items. With this broad set of questions, we intend to provide an overview of people's concerns and preferences to serve as a starting point for future research. Table 2 lists the five question blocks. We first asked about automation-related subjective job risk, followed by questions eliciting general views on AI. We next asked about social policy preferences. The order of these questions was randomized within question blocks. Next, we obtained respondents' political self-identification on a left-right scale. The survey concluded with an open text question about people's priorities when it comes to dealing with the effects of AI on the labour market.

Table 2: Outcome measures

Questions	**Measurement**
Subjective risk: How likely, or unlikely, do you think it is that the following will happen to your job (or job opportunities) over the next 5 years? • My job will be replaced by an artificial intelligence, algorithm, computer software, or robot. • I will lose my job because I am not good enough with new technologies or because I will be replaced by someone with better technological skills.	5-point scale, dichotomized
General attitudes: Please tell us to what extent you agree or disagree with the following statements. • People like me are likely to benefit from AI. • Products and services using AI have more benefits than drawbacks. • The government should limit the use of some forms of AI. • AI poses a serious threat to humanity.	5-point scale, dichotomized
Social policy preferences: People have different views on what the responsibilities of governments should or should not be. For each of the following, please tell us if it should, or should not be the government's responsibility to... • ...ensure a job for everyone who wants one? • ...ensure a reasonable standard of living for the unemployed? • ...ensure training opportunities for everyone who wants to upgrade their skills?	5-point scale
Political self-identification: On a scale from 0 to 10, where 0 is "extremely left-wing" and 10 is "extremely right-wing", how would you describe your political orientation?	11-point scale
Policy priorities: Please briefly tell us in your own words what you think is most important when it comes to dealing with the effects of artificial intelligence on the labour market.	Open-text

We dichotomize the outcome variables in the first two question blocks and estimate linear probability models which are straightforward to interpret as the probability of agreeing with a statement. For the social policy and political self-identification questions, we retain the original 5-point and 11-point scales, so that the coefficients represent a point change on the respective scale. For the main analyses, we regress the outcome on a treatment dummy. To analyze heterogeneous effects, we interact the treatment dummy with a dummy for the characteristic of interest.

Given our research design, the effects sizes we estimate are likely to be conservative: since the entire sample is recruited from people with a ChatGPT account, even the control group has at least some prior awareness of the capabilities of generative AI. This was necessary since at the time it was not possible to embed ChatGPT directly in the survey. Insofar as there is a "wow" effect for those who are newly exposed to the technology, we should therefore find it more subdued than for fully novice users. However, with a view to the future adoption of AI in the workplace, this design strikes us as a more realistic scenario than assuming zero prior knowledge of AI. Furthermore, individuals who are broadly aware of the capabilities of generative AI are more likely to have considered its societal implications and to give non-random answers to our questions (Converse, 2006).

We round off the analysis by performing sentiment analysis and running structural topic models on the open text question at the end of the survey. Questions where respondents can explain their reasoning contribute additional nuance to survey data as they are less guided than multiple choice questions. Modern text analysis techniques can then be used to elicit first-order concerns regarding policy issues, akin to their widespread use on non-survey data (Ferrario and Stantcheva, 2022). Sentiment analysis can be used to judge the overall tone of a response. Topic modelling can identify latent themes and so indicate for example which issues might become politically salient. These additional analyses help unpack our headline findings by providing insights into the reasoning of the respondents and their policy priorities, which can in turn inform the design of future studies.

Results

Evidence Supports the Sociotropic Model

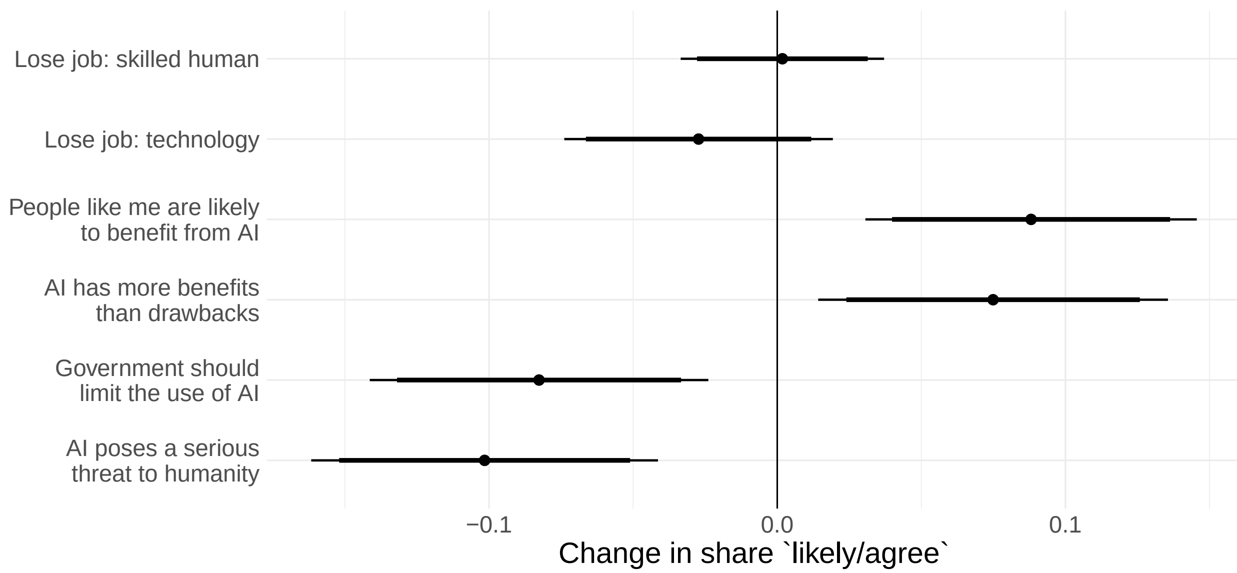
Existing research on the technology-social policy nexus reports ambiguous findings (Gallego and Kurer, 2022; Weisstanner, 2023). Thus, an experimental intervention that directly exposes individuals to generative AI in a work-like setting, highlighting the capabilities and making tangible the risks of generative AI, followed by questions that are widely used in welfare state scholarship, promises important insights.

Overall, the evidence from our study supports the sociotropic model over the risk-insurance model. Figure 2 shows that, compared to the control group, respondents in the treatment group are on average unaffected in their perceptions of subjective risk, similar to what Busemeyer et al. (2023) found for ICT use, and supporting H2.1 over H1.1. There are two plausible explanations for this finding: either respondents found generative AI so useful that they expect to benefit from its proliferation, not making the leap to consider possible substituting effects, or they found the capabilities of ChatGPT underwhelming, eliminating it as a potential source of job risk — but the vast majority of AI users rated it as (very) useful which makes the first scenario seem more likely. Still, it is surprising as the questions explicitly ask respondents about possible risks associated with AI technology. The absence of a link between AI use and subjective risk deals a major blow to the risk-insurance framework as this is the first step in the causal chain posited by the model (see Figure 1). Moreover, H2.2 is supported over H1.2, as people in the treatment group express more positive attitudes towards AI. Treated respondents are 5 to 8 percentage points more likely to see AI as a positive force for themselves and in general, and 7 to 10 percentage points less likely to see AI as a threat which should be restricted by the government. That treated respondents are more likely to say that they expect to benefit personally from using AI further indicates that they did not consider the potential second-order effects of a productivity-increasing technology that is capable of substituting vast number of workers. Interestingly, that people become less supportive of government limiting the use of some forms of AI superficially appears to conflict with Gallego et al. (2022) who found that concerned survey respondents in Spain advocated for policies slowing down technological change over compensatory policies. However, they look at concern about technology rather than individual technology use, which, as we have seen, need not be correlated. Thus, at least on average, our intervention did not affect people's risk perceptions while shifting attitudes towards AI in a positive direction. Both of these findings contradict the risk-insurance framework but are consistent with the sociotropic model of AI and politics.

Figure 3 shows that the effect of AI use on social policy preferences is in line with the theoretical predictions of the sociotropic model as well (H2.3). Respondents in the control group are marginally significantly more likely to advocate for the government to provide for the unemployed and offer a job to everyone who wants one. We find a larger and highly statistically significant effect on support for government-provided re-training programs. This is notable as scholarship in the pre-AI era has fairly consistently found that individuals affected by technological change favor compensatory policies over activating measures that allow them to adapt (Weisstanner, 2023). They are also significantly more left-leaning than respondents in the control group, highlighting the dominance of economic or "first dimension" concerns in policy responses to AI.[8] These results paint a consistent picture in line with the sociotropic model. It appears that AI users

[8]Political ideology is commonly considered a fairly stable trait and is more often used as an independent variable in explaining

Figure 2: Effect of AI Use on Expectations and Attitudes



Note: The figure shows estimated treatment effects based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). N = 1,041 in all models. Full model output in Table D3.

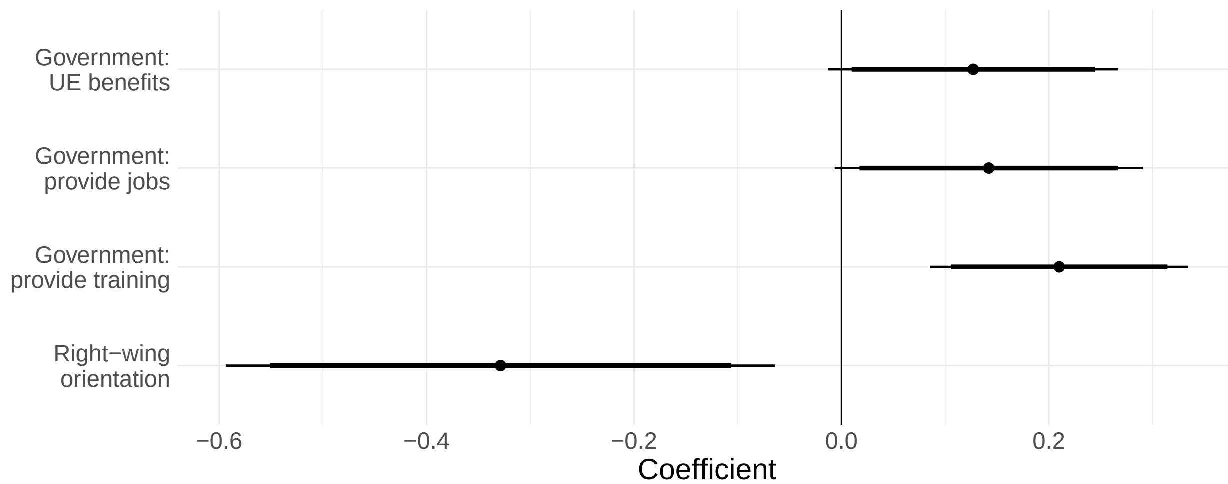
consider themselves to be insulated from potential negative employment effects but, in light of an alarmist discourse and after experiencing the capabilities of generative AI first-hand, nevertheless acknowledge a need for redistributive state action. Further research is necessary to more thoroughly test this proposition, for example by explicitly testing the mechanisms using causal mediation analysis (Tingley et al., 2014). Our initial findings suggest that there may be a window of opportunity during which the (expected) winners from the proliferation of generative AI support social policies that are presented as necessary to compensate the losers from AI.

The models shown so far include survey weights but no controls, as the participants were randomly assigned to the treatment and we have good balance between the groups in terms of socio-demographic characteristics (see Table B2). However, we show in Table G14 and Table G15 that the results are robust to including pre-treatment individual-level covariates (age, gender, education, and income). The results are furthermore robust to excluding respondents who failed an attention check question or respondents who did not comply with the instructions to use (or refrain from using) AI (results available upon request). Another concern is that by virtue of having a ChatGPT account, our respondents might be highly selected from a

redistributive or tax preferences (see, e.g., Fernández-Albertos and Kuo, 2018; Armingeon and Weisstanner, 2022). While it is unlikely that our treatment fundamentally altered people's ideological commitments, it may reflect an attempt to avoid cognitive dissonance. Since we asked the ideology question after the social policy questions, treated respondents may have felt compelled to describe themselves as more left-wing after expressing support for left-wing social policy positions. Causal mediation analysis supports this interpretation (Tingley et al., 2014).

technological avant-garde. Note, however, that our participants are still overwhelmingly novice users: only 30% report using ChatGPT at least once a week, and 23% say they never use it despite having an account, with the rest being infrequent users. If we exclude frequent users, we obtain qualitatively identical results, with some reductions in statistical significance due to the smaller sample, as we show in Table G16 and Table G17. Hence, we are confident in the robustness and generalizability of our main results.

Figure 3: Effect of AI use on social policy preferences



Note: The figure shows estimated treatment effects based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). N = 1,041 in all models. Full model output in Table D4.

Heterogeneous Effects of Exposure to AI

Digging deeper, we find evidence that these aggregate results mask important heterogeneity between groups. Inequalities by socio-demographic characteristics such as gender, education, age, or occupation are some of the most persistent features of modern labor markets and wider societies (Goldin, 2014; Blau and Kahn, 2017; Weisstanner and Armingeon, 2020; Brunello, Fort and Weber, 2019; Haslberger, 2021). Exposure to, familiarity with, and pre-existing attitudes towards novel technologies by and large vary along the same lines (OECD, 2018; Combet, 2024; Carvajal, Franco and Isaksson, 2024; Haslberger, Gingrich and Bhatia, 2023). It is therefore plausible that exposure to AI affects the attitudes and preferences of members of various groups differently. For reasons of space, we focus on gender and education differences in the main text, and provide results for age and occupations in Appendix F. The next set of figures explores these possibilities by interacting our treatment dummy with the respective demographic indicators. Note that the interaction coefficients can be interpreted as the difference in differences of the treatment effect on the respective groups.

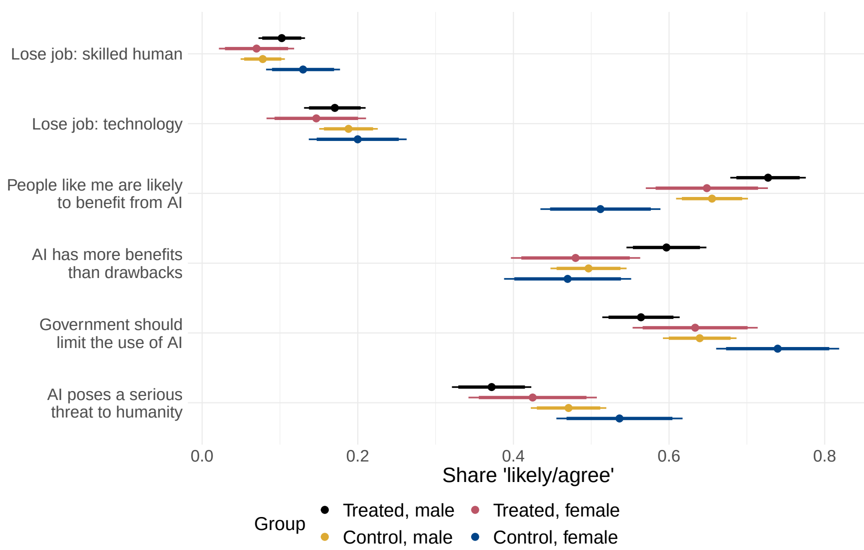
Gender Differences

Figure 4 shows some baseline differences between the sexes regarding automation fears and attitudes, but similar effects in terms of size and direction of the treatment. There is only one exception: while in the control group, women are slightly more likely to fear replacement by someone with better technology skills than men (13% vs 8%), this relationship is reversed in the treatment group (7% vs 10%). Thus, the null finding in Figure 2 is the result of two opposing effects cancelling each other out. By contrast, we find neither treatment nor gender differences in people's expectation that they will lose their job to technology directly. For both perceived risk questions, and regardless of treatment status, only a small number of respondents deem it (very) likely that their job prospects will deteriorate due to new technologies. The share of concerned respondents is also substantially lower than in any country included in the original OECD Risks that Matter survey (see, for example, Busemeyer et al. 2023) — but this survey may show inflated levels of concern due to priming effects. Nevertheless, this clearly does not align with the narrative that AI will eliminate vast numbers of jobs — or the soon-to-be-jobless are oblivious of what is in store for them.

In the four questions on expectations surrounding AI, we see that in the control group, women are significantly less likely than men to think that they will personally benefit from AI, and significantly more likely to think that the government should limit the use of some forms of AI. While the treatment main effects (that is, the effects on men) are statistically significant in all cases, we find no significant interaction effects. This indicates that the treatment affects the expectations of women and men similarly. However, it should be noted that some of the interactions have substantively meaningful effect sizes but narrowly fail to reach statistical significance. For example, the treatment increases men's likelihood of agreeing that products and services using AI have more benefits than drawbacks by 10 percentage points, while the increase for women is only about 1 percentage point. Yet, due to the larger standard errors in the smaller female group, the interaction is not significant even though clearly substantively meaningful. Thus, a larger and more balanced sample may well uncover more gender differences.

The results for social policy preferences and political orientation align with the findings for risks and expectations (see Figure 5). In the control group, men are less supportive of unemployment benefits than women (by about 0.3 points on a 5-point scale). However the treatment increases men's support by approximately 0.3 points and reduces that of women by the same amount, so that they effectively swap position. This is interesting in light of women's greater baseline scepticism towards AI, but is in line with the finding that women become less worried about their job prospects after using AI, while for men, if anything, the opposite is true Gingrich and Kuo (2022).[9] The (borderline significant) effect on support for job programs is similar in magnitude for both sexes and there are no baseline differences either. However, the aggregate treatment effect on support for government provision of re-training opportunities is entirely driven by men, whose support increases by 0.3 points, while women's views remain unaffected. Finally, while men are politically significantly to the right of women, both groups move to the left by the same amount after being exposed to our treatment. Taken together, these results suggest that men are the main drivers of the sociotropic effects we found in the pooled analysis. Whether this has to do with men's greater recent experience with technological displacement is an interesting area for further study, especially as recent estimates suggest that women may be more exposed to disruption from generative AI (Gmyrek, Berg and Bescond, 2023).

Figure 4: Gender differences in effect on expectations and attitudes



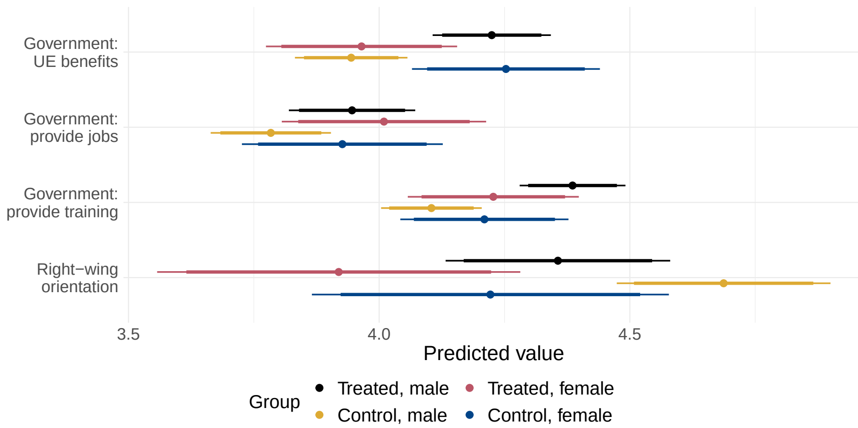
Note: The figure shows predicted values based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). N = 1,041 in all models. Full model output in Table D5.

Education Differences

We now consider heterogeneous effects by education in Figure 6. In the control group, respondents with a university degree are slightly more worried about losing their job to a tech-savvy human than those without (11% vs 6%), but there is a significant negative interaction, reversing the relationship in the treatment group (8% vs 14%). While the change is not statistically significant for degree holders, it is for those without university education. As above, neither treatment status nor education affect the likelihood of worrying

[9] This also implies that the borderline significant aggregate effect in Figure 3 is due to the over-representation of men the sample.

Figure 5: Gender differences in effect on social policy preferences

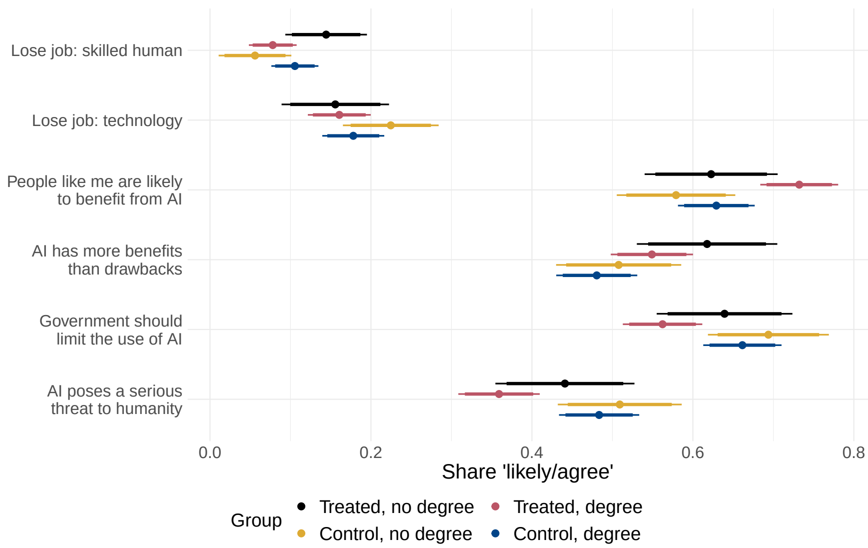


Note: The figure shows predicted values based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). N = 1,041 in all models. Full model output in Table D6.

about job loss to technology. For the expectation questions, interestingly, we do not find baseline differences by education. The treatment makes both educational groups more likely to see AI as beneficial and less likely to see it as a threat or advocate government limitations. The treatment effects appear to be stronger on respondents with a university degree, but the insignificant interactions do not allow us to reject the null hypothesis that the treatment affects respondents equally regardless of education. Thus, except for the fear of job loss to more skilled workers, we find neither level differences nor differential treatment effects between educational groups.

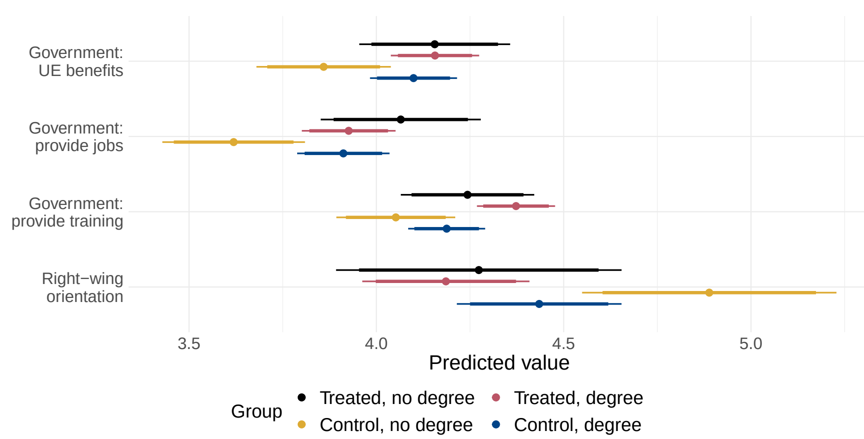
When it comes to social policy preferences, Figure 7 shows that initial preference gaps essentially disappear in the treated group. In the control group, respondents without a university degree tend to be *less* supportive of government provision of benefits, jobs, and training, and place themselves further to the right politically, than degree holders. All of these gaps disappear in the treatment group, where the views of the more and less educated are statistically indistinguishable. This narrowing is due to stronger treatment effects on the non-university-educated group, while the only significant treatment effect on the university-educated group is an increase in support for government-provided training opportunities. Thus, even though only one of the interaction coefficients (for government job provision) is statistically significant in our sample, Figure 7 shows a clear convergence after exposure to AI. The evidence for a convergence of policy views between university-educated and non-university-educated AI users should be treated as suggestive until successful replication with a larger sample. However, it indicates one way in which AI may ameliorate, rather than accentuate, existing polarization in society and thus offers a cautiously optimistic message. Given that gender and education are two of the dimensions along which many democracies have seen strong polarization in recent years, the relatively limited heterogeneous effects of AI exposure also corroborate research from the US which argues that AI is not (yet) heavily politicized. The findings for age and occupations reported in Appendix F point in the same direction. Yet, whether this state of affairs can last remains to be seen.

Figure 6: Education differences in effect on expectations and attitudes



Note: The figure shows predicted values based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). N = 1,027 in all models. Full model output in Table D7.

Figure 7: Education differences in effect on social policy preferences



Note: The figure shows predicted values based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). N = 1,027 in all models. Full model output in Table D8.

Evidence from Quantitative Text Analysis

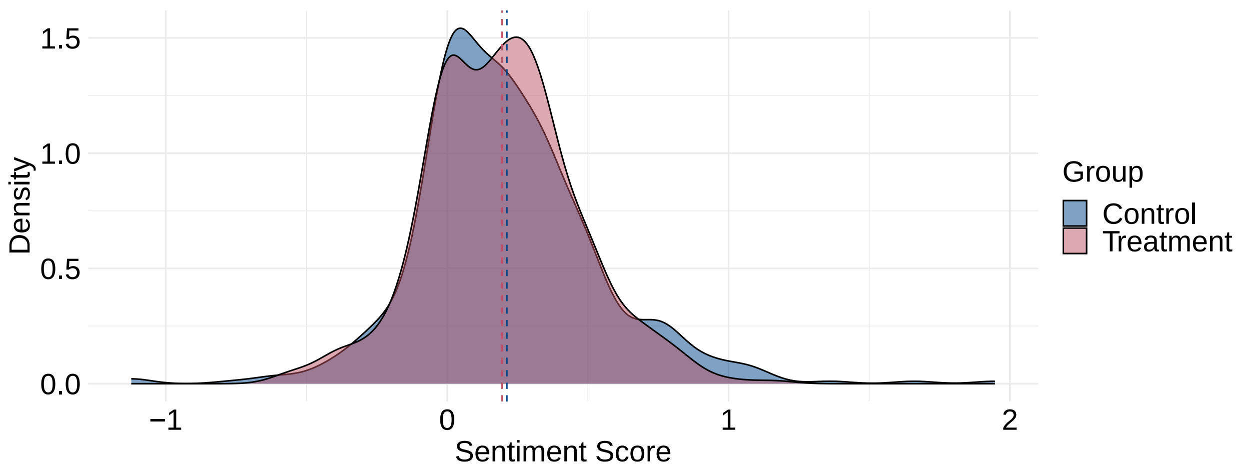
We asked people at the very end of the survey to "please briefly tell us in your own words what you think is most important when it comes to dealing with the effects of artificial intelligence on the labour market." People in both groups tended to take this question seriously and provided thoughtful and considered responses despite the length of the survey. However, some respondents used ChatGPT to answer this question. We use a triangulation approach using qualitative assessment, similarity scores, and AI detection software to identify and exclude these answers in the analyses presented here.[10] We identified one answer in the control group and 37 in the treatment group as likely AI-generated, leaving a corpus of 1,003 answers for our analyses (533 in the control group and 470 in the treatment group). Answers in the treatment group were on average 37 words long, compared to 33 words in the control group, a borderline statistically significant difference ($p < 0.1$).

We first perform a sentiment analysis to gauge how our respondents feel about artificial intelligence in the labor market. We use the *sentimentr* package to score each sentence according to its sentiment polarity and then aggregate over respondents. Overall sentiment of the answers is slightly positive, with an average value of 0.18. This is in line with the finding that a majority of respondents agree with the statement that AI has more benefits than drawbacks while few fear for their job, and suggests a cautious optimism regarding the implications of generative AI for the labor market. Sentiment is virtually indistinguishable in the control group and the treatment group, as shown in Figure 8. We also find no statistically significant differences based on any of the key socio-demographic markers (age, sex, education, occupation). Overall, while the sentiment analysis is no help in adjudicating between H1.2 and H2.2, it adds to the evidence that people in the UK are cautiously optimistic when it comes to dealing with the effects of artificial intelligence on the labour market.

We next attempt to identify the latent themes in our participants' responses. Our question encouraged participants to think independently about the challenges AI poses for the labor market and how they could be addressed. Other than directing their focus to the labor market, we did not guide the respondents in any way. They may of course have taken some cues from the earlier parts of the survey, but we were careful not to present AI technology in a one-sided way. This makes the question particularly suitable for structural topic modelling. We use the *stm* package in R and experimented with different specifications before eventually settling on a model with three topics and which includes the treatment indicator as a covariate. This specification avoids repetition while still distinguishing between substantively different sets

[10] See Appendix E for details on the selection process. Including likely ChatGPT-generated answers would inflate the differences between treatment and control group, as all but one of the AI-generated answers are in the treatment group and ChatGPT generally provides very positive answers. Interestingly, the demographic characteristics of the respondents who used ChatGPT to answer the open-ended question are very similar to the full sample.

Figure 8: Sentiment scores by treatment group



Note: Sentiment scores calculated using the *sentimentr* package.

Table 3: Exemplary responses for the different topics

Topic 1: Adaptation and intervention
There will be a massive need for retraining and income support while people's jobs are displaced by new technology and new jobs are created. I support government intervention to make sure this happens.
The effects of AI need to be balanced in a way where it acts as an aid for work without resulting in large-scale job losses or loss of income. Likewise an increase in the use of AI will create a load of new jobs and new positions in areas that may not have been created yet, depending on how far AI revolutionises society and the workplace.
Ensuring that people that lose their jobs to AI have a viable path back to employment and are helped whilst searching for a new job. Ideally though, people would be given the opportunity to gain new skills within their existing workplace in the event that their job is no longer needed due to AI, because this could be the best outcome for everyone involved.
Topic 2: Ethics and safety
As with any other technology, deploying AI should be done in a responsible manner, it can be a powerful tool or a powerful enabler of other technologies, however, history has shown us that powerful technologies can also be used to exploit and harm, a good example of that is the internet itself.
It is important that AI is constructed ethically, not just for the labour market but for all aspects of life. In its current form it is a useful productivity tool and should be used as such but it has capabilities that eventually we won't be able to understand and should be treated and developed very carefully
I work I an academic library. Safeguarding user data and ensuring transparent AI algorithms are essential to maintain trust and integrity. Additionally, there is a need for librarians, educators, and researchers to acquire AI-related skills and knowledge to harness the full potential of these technologies responsibly.
Topic 3: Limitations
The important thing with ai is that it should be used as a tool to aid productivity rather than as an outright replacement for skills. AI has not yet advanced to a stage that would be referred to as general artificial intelligence and are more advanced language models. They can be used to great effect as a tool but in their current state are not appropriate to be used as an straightforward replacement for human lanour
This is evolving and we are touching the surface and we don't know as yet how much of an impact this will be. Some jobs will always be human ie health care and hospitality some industries will benefit too
It is important to understand that LLM "artificial intelligence" tools that exist today are not "AI" in any real sense, but instead are regurgitating information that they have scavenged to provide outputs that *look like* what they think the user might like to have in response to a query. Their usefulness is therefore extremely limited, since these "AIs" tend to structure responses to fit their "expectations" of a good response, rather than what actually is a good response or is even factually correct or congruent with the real world.

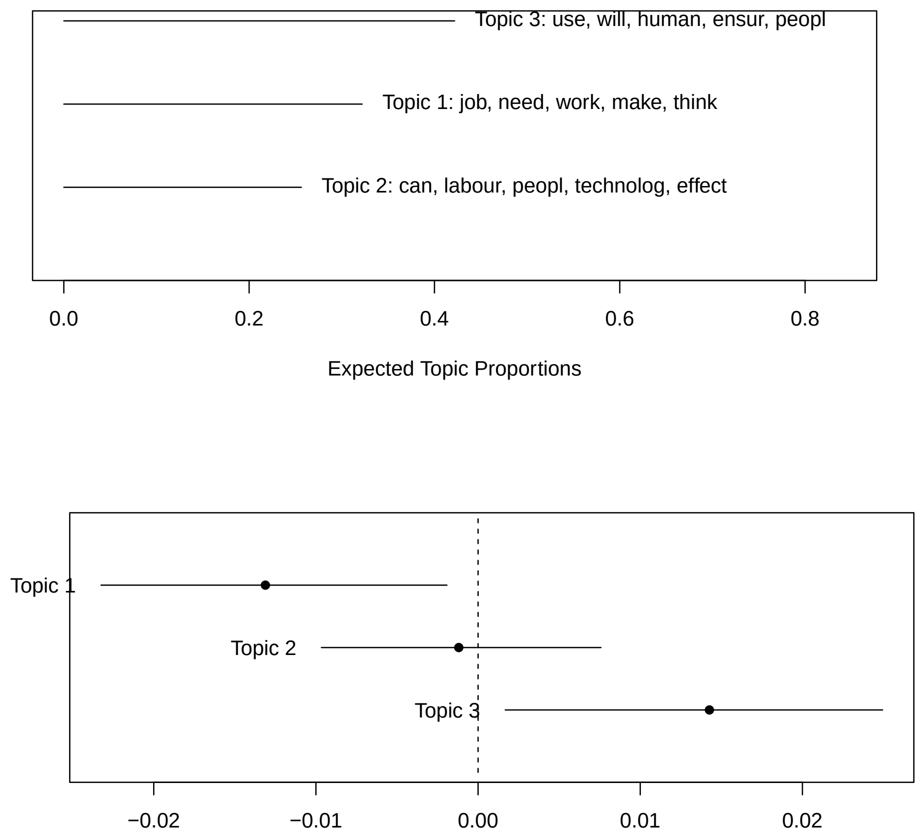
of ideas. The relatively small number of topics reflects the brevity and similarity of many of the open-ended responses.[11] The top panel of Figure 9 shows the three topics and their expected proportions in the answers. Topic 1, with an average topic loading of 0.33, revolves around the need for intervention to deal with job losses, for example through retraining policies and income support. Topic 2 highlights safety and ethics concerns alongside the benefits of AI. This topic has the lowest prevalence with an average topic loading of 0.25. The most prominent topic, with an average loading of 0.42, is topic 3, which stresses the technical limitations of current AI technology and expresses scepticism at its capability to replace humans on a large scale. Table 3 shows 3 exemplary responses for each of the topics. Note that while the selected responses load highly on the respective topic, they do not exclusively talk about one issue (the loadings for each answer sum up to 1).

The bottom panel of Figure 9 shows how the treatment affects the salience of the different topics. We find that respondents in the treatment group are significantly less likely to engage with topic 1 in their answers compared to the control group. The effect amounts to a reduction in prevalence of 1.3 percentage points, or approximately 4 percent. Answers that load highly on this topic tend to talk about the need to prevent or deal with job losses, especially through policies that promote skill acquisition or income support. The treatment does not affect the salience of the safety and ethics concerns central to topic 2. However, the answers of treated respondents load significantly higher on topic 3, which focuses on the current limitations of AI. This effect is similar in size to that on topic 1, at approximately 1.4 percentage points. This reflects the fact that topic 3 is to some extent the mirror image of topic 1. It is at first surprising that the treatment leads to lower loadings on topic 1, which makes explicit mention of government intervention through social policy. However, this might be explained by the greater emphasis on topic 3: treated respondents might be less worried that the interventions central to topic 1 will be needed on a large scale. In a model including key socio-demographic covariates (age, sex, education, occupation), we obtain essentially the same topics and estimates of the treatment effects. Table E9 shows the results. Older respondents' answers tend to load more heavily on topic 2 and less strongly on topic 3, while the answers of those in professional and managerial occupations load less heavily on topic 1 and more strongly on topic 3. In additional analyses, we find no effect of political orientation on the prevalence of any of the topics, but since political orientation is measured post-treatment, this should not be overinterpreted (Montgomery, Nyhan and Torres, 2018). Nevertheless, it is consistent with existing research from the US which argues that AI is not politicized yet (Margalit and Raviv, 2023; Schiff, Schiff and Jacobson, 2024).

Overall, while the text analysis does not allow us to directly test our hypotheses, it provides complementary insights into people's reasoning about the possible consequences of generative AI. The sentiments and

[11] Additional details from the STM analysis are reported in Appendix E.

Figure 9: Structural topic model



Note: Structural topic model estimated with the *stm* package in R. The top panel shows the topic prevalence. The bottom panel shows the effect of the treatment on topic prevalence.

underlying themes we discovered by and large reinforce our conclusions from the inferential analysis. We found, against our pre-registered expectations, that respondents who were encouraged to use AI in our tasks were no more worried about job loss than untreated participants, and generally expressed more positive views about AI. While the sentiment analysis did not echo this finding, as we found no treatment effect, the STM analysis is consistent with it. Assignment to our treatment is associated with lower loads on topic 1, which is most pessimistic with regard to labor market effects, and with higher loads on topic 3, which highlights the limitations of AI and the continued need for human labor. Due to this pessimism, treated respondents are less likely to engage with topic 1 despite supporting various compensatory and activating social policies which the topic also mentions. The text analysis therefore reinforces the picture that first-hand experience with AI use tends to dispel fears and leads to a more positive outlook overall.

Discussion and Conclusion

In this article, we presented experimental evidence on the implications of generative AI use for political preferences and expectations. Based on a near-representative sample of working-age adults from the UK, we document that respondents who were randomly assigned to use generative AI in three short text-based tasks became more positively disposed towards the technology while also developing more progressive social policy preferences. Analyzing text data from an open-ended question, we find that amidst overall positive sentiment, treated respondents are less likely to mention job losses and more likely to allude to the limitations of current generative AI technology. We argue that this supports a model whereby the primary effect of AI technology is on sociotropic preferences, rather than the risk-insurance framework that emphasizes individual material mechanisms.

In more detail, direct exposure to AI through our treatment did not increase anxiety about people's own job prospects — they are no more likely to worry about losing their job to someone with better technology skills or to technology outright. We also find a significant increase in the share of people who expect to benefit from AI and a reduction in calls for government limitations on the use of AI and in the share of people who consider AI a threat to humanity. Thus, people overall became more positively disposed towards AI, which went against our initial expectations. At the same time, treated respondents showed increased support for progressive social policies such as unemployment benefits, a job guarantee, and continued training schemes, and described themselves as more left-wing. In light of the vast social inequalities that characterize modern societies, we also shed light on the interaction between AI use and gender and education. We show that gender and (to a lesser extent) education gaps in attitudes towards AI and social policy exist, some of which are affected by the treatment. For example, women became less worried about their job prospects and less supportive of unemployment benefits, while the opposite is true for men. Changes in policy preferences and political orientation were concentrated in less educated respondents, closing the gaps in support for progressive policies. Overall, the support for a model of sociotropic preferences over the self-interest-based risk-insurance framework may reflect uncertainty about who the relative winners and losers of widespread adoption of AI will be — currently, forming self-interested views is only possible to a limited extent.

We furthermore used an open-ended question to elicit people's first-order concerns with regard to the effects of AI on the labor market (Ferrario and Stantcheva, 2022). We first show that overall sentiment is slightly positive regardless of treatment status. We then employ structural topic models to uncover the latent themes respondents engage with. We identify three main topics: adaptation and intervention, ethics and safety, and limitations. Treated respondents tend to put less emphasis on adaptation and intervention and more on the limitations of current AI technology. We can also corroborate existing evidence from the US that the question of how to best deal with AI is not subject to partisan cueing yet, perhaps offering a rare opportunity for bipartisan consensus on regulation for the common good (Margalit and Raviv, 2023). Overall, the nuanced responses to the open-ended question provided valuable insights into people's priorities and concerns and can be helpful for refining future questionnaires with new, relevant questions.

While this article provides important and novel evidence, it has a number of limitations. First, as we already pointed out, the sample is not fully representative of the UK working age population. This is due to methodological limitations at the time of the survey, in particular the requirement, due to privacy concerns, to sample people who already had a ChatGPT account. In future work, we intend to replicate our analyses with a larger and more representative sample. Closely related, this study focused on the UK, with a particular set of liberal labor market institutions. It is plausible that people's response to contact with new technology might differ in countries with a more universalist approach to social policy. Comparative work will therefore be necessary to establish the external validity of our findings. Furthermore, due to time limitations, our survey included a rather limited set of policy questions. While these are frequently used in welfare state scholarship (Vlandas, 2020), future work will benefit from a wider range of questions — in particular, questions that highlight trade-offs (Busemeyer and Tober, 2022; Magistro et al., 2024).

Another possible objection concerns the nature of our experimental treatment itself. Our idea in designing the treatment was to go beyond the frequently used information treatments and directly expose people to generative AI in an RCT-like setting to allow them to see for themselves the risks the new technology entails. Of course, with this setup, we have less control over what people take away from the intervention — in this case, it appears that they above all found AI useful, but did not make the leap to seeing their jobs at risk. While this might be construed as a shortcoming of the study design, we contend that it is itself an interesting insight which suggests avenues for further research. For example, it could be insightful to combine an experience treatment like ours with a traditional information treatment to study the impact of potentially conflicting information on attitudes towards AI. In short, this study entered uncharted territory which necessitated some design choices that can be criticized. Nevertheless, we believe that it provides important insights and that it can serve as a useful starting point for further research into the political implications of generative AI.

Scholars interested in labor markets, redistribution and the welfare state, or voting behavior and extremism will all have to grapple with the implications of AI. The findings detailed in this article provide some guiding insights to inform the emerging research agenda on the politics of AI. First, it will be imperative for political scientists to closely monitor the literature in adjacent disciplines — particularly economics, management, and computer science — to remain informed about the current capabilities and uses of AI. Second, as long as uncertainty about the economic winners and losers of labor market changes is pervasive, theories that rely on material interest to predict political consequences are likely to provide limited insights. For the same reason, however, one should not conclude based on our study alone that self-interest as in the risk-insurance framework will play no role in the politics of AI. It may well regain its explanatory power once the relative winners and losers are clear. Finally, there are some methodological lessons as well. Especially in an emerging field, we believe that text-as-data can be invaluable for making sense of counterintuitive results and theorybuilding. Furthermore, our study highlights the pros and cons of different experimental treatments. While undoubtedly stronger than priming or information treatments, with a treatment like ours the experimenter also has less control over what respondents take away from the treatment. This can create ambiguity in the interpretation of the results, which must be weighed against the greater realism of the treatment. We conclude by expressing our hope that this article will contribute to the emerging research agenda on the political implications of generative AI and that it will inspire further research on the topic.

References

Adler, David and Ben Ansell. 2020. "Housing and populism." West European Politics 43(2):344–365. Publisher: Routledge.

Armingeon, Klaus and David Weisstanner. 2022. "Objective Conditions Count, Political Beliefs Decide: The Conditional Effects of Self-Interest and Ideology on Redistribution Preferences." Political Studies 70(4):887–900.

Barabas, Jason and Jennifer Jerit. 2010. "Are Survey Experiments Externally Valid?" American Political Science Review 104(2):226–242.

Blau, Francine D. and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." Journal of Economic Literature 55(3):789–865.

Brunello, Giorgio, Margherita Fort and Guglielmo Weber. 2019. "Changes in Compulsory Schooling, Education and the Distribution of Wages in Europe." The Economic Journal 119(536):516–539.

Brynjolfsson, Erik, Danielle Li and Lindsey R. Raymond. 2023. "Generative AI at Work." NBER Working Paper Series (31161).

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):751–770.

Busemeyer, Marius R., Mia Gandenberger, Carlo Knotz and Tobias Tober. 2023. "Preferred policy responses to technological change: Survey evidence from OECD countries." Socio-Economic Review 21(1):593–615.

Busemeyer, Marius R. and Tobias Tober. 2022. "Dealing with Technological Change: Social Policy Preferences and Institutional Context." Comparative Political Studies p. 001041402211393.

Carvajal, Daniel, Catalina Franco and Siri Isaksson. 2024. "Will Artificial Intelligence Get in the Way of Achieving Gender Equality?" SSRN Electronic Journal .

Cavaillé, Charlotte. 2023. Fair Enough? Support for Redistribution in the Age of Inequality. Cambridge: Cambridge University Press.

Ciani, Emanuele, Louis Freget and Thomas Manfredi. 2021. Learning about inequality and demand for redistribution: A meta-analysis of in-survey informational experiments. OECD Papers on Well-being and Inequalities 02. Series: OECD Papers on Well-being and Inequalities Volume: 02.

Combet, Benita. 2024. "Women's aversion to majors that (seemingly) require systemizing skills causes gendered field of study choice." European Sociological Review 40(2):242–257.

Compton, Mallory E. and Christine S. Lipsmeyer. 2019. "Everybody Hurts Sometimes: How Personal and Collective Insecurities Shape Policy Preferences." The Journal of Politics 81(2):539–551.

Converse, Philip E. 2006. "The nature of belief systems in mass publics (1964)." Critical Review 18(1-3):1–74.

Dell'Acqua, Fabrizio, Edward McFowland Iii, Ethan Mollick, Hila Lifshitz-Assaf, Katherine C Kellogg, Saran Rajendran, Lisa Krayer, François Candelon and Karim R Lakhani. 2023. "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality." SSRN Electronic Journal .

Dermont, Clau and David Weisstanner. 2020. "Automation and the future of the welfare state: basic income as a response to technological change?" Political Research Exchange 2(1):1757387.

Felten, Edward W., Manav Raj and Robert Seamans. 2023a. "How will Language Modelers like ChatGPT Affect Occupations and Industries?" SSRN Electronic Journal .

Felten, Edward W., Manav Raj and Robert Seamans. 2023b. "Occupational Heterogeneity in Exposure to Generative AI." SSRN Electronic Journal .

Fernández-Albertos, José and Alexander Kuo. 2018. "Income Perception, Information, and Progressive Taxation: Evidence from a Survey Experiment." Political Science Research and Methods 6(1):83–110.

Ferrario, Beatrice and Stefanie Stantcheva. 2022. "Eliciting People's First-Order Concerns: Text Analysis of Open-Ended Survey Questions." AEA Papers and Proceedings 112:163–169.

Frey, Carl Benedikt and Michael Osborne. 2023. "Generative AI and the Future of Work: A Reappraisal." Brown Journal of World Affairs XXX(1).

Funk, Carolyn L. 2000. "The Dual Influence of Self-Interest and Societal Interest in Public Opinion." Political Research Quarterly 53(1):37–62.

Future of Life Institute. 2023. "Policymaking In The Pause: What can policymakers do now to combat risks from advanced AI systems?".

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 and Thomas Kurer. 2022. "Automation, Digitalization, and Artificial Intelligence in the Workplace: Implications for Political Behavior." Annual Review of Political Science 25(1):463–484.

Gallego, Aina, Thomas Kurer and Nikolas Schöll. 2022. "Neither Left Behind nor Superstar: Ordinary Winners of Digitalization at the Ballot Box." The Journal of Politics 84(1):418–436.

Gingrich, Jane and Alexander Kuo. 2022. Gender, Technological Risk, and Political Preferences. In Digitalization and the Welfare State, ed. Marius R. Busemeyer, Achim Kemmerling, Paul Marx and Kees van Kersbergen. Oxford: Oxford University Press.

Gingrich, Jane and Ben Ansell. 2012. "Preferences in Context: Micro Preferences, Macro Contexts, and the Demand for Social Policy." Comparative Political Studies 45(12):1624–1654.

Gmyrek, Pawel, Janine Berg and David Bescond. 2023. "Generative AI and Jobs: A global analysis of potential effects on job quantity and quality." ILO Working Paper 96.

Goldin, Claudia. 2014. "A grand gender convergence: Its last chapter." American Economic Review 104(4):1091–1119.

Hakim, Catherine. 1982. "The Social Consequences of High Unemployment." Journal of Social Policy 11(4):433–467.

Haslberger, Matthias. 2021. "Routine-biased technological change does not always lead to polarisation: Evidence from 10 OECD countries, 1995–2013." Research in Social Stratification and Mobility 74(4):100623.

Haslberger, Matthias, Jane Gingrich and Jasmine Bhatia. 2023. "No Great Equalizer: Experimental Evidence on AI in the UK Labor Market." SSRN Electronic Journal .

Haslberger, Matthias, Patrick Emmenegger and Niccolo Durazzi. 2023. The Missing Link: Technological Change, Dual VET, and Social Policy Preferences. In EPSA 2023. Glasgow: .

Hui, Xiang, Oren Reshef and Luofeng Zhou. 2023. "The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market." SSRN Electronic Journal .

Im, Zhen Jie. 2021. "Automation risk and support for welfare policies: How does the threat of unemployment affect demanding active labour market policy support?" Journal of International and Comparative Social Policy 37(1):76–91.

Kam, Cindy D. and Yunju Nam. 2008. "Reaching Out or Pulling Back: Macroeconomic Conditions and Public Support for Social Welfare Spending." Political Behavior 30(2):223–258.

Kitschelt, Herbert and Philipp Rehm. 2014. "Occupations as a Site of Political Preference Formation." Comparative Political Studies 47(12):1670–1706.

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 Silja Häusermann. 2022. Automation Risk, Social Policy Preferences, and Political Participation. In Digitalization and the Welfare State, ed. Marius R. Busemeyer, Achim Kemmerling, Paul Marx and Kees van Kersbergen. Oxford: Oxford University Press.

Law School Admission Council. 2007. "The Official LSAT Preptest.". Form 8LSN75.

Magistro, Beatrice, Peter Loewen, Bart Bonikowski, Sophie Borwein and Blake Lee-Whiting. 2024. "Attitudes toward automation and the demand for policies addressing job loss: the effects of information about trade-offs." Political Science Research and Methods pp. 1–16.

Margalit, Yotam and Shir Raviv. 2023. "The Politics of Using AI in Public Policy: Experimental Evidence." SSRN Electronic Journal .

Montgomery, Jacob M., Brendan Nyhan and Michelle Torres. 2018. "How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It." American Journal of Political Science 62(3):760–775.

Noy, Shakked and Whitney Zhang. 2023. "Experimental evidence on the productivity effects of generative artificial intelligence." Science (381):187–192.

OECD. 2018. Bridging the Digital Gender Divide - Include, Upskill, Innovate. Paris: OECD.

Raviv, Shir. 2023. "When Do Citizens Resist The Use of Algorithmic Decision-making in Public Policy? Theory and Evidence." SSRN Electronic Journal .

Rehm, Philipp. 2011. "Social Policy by Popular Demand." World Politics 63(2):271–299.

Rueda, David and Daniel Stegmueller. 2016. "The Externalities of Inequality: Fear of Crime and Preferences for Redistribution in Western Europe." American Journal of Political Science 60(2):472–489.

Sacchi, Stefano, Dario Guarascio and Silvia Vannutelli. 2020. Risk of technological unemployment and support for redistributive policies. In The European Social Model under Pressure, ed. Romana Careja, Patrick Emmenegger and Nathalie Giger. Springer VS Wiesbaden pp. 277–295.

Schiff, Daniel S, Kaylyn Jackson Schiff and Robin Jacobson. 2024. The Politics of AI: Will Bipartisanship Last or Is Polarization Inevitable? Chicago: .

Støstad, Morten Nyborg and Frank Cowell. 2024. "Inequality as an externality: Consequences for tax design." Journal of Public Economics 235:105139.

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

Tingley, Dustin, Teppei Yamamoto, Kentaro Hirose, Luke Keele and Kosuke Imai. 2014. "**mediation** : *R* Package for Causal Mediation Analysis." Journal of Statistical Software 59(5).

Vlandas, Tim. 2020. "The Political Consequences of Labor Market Dualization: Labor Market Status, Occupational Unemployment and Policy Preferences." Political Science Research and Methods 8(2):362–368.

Weisstanner, David. 2023. Technological Change and Labour Market Policy Preferences. In Handbook of Labour Market Policy in Advanced Democracies, ed. Daniel Clegg and Niccolo Durazzi. Cheltenham: Edward Elgar Publishing.

Weisstanner, David and Klaus Armingeon. 2020. "How redistributive policies reduce market inequality: education premiums in 22 OECD countries." Socio-Economic Review 18(3):839–856.

Wu, Nicole. 2022. "Misattributed blame? Attitudes toward globalization in the age of automation." Political Science Research and Methods 10(3):470–487.

Wu, Nicole. 2023. ""Restrict foreigners, not robots": Partisan responses to automation threat." Economics & Politics 35(2):505–528.

Zhang, Baobao. 2022. No Rage Against the Machines: Threat of Automation Does Not Change Policy Preferences. In Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society. Oxford United Kingdom: ACM pp. 856–866.

Appendix: For Online Publication

Contents

A Ethics

The research is conducted so it maintains the highest research ethics, including compliance with the *Principles and Guidance for Human Subjects Research* adopted by the APSA Council in 2020. The project was conducted with approval from the University of Oxford Department of Social Policy and Intervention's Ethics Committee (Reference number: SPI_C1A_23_004).

All participants had to fill in a participation form at the beginning of the survey. The form informed participants that they were taking part in a research study and that their answers would be used as part of an academic research project examining the implications of generative artificial intelligence. Participants were informed that they may be asked to use the free third-party software ChatGPT, but that they would not be required to do so. Moreover, the form, amongst other things, told the participants that their participation was voluntary, that they could end their participation at any time, that their answers would be anonymized and stored in a secure way, and that they have the right to withdraw their answers at a later point if they so desire. The form also contained the reference number of the ethics approval, enabling participants to raise any concerns.

We did not engage in deception or intervene in political processes, nor was our study and the material used in it likely to trigger strong emotions, intense psychological stress, or traumatic experiences beyond everyday experiences. Furthermore, our study did not include physical interventions and did not expose participants to exceptional risks. The confidentiality of participants in our study is protected throughout all stages of the research process. The data is stored securely and any data used for analyses and provided for replication purposes are de-identified and anonymized.

Our research respects participants' autonomy and does not involve vulnerable individuals or groups, such as minors, patients, or persons unable to give informed consent. All study participants are volunteers recruited and compensated by the survey company YouGov. YouGov do not pay them directly but give them points, which can be redeemed for cash or prizes. In light of the length of the survey and the higher cognitive load than normal surveys, we offered compensation at twice the normal rate for a survey of this length, as well as performance-based incentives. Participants therefore received a fair compensation for their participation.

B Sampling

To ensure the feasibility of the study, YouGov asked 34,211 people between 14$^{\text{th}}$ April and 4$^{\text{th}}$ May 2023 whether they have a ChatGPT account. The study sample is recruited from the 5,350 respondents who stated that they have an account (free or paid). Below we list the feasibility question and the distribution of answers.

- ChatGPT is an AI-based computer program that can generate human-like text. Before taking this survey, had you ever used ChatGPT?
  - Yes, and I have a paid account (n=865)
  - Yes, and I have a free account (n=4485)
  - Yes, but I used someone else's account (n=774)
  - No, I have never used ChatGPT (n=28087)

While we acknowledge that this introduces some selection issues, we employ quotas for age, gender, income, and region to ensure a broadly representative sample. Given the user pool at the time, we did not manage to obtain a fully representative sample. As Table B1 shows, the sample is more male and more educated than our target population, the UK working age population. However, Table B2 shows good balance between the treatment and control groups.

Table B1: Sample characteristics

Characteristic	**Count/Mean (%/SD) Treatment**	**Count/Mean (%/SD) Control**
Sex: Female	165 (32.7%)	170 (31.7%)
Sex: Male	339 (67.3%)	367 (68.3%)
Uni: Yes	393 (79.1%)	421 (79.4%)
Uni: No	104 (20.9%)	109 (20.6%)
Professional: Yes	346 (68.7%)	355 (66.1%)
Professional: No	158 (31.4%)	182 (33.9%)
Age	40.5 (11.2)	39.8 (11.3)
HH Income	10.9 (3.2)	10.6 (3.4)
Skills Average	1.9 (0.5)	2.0 (0.6)

Table B2: Balance Tests

Variable name	Type	Std. Mean Dif.
Age	C	0.0616
Sex: Male	B	-0.0108
Degree	B	-0.0036
Degree: NA	B	0.0009
Professional occupation	B	0.0254
HH income	C	0.1017*
HH income: NA	B	-0.0116
Skills average	C	-0.0460
Skills average: NA	B	0.0010
Education: GCSEs / O-Levels or none	B	0.0046
Education: A-Levels or equivalent	B	-0.0010
Education: Undergraduate	B	-0.0216
Education: Postgraduate	B	0.0180

Note: B = binary variable; C = continuous variable. Sample sizes: Control: 537; Treatment: 504.

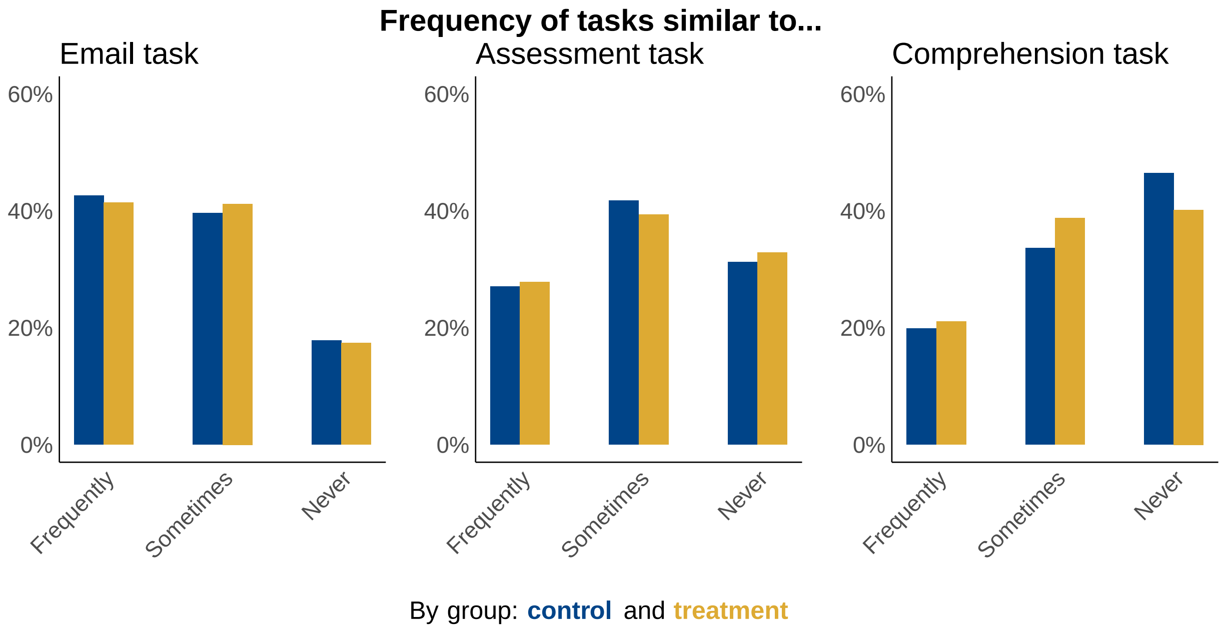
We pre-tested a version of the survey which included a control group in which we did not prime participants about AI at all. Since answer quality in this group was generally lower and a substantial share of respondents used ChatGPT, we did not include this control group in the final version of the survey.

C Tasks

In the first task, participants were presented with an email addressing a hypothetical workplace dispute and were asked to make any improvements they deemed appropriate. The email included deliberate grammatical and spelling errors and was written in a harsh, unprofessional tone. The evaluation criteria consider whether the errors have been corrected and whether the tone of the email has been made more constructive. In the second task, respondents evaluated the persuasiveness of two texts of about 500 words, each presenting opposing views on the merits of a universal basic income. While there was no correct answer to this task, the answers were graded based on whether they provided a well-reasoned and comprehensive assessment. In the final task we asked participants to answer three short questions about a complex text. The questions are adapted from a publicly available LSAT practice test (Law School Admission Council, 2007). We refer to the three tasks as the email task, assessment task, and comprehension task, respectively. The three tasks increase in complexity and are of a sufficiently general nature so that not only knowledge workers are likely to be familiar with them. Nevertheless, with increasing complexity, a growing share of respondents reports that they never perform similar tasks, as Figure C1 shows. However, there are no systematic differences between treatment groups. While not covering the entire range of use cases for AI, our design allows us to investigate the impact of AI use for tasks of varying complexity in a near-representative sample of the working age population.

D Full Models

Figure C1: Familiarity of respondents with the tasks



Note: The figure shows whether respondents perform the task frequently (once a week or more often), sometimes, or never.

Table D3: Full output to Figure 2

	Dependent variable:					
	Lose job 1	Lose job 2	Benefit 1	Benefit 2	Gov. limit	AI threat
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.002	−0.027	0.088***	0.075**	−0.083***	−0.102***
	(0.018)	(0.024)	(0.029)	(0.031)	(0.030)	(0.031)
Constant	0.092***	0.191***	0.617***	0.489***	0.666***	0.488***
	(0.012)	(0.016)	(0.020)	(0.021)	(0.021)	(0.021)
Observations	1,041	1,041	1,041	1,041	1,041	1,041
R^2	0.00001	0.001	0.009	0.006	0.007	0.010
F Stat. (df = 1; 1039)	0.010	1.322	9.037***	5.866**	7.622***	10.967***

Note: The models model agreement with the following statements: **Replace 1**: “I will lose my job because I am not good enough with new technologies or because I will be replaced by someone with better technological skills.” **Replace 2**: “My job will be replaced by an artificial intelligence, algorithm, computer software, or robot.” **Benefit 1**: “People like me are likely to benefit from AI.” **Benefit 2**: “Products and services using AI have more benefits than drawbacks.” **Gov. limit**: “The government should limit the use of some forms of AI.” **AI threat**: “AI poses a serious threat to humanity.” All models include weights. *p<0.1; **p<0.05; ***p<0.01.

Table D4: Full output to Figure 3

	Dependent variable:			
	Left-right	Training	Job guarantee	UE benefits
	(1)	(2)	(3)	(4)
Treatment	−0.329**	0.210***	0.142*	0.127*
	(0.135)	(0.063)	(0.076)	(0.071)
Constant	4.564***	4.132***	3.822***	4.026***
	(0.094)	(0.044)	(0.052)	(0.049)
Observations	1,041	1,041	1,041	1,041
R^2	0.006	0.010	0.003	0.003
F Statistic (df = 1; 1039)	5.928**	10.948***	3.517*	3.183*

Note: The models refer to the following questions: **Left-right**: Self placement on 0 - 10 left-right scale of political orientation. **Training**: "Should it be the government's responsibility to ensure training opportunities for everyone who wants to upgrade their skills?" **Job guarantee**: "Should it be the government's responsibility to ensure a job for everyone who wants one?" **UE benefits**: "Should it be the government's responsibility to ensure a reasonable standard of living for the unemployed?" All models include weights. *p<0.1; **p<0.05; ***p<0.01.

Table D5: Full output to Figure 4

	Dependent variable:					
	Lose job 1	Lose job 2	Benefit 1	Benefit 2	Gov. limit	AI threat
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.024	−0.018	0.072**	0.100***	−0.075**	−0.099***
	(0.021)	(0.028)	(0.034)	(0.036)	(0.035)	(0.036)
Female	0.052*	0.012	−0.143***	−0.027	0.100**	0.065
	(0.028)	(0.037)	(0.046)	(0.048)	(0.047)	(0.048)
Treatment x Female	−0.084**	−0.036	0.065	−0.090	−0.031	−0.013
	(0.040)	(0.054)	(0.066)	(0.069)	(0.067)	(0.069)
Constant	0.078***	0.188***	0.655***	0.497***	0.639***	0.471***
	(0.014)	(0.019)	(0.024)	(0.025)	(0.024)	(0.025)
Observations	1,041	1,041	1,041	1,041	1,041	1,041
R^2	0.004	0.002	0.021	0.011	0.014	0.013
F Stat. (df = 3; 1037)	1.543	0.602	7.242***	3.900***	4.766***	4.654***

Note: The models model agreement with the following statements: **Replace 1**: "I will lose my job because I am not good enough with new technologies or because I will be replaced by someone with better technological skills." **Replace 2**: "My job will be replaced by an artificial intelligence, algorithm, computer software, or robot." **Benefit 1**: "People like me are likely to benefit from AI." **Benefit 2**: "Products and services using AI have more benefits than drawbacks." **Gov. limit**: "The government should limit the use of some forms of AI." **AI threat**: "AI poses a serious threat to humanity." All models include weights. *p<0.1; **p<0.05; ***p<0.01.

Table D6: Full output to Figure 5

	Dependent variable:			
	Left-right	Training	Job guarantee	UE benefits
	(1)	(2)	(3)	(4)
Treatment	0.280***	0.162*	0.281***	−0.331**
	(0.083)	(0.089)	(0.074)	(0.158)
Female	0.309***	0.143	0.106	−0.465**
	(0.111)	(0.119)	(0.100)	(0.211)
Treatment x Female	−0.568***	−0.079	−0.264*	0.028
	(0.160)	(0.171)	(0.143)	(0.303)
Constant	3.944***	3.784***	4.104***	4.687***
	(0.057)	(0.061)	(0.051)	(0.109)
Observations	1,041	1,041	1,041	1,041
R^2	0.015	0.005	0.014	0.014
Adjusted R^2	0.012	0.002	0.011	0.011
Residual Std. Error (df = 1037)	1.142	1.221	1.022	2.169
F Statistic (df = 3; 1037)	5.358***	1.742	4.827***	4.956***

Note: The models refer to the following questions: **Left-right**: Self placement on 0 - 10 left-right scale of political orientation. **Training**: "Should it be the government's responsibility to ensure training opportunities for everyone who wants to upgrade their skills?" **Job guarantee**: "Should it be the government's responsibility to ensure a job for everyone who wants one?" **UE benefits**: "Should it be the government's responsibility to ensure a reasonable standard of living for the unemployed?" All models include weights. *p<0.1; **p<0.05; ***p<0.01.

Table D7: Full output to Figure 6

	Dependent variable:					
	Lose job 1	Lose job 2	Benefit 1	Benefit 2	Gov. limit	AI threat
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.088**	−0.069	0.044	0.110*	−0.054	−0.068
	(0.035)	(0.046)	(0.056)	(0.060)	(0.058)	(0.059)
Degree	0.050*	−0.047	0.050	−0.027	−0.032	−0.026
	(0.027)	(0.036)	(0.045)	(0.047)	(0.046)	(0.047)
Treatment x Degree	−0.116***	0.052	0.059	−0.041	−0.045	−0.056
	(0.041)	(0.053)	(0.066)	(0.070)	(0.068)	(0.069)
Constant	0.056**	0.225***	0.579***	0.508***	0.694***	0.509***
	(0.023)	(0.030)	(0.038)	(0.040)	(0.038)	(0.039)
Observations	1,027	1,027	1,027	1,027	1,027	1,027
R^2	0.008	0.003	0.015	0.008	0.011	0.015
F Stat. (df = 3; 1023)	2.741**	1.178	5.191***	2.793**	3.890***	5.287***

Note: The models model agreement with the following statements: **Replace 1**: "I will lose my job because I am not good enough with new technologies or because I will be replaced by someone with better technological skills." **Replace 2**: "My job will be replaced by an artificial intelligence, algorithm, computer software, or robot." **Benefit 1**: "People like me are likely to benefit from AI." **Benefit 2**: "Products and services using AI have more benefits than drawbacks." **Gov. limit**: "The government should limit the use of some forms of AI." **AI threat**: "AI poses a serious threat to humanity." All models include weights. *p<0.1; **p<0.05; ***p<0.01.

Table D8: Full output to Figure 7

	Dependent variable:			
	Left-right	Training	Job guarantee	UE benefits
	(1)	(2)	(3)	(4)
Treatment	0.296**	0.446***	0.191	−0.615**
	(0.137)	(0.146)	(0.122)	(0.260)
Degree	0.240**	0.293**	0.136	−0.454**
	(0.109)	(0.116)	(0.096)	(0.206)
Treatment x Degree	−0.239	−0.432**	−0.006	0.366
	(0.161)	(0.171)	(0.143)	(0.305)
Constant	3.859***	3.619***	4.052***	4.889***
	(0.091)	(0.097)	(0.081)	(0.173)
Observations	1,027	1,027	1,027	1,027
R^2	0.008	0.010	0.012	0.012
Adjusted R^2	0.005	0.008	0.009	0.009
Residual Std. Error (df = 1023)	1.149	1.220	1.016	2.175
F Statistic (df = 3; 1023)	2.675**	3.608**	4.217***	4.016***

Note: The models refer to the following questions: **Left-right**: Self placement on 0 - 10 left-right scale of political orientation. **Training**: "Should it be the government's responsibility to ensure training opportunities for everyone who wants to upgrade their skills?" **Job guarantee**: "Should it be the government's responsibility to ensure a job for everyone who wants one?" **UE benefits**: "Should it be the government's responsibility to ensure a reasonable standard of living for the unemployed?" All models include weights. *p<0.1; **p<0.05; ***p<0.01.

E Additional Details on the Text Analysis

Excluding AI-generated responses

We used three methods to identify open-ended answers that were likely AI-generated. First, two authors undertook a qualitative review of responses to the question: "please briefly tell us in your own words what you think is most important when it comes to dealing with the effects of artificial intelligence on the labour market." AI-generated responses tended to be longer and contain repetitive language and structure. We compiled an initial list of responses with similar characteristics.

Next, we calculated Cosine Similarity Scores for each response using the following ChatGPT-generated answer as a reference document:

"When dealing with the effects of artificial intelligence on the labor market, the most important factors include:

1. *Worker Retraining and Education: Providing continuous learning opportunities and skill development programs to help workers adapt to new technologies and transition into new roles.*

2. *Policy and Regulation: Implementing policies that protect workers' rights, ensure fair wages, and regulate the integration of AI to prevent job displacement without adequate support.*

3. *Economic and Social Support: Establishing robust social safety nets, such as unemployment benefits and healthcare, to support workers affected by job automation.*

4. *Collaboration between Stakeholders: Encouraging collaboration between governments, businesses, educational institutions, and workers to create a cohesive approach to integrating AI into the labor market.*

5. *Ethical Considerations: Ensuring AI is used ethically and responsibly, with a focus on augmenting human work rather than merely replacing it, to promote an inclusive and fair labor market."*

We found a high degree of overlap between responses with the highest similarity scores and those identified through our initial qualitative review. We did a second qualitative review of responses identified as having high similarity scores and added additional responses to our initial list.

Finally, we ran all suspected AI-generated responses through Turnitin's AI detection service. Once again, we found a high degree of overlap between responses identified as AI-generated in Turnitin and those identified through qualitative review and through high similarity scores, giving us a high degree of confidence that these responses are indeed AI-generated. Our final list of exclusions is comprised of a total of 38 answers that are with a high probability generated by ChatGPT. We excluded these responses from our text analysis.

Additional Details about STM analysis

We used the *stm* package in R to generate topic models that we used to explore differences in which topics various subgroups of interest engaged with. We used the *findThoughts* function and qualitative review to determine the appropriate number of topics, beginning with 10 topics and gradually reducing the number of topics until we settled on an appropriate number (three). We judged that three topics avoided repetition, while still distinguishing between substantively different sets of ideas. The relatively small number of topics reflects the brevity and similarity of many of the open-ended responses in our survey.

F Additional Heterogeneous Effects

While gender and education are arguably the most prominent factors in the literature on labor market stratification, there are two more margins that are of particular interest with regard to generative AI: age and occupational background. Older individuals tend to be less familiar with novel technologies and professional and managerial occupations tend to require much more use of novel technologies than other occupations. This might in turn shape how people belonging to these groups react politically to the emergence of generative AI.

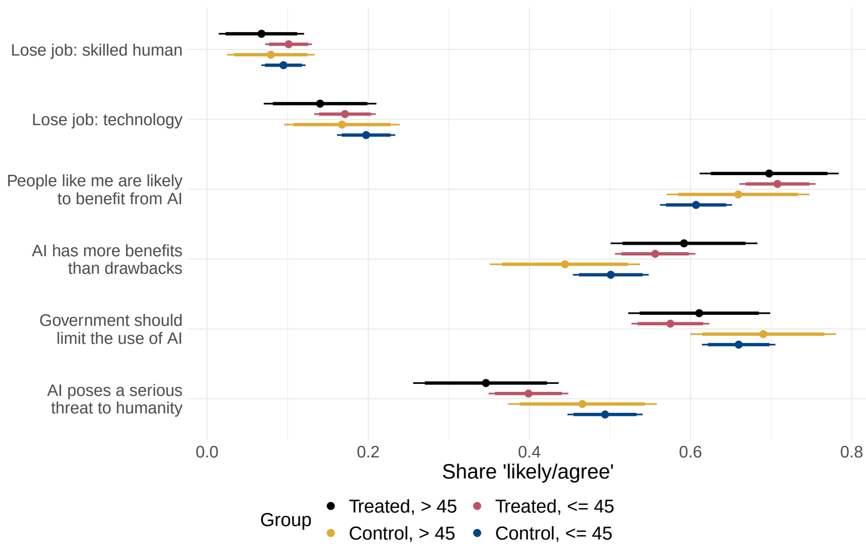
Table E9: Prevalence of topics, model with full controls

Variable	Topic 1	Topic 2	Topic 3
Treatment	-0.0129918^{**}	-0.0024186	0.0154090^{**}
	(0.0047250)	(0.0038277)	(0.0051061)
Age	0.0002504	0.0007208^{***}	-0.0009732^{***}
	(0.0002523)	(0.0002098)	(0.0002495)
Female	-0.0004633	0.0039519	-0.0034586
	(0.0052343)	(0.0044990)	(0.0060452)
Degree	-0.0114715	0.0004410	0.0110543
	(0.0064216)	(0.0057655)	(0.0077253)
Professional	-0.0159711^{**}	-0.0030007	0.0189229^{**}
	(0.0060635)	(0.0055620)	(0.0065680)
Intercept	0.3381209^{***}	0.2290809^{***}	0.4328775^{***}
	(0.0117124)	(0.0101228)	(0.0123889)

Age Differences

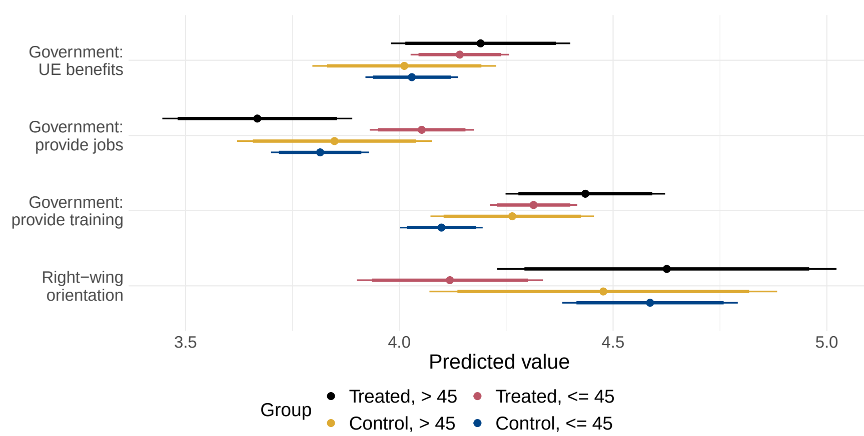
The effects of exposure to new technology might differ by age. Younger individuals tend to be more familiar with technology and have fewer reservations about using technology at work. Older workers, with fewer years left until retirement, might also be less able or willing to learn to use new technologies. They might therefore be more worried about losing their job and less likely to see technology as a good thing. However, there is little evidence for this in our sample, as Figure F2 shows. There are no significant baseline differences in the control group between people aged up to 45 years and older respondents, and no significant interactions. This is surprising and indicates that the middle-aged and older ChatGPT users that we were able to sample might be selected from a particularly tech-savvy section of this group. With regard to social policy preferences and ideology, we do see some differences. Here, the treatment has opposite effects on support for government job provision and political orientation. While younger workers show significant increases in support for job programs and an attendant shift to the left, the treatment has opposite, but insignificant effects on older workers. This also means that the interaction is (borderline) significant in both cases. All in all, the analysis of heterogeneous effects by age does not show a highly consistent picture. A dedicated analysis with a larger and more representative sample will be necessary to arrive at firm conclusions about how age moderates the effects of exposure to generative AI on attitudes and political preferences.

Figure F2: Age differences in effect on expectations and attitudes



Note: The figure shows predicted values based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). N = 1,027 in all models. Full model output in Table F10.

Figure F3: Age differences in effect on social policy preferences



Note: The figure shows predicted values based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). N = 1,027 in all models. Full model output in Table F11.

Table F10: Full output to Figure F2

	Dependent variable:					
	Lose job 1	Lose job 2	Benefit 1	Benefit 2	Gov. limit	AI threat
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.024	−0.025	0.164*	−0.038	−0.090	−0.070
	(0.056)	(0.074)	(0.091)	(0.096)	(0.093)	(0.096)
> 45 years	−0.016	−0.030	0.052	−0.057	0.030	−0.028
	(0.031)	(0.041)	(0.050)	(0.053)	(0.052)	(0.053)
Treatment x > 45 years	−0.018	−0.001	−0.063	0.093	0.005	−0.025
	(0.044)	(0.058)	(0.071)	(0.075)	(0.073)	(0.074)
Constant	0.110***	0.227***	0.554***	0.558***	0.629***	0.522***
	(0.039)	(0.052)	(0.064)	(0.068)	(0.065)	(0.067)
Observations	1,041	1,041	1,041	1,041	1,041	1,041
R^2	0.001	0.002	0.010	0.007	0.008	0.012
F Statistic (df = 3; 1037)	0.489	0.811	3.382**	2.486*	2.817**	4.085***
Note:						*p<0.1; **p<0.05; ***p<0.01

Table F11: Full output to Figure F3

	Dependent variable:			
	UE benefits	Job guarantee	Training	Left-right
	(1)	(2)	(3)	(4)
Treatment	0.046	0.656***	0.260	−1.085**
	(0.222)	(0.235)	(0.198)	(0.421)
> 45 years	−0.018	0.034	0.166	−0.109
	(0.123)	(0.130)	(0.109)	(0.232)
Treatment x > 45 years	0.066	−0.418**	−0.044	0.617*
	(0.173)	(0.183)	(0.154)	(0.327)
Constant	4.047***	3.781***	3.933***	4.696***
	(0.156)	(0.165)	(0.138)	(0.294)
Observations	1,041	1,041	1,041	1,041
R^2	0.003	0.012	0.014	0.011
F Statistic (df = 3; 1037)	1.119	4.161***	4.838***	3.669**
Note:				*p<0.1; **p<0.05; ***p<0.01

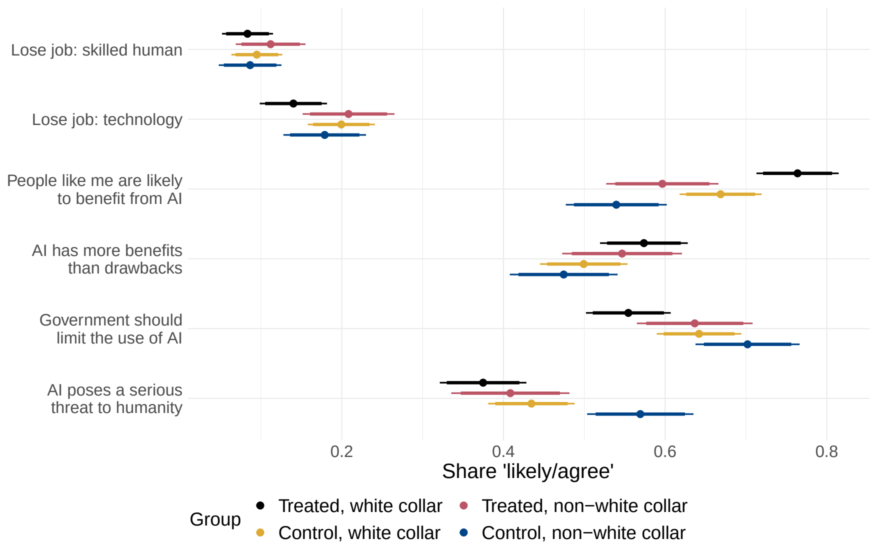
Occupational Differences

Finally, we consider heterogeneous effects by occupation. We distinguish between professional and managerial occupations (ISCO groups 1 and 2), which we refer to as "white collar", and the remaining occupational groups. Much commentary and early research argues that generative AI is likely to lead to substitution in different occupational groups than previous waves of automation. Felten, Raj and Seamans (2023a) argue that white collar workers are likely to be most exposed to generative AI. Hui, Reshef and Zhou (2023) show that white collar contract work declined both in quantity and remuneration after the introduction of ChatGPT. Thus, occupations that used to be among the winners in recent automation processes might now find themselves on the losing end. If this were true, it could have profound implications for political behavior. Our results show, however, that the people who supposedly should be worried, are not.

Figure F4 shows that while treated white collar workers see an insignificant reduction in subjective risk, non-white collar workers experience a small increase. The interaction is marginally statistically significant ($p < 0.1$) in the case of losing one's job to technology. There are no further significant interactions, indicating that, despite some baseline differences in attitudes towards AI, exposure has similar effects on white collar and non-white collar workers. When it comes to social policy preferences, both baseline differences and differences in treatment effects are minor. The reactions of non-white collar workers tend to be ever so slightly more pronounced, but these differences are far from statistical significance. However, the shift to the political left is virtually entirely driven by non-white collar workers. This suggests two non-mutually exclusive interpretations. First, white collar workers seem to be oblivious of the difficulties that generative AI, at least in the opinion of many experts, will cause them. At the same time as predictions abound that people working in professional and managerial occupations will face increased insecurity, workers who experience the technology remain unconcerned. Second, the fact that non-white collar workers show slightly stronger reactions indicates that their views and preferences may be less firmly held, owing to less information. It could, however, also be an artifact of the smaller sample size in this group. Overall, the heterogeneous effects of our treatment on people in different occupational groups are interesting as they do not align with the prominent narrative about who should be affected by generative AI, but again, more evidence will be necessary before firm conclusions can be drawn.

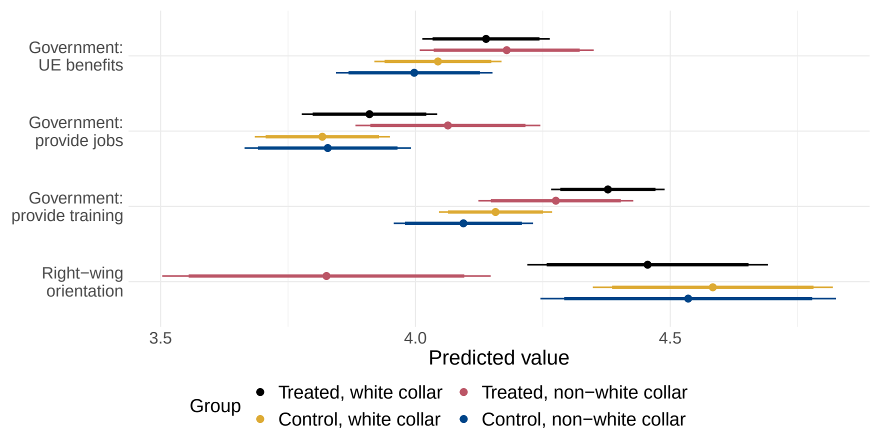
G Additional Analyses

Figure F4: Occupational differences in effect on expectations and attitudes



Note: The figure shows predicted values based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). N = 1,027 in all models. Full model output in Table F12.

Figure F5: Occupational differences in effect on social policy preferences



Note: The figure shows predicted values based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). N = 1,027 in all models. Full model output in Table F13.

Table F12: Full output to Figure F4

	Dependent variable:					
	Lose job 1	Lose job 2	Benefit 1	Benefit 2	Gov. limit	AI threat
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.025	0.029	0.057	0.072	−0.065	−0.161***
	(0.030)	(0.039)	(0.048)	(0.051)	(0.049)	(0.050)
Professional	0.008	0.021	0.129***	0.025	−0.060	−0.135***
	(0.025)	(0.034)	(0.041)	(0.044)	(0.042)	(0.043)
Treatment x Professional	−0.037	−0.089*	0.038	0.002	−0.022	0.101
	(0.037)	(0.049)	(0.060)	(0.064)	(0.062)	(0.063)
Constant	0.087***	0.179***	0.540***	0.475***	0.702***	0.569***
	(0.020)	(0.026)	(0.032)	(0.034)	(0.033)	(0.034)
Observations	1,041	1,041	1,041	1,041	1,041	1,041
R^2	0.001	0.005	0.031	0.006	0.012	0.020
F Statistic (df = 3; 1037)	0.406	1.770	11.227***	2.170*	4.320***	7.101***
Note:						*p<0.1; **p<0.05; ***p<0.01

Table F13: Full output to Figure F5

	Dependent variable:			
	UE benefits	Job guarantee	Training	Left-right
	(1)	(2)	(3)	(4)
Treatment	0.181	0.236*	0.181*	−0.710***
	(0.117)	(0.124)	(0.104)	(0.221)
Professional	0.047	−0.011	0.063	0.048
	(0.101)	(0.107)	(0.090)	(0.190)
Treatment x Professional	−0.087	−0.143	0.039	0.582**
	(0.148)	(0.157)	(0.131)	(0.279)
Constant	3.998***	3.828***	4.094***	4.535***
	(0.078)	(0.083)	(0.070)	(0.148)
Observations	1,041	1,041	1,041	1,041
R^2	0.003	0.005	0.012	0.015
F Statistic (df = 3; 1037)	1.177	1.776	4.190***	5.206***
Note:				*p<0.1; **p<0.05; ***p<0.01

Table G14: Threat perceptions and expectations: Models with controls

	Dependent variable:					
	Replace 1	Replace 2	Benefit 1	Benefit 2	Gov. limit	AI threat
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	−0.005	−0.037	0.066**	0.059*	−0.088***	−0.115***
	(0.019)	(0.026)	(0.031)	(0.033)	(0.032)	(0.033)
Female	0.006	−0.008	−0.102***	−0.077**	0.076**	0.055
	(0.022)	(0.029)	(0.035)	(0.038)	(0.036)	(0.037)
Degree	−0.001	−0.015	0.013	−0.068*	−0.028	−0.033
	(0.023)	(0.030)	(0.035)	(0.038)	(0.037)	(0.038)
Age	−0.001	−0.002	−0.001	−0.002	−0.0001	−0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Income	0.004	0.003	0.024***	0.015***	−0.011**	−0.009*
	(0.003)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Constant	0.111***	0.256***	0.504***	0.517***	0.767***	0.615***
	(0.040)	(0.052)	(0.063)	(0.068)	(0.065)	(0.067)
Observations	914	914	914	914	914	914
R^2	0.003	0.005	0.054	0.021	0.025	0.026
F Statistic (df = 5; 908)	0.603	0.975	10.392***	3.990***	4.700***	4.771***

Note: The models model agreement with the following statements. **Replace 1**: "I will lose my job because I am not good enough with new technologies or because I will be replaced by someone with better technological skills." **Replace 2**: "My job will be replaced by an artificial intelligence, algorithm, computer software, or robot." **Benefit 1**: "People like me are likely to benefit from AI." **Benefit 2**: "Products and services using AI have more benefits than drawbacks." **Gov. limit**: "The government should limit the use of some forms of AI." **AI threat**: "AI poses a serious threat to humanity." *p<0.1; **p<0.05; ***p<0.01.

Table G15: Social policy preferences: Models with controls

	Dependent variable:			
	Left-right	Training	Job guarantee	UE benefits
	(1)	(2)	(3)	(4)
Treatment	−0.445***	0.154**	0.138*	0.105
	(0.141)	(0.067)	(0.082)	(0.077)
Female	−0.305*	−0.039	0.082	0.044
	(0.161)	(0.077)	(0.094)	(0.088)
Degree	−0.374**	0.114	0.135	0.115
	(0.163)	(0.078)	(0.095)	(0.090)
Age	0.001	0.007**	−0.007*	0.001
	(0.007)	(0.003)	(0.004)	(0.004)
Income	0.070***	0.007	−0.005	−0.010
	(0.021)	(0.010)	(0.012)	(0.012)
Constant	4.317***	3.792***	4.021***	3.980***
	(0.288)	(0.138)	(0.168)	(0.158)
Observations	914	914	914	914
R^2	0.032	0.018	0.011	0.005
F Statistic (df = 5; 908)	6.050***	3.398***	2.100*	0.859

Note: The models refer to the following questions. **Left-right**: Self placement on 0 - 10 left-right scale of political orientation. **Training**: "Should it be the government's responsibility to ensure training opportunities for everyone who wants to upgrade their skills?" **Job guarantee**: "Should it be the government's responsibility to ensure a job for everyone who wants one?" **UE benefits**: "Should it be the government's responsibility to ensure a reasonable standard of living for the unemployed?" *p<0.1; **p<0.05; ***p<0.01.

Table G16: Threat perceptions and expectations: Models excluding frequent users

	Dependent variable:					
	Replace 1	Replace 2	Benefit 1	Benefit 2	Gov. limit	AI threat
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.007	−0.035	0.066*	0.067*	−0.079**	−0.100***
	(0.020)	(0.027)	(0.037)	(0.037)	(0.035)	(0.037)
Constant	0.072***	0.174***	0.556***	0.449***	0.701***	0.511***
	(0.013)	(0.018)	(0.024)	(0.024)	(0.023)	(0.024)
Observations	732	732	732	732	732	732
R^2	0.0002	0.002	0.004	0.004	0.007	0.010
Adjusted R^2	−0.001	0.001	0.003	0.003	0.006	0.008
Residual Std. Error (df = 730)	0.260	0.361	0.486	0.493	0.464	0.491
F Statistic (df = 1; 730)	0.132	1.603	3.183*	3.245*	5.093**	7.262***

Note: The models model agreement with the following statements. **Replace 1**: "I will lose my job because I am not good enough with new technologies or because I will be replaced by someone with better technological skills." **Replace 2**: "My job will be replaced by an artificial intelligence, algorithm, computer software, or robot." **Benefit 1**: "People like me are likely to benefit from AI." **Benefit 2**: "Products and services using AI have more benefits than drawbacks." **Gov. limit**: "The government should limit the use of some forms of AI." **AI threat**: "AI poses a serious threat to humanity." *p<0.1; **p<0.05; ***p<0.01.

Table G17: Social policy preferences: Models excluding frequent ChatGPT users

	Dependent variable:			
	Left-right	Training	Job guarantee	UE benefits
	(1)	(2)	(3)	(4)
Treatment	−0.413***	0.226***	0.084	0.092
	(0.153)	(0.075)	(0.091)	(0.088)
Constant	4.435***	4.156***	3.841***	4.019***
	(0.100)	(0.049)	(0.059)	(0.058)
Observations	732	732	732	732
R^2	0.010	0.012	0.001	0.001
Adjusted R^2	0.009	0.011	−0.0002	0.0001
Residual Std. Error (df = 730)	2.024	0.994	1.198	1.163
F Statistic (df = 1; 730)	7.273***	9.041***	0.869	1.083

Note: The models refer to the following questions. **Left-right**: Self placement on 0 - 10 left-right scale of political orientation. **Training**: "Should it be the government's responsibility to ensure training opportunities for everyone who wants to upgrade their skills?" **Job guarantee**: "Should it be the government's responsibility to ensure a job for everyone who wants one?" **UE benefits**: "Should it be the government's responsibility to ensure a reasonable standard of living for the unemployed?"
*p<0.1; **p<0.05; ***p<0.01.