

Rage against the machine? Generative AI exposure, subjective risk, and policy preferences

Matthias Haslberger, Jane Gingrich & Jasmine Bhatia

To cite this article: Matthias Haslberger, Jane Gingrich & Jasmine Bhatia (10 Sep 2025): Rage against the machine? Generative AI exposure, subjective risk, and policy preferences, Journal of European Public Policy, DOI: [10.1080/13501763.2025.2554903](https://doi.org/10.1080/13501763.2025.2554903)

To link to this article: <https://doi.org/10.1080/13501763.2025.2554903>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 10 Sep 2025.



[Submit your article to this journal](#)



[View related articles](#)



[View Crossmark data](#)

Rage against the machine? Generative AI exposure, subjective risk, and policy preferences

Matthias Haslberger^a, Jane Gingrich^b and Jasmine Bhatia^c

^aUniversity of St.Gallen, St. Gallen, Switzerland; ^bUniversity of Oxford, Oxford, UK; ^cBirkbeck, University of London, London, UK

ABSTRACT

How does novel technology change public policy demands? Scholars interested in the effect of automation on policy preferences have commonly argued that exposure to automation technology increases subjective risk, which in turn predicts demand for insurance. Generative AI potentially challenges this dynamic. Based on a pre-registered online experiment with a sample of 1041 UK working-age adults we show that direct exposure to generative AI in realistic work tasks does not increase subjective risk but strengthens support for activating social policy. To understand this constellation of attitudes, we argue that exposure to technology may activate sociotropic preferences to support individuals who might be negatively affected by AI. Text analysis shows cautious optimism and thoughtful engagement with the implications of AI for work and social policy. Our findings suggest that the current uncertainty over the relative winners and losers from AI opens a window of opportunity to expand activating social policies.

ARTICLE HISTORY Received 19 March 2025; Accepted 25 August 2025

KEYWORDS Social policy; artificial intelligence; subjective risk; technological change; survey experiment; text analysis

Introduction

Does using generative AI make people more or less concerned about disruptive effects – either for their own jobs or those of others? How do people think that the use of AI technology should 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 labour market

CONTACT Matthias Haslberger  matthias.haslberger@unisg.ch  University of St.Gallen, Müller-Friedberg-Strasse 8, St. Gallen, 9000, Switzerland

 Supplemental data for this article can be accessed online at <http://dx.doi.org/10.1080/13501763.2025.2554903>.

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

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 large sample of 1041 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 subjective risk perceptions, general attitudes towards generative AI, and social policy preferences. This study, which we fielded in July 2023, before the mass adoption of generative AI in most workplaces, allows us to observe the causal effects of exposure to AI technology in realistic tasks on political preferences in a sample of the general population. As such, it complements existing studies that examine occupation-specific samples (Brynjolfsson *et al.*, 2025; Dell'Acqua *et al.*, 2023; Noy & Zhang, 2023), or do not consider political preferences (Brynjolfsson *et al.*, 2025; Dell'Acqua *et al.*, 2023; Haslberger *et al.*, 2025). 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 labour 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 public debate on possible dangers of AI, we find that individuals who are exposed to generative AI (ChatGPT) in our treatment do not perceive their jobs to be at greater risk of replacement by technology or a more tech-savvy person. Furthermore, treated respondents 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. Although exposure to AI does not increase perceived subjective economic risks, we do find that treated respondents express greater support for activating social policy. This constellation of preferences contradicts established work on the relationship between technology and policy preferences, which suggests that the relationship between technology and policy demands largely runs through changing risk perceptions (see, e.g., Gallego & Kurer, 2022). To make sense of these findings, we argue that AI exposure may activate sociotropic concerns, whereby individuals who expect to benefit from the technology see it as providing them with personal gains but leading to collective disruption. Hence, they support government intervention for those who – through no fault of their own – are negatively affected by AI. Interaction analyses suggest that sociotropic motivations are more pronounced among men, whereas there are no heterogeneous effects by occupation.

The paper is structured as follows. Section ‘Argument’ 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 ‘Methods’ outlines the design of the survey and our analytical approach. In Section ‘Results’, we present our findings, and a final section concludes by discussing the generalizability of the results and ways forward for research on the political implications of generative AI.

Argument

In this paper we study the effect of exposure to generative AI on subjective risk and policy preferences. We lay out our argument in three steps. First, we discuss how the literature before the advent of AI has conceptualized the link between technology, risk, and policy preferences, and the empirical evidence for this ‘risk-insurance framework.’ Second, we discuss the application of this framework to AI. In a third step, we propose an alternative perspective and argue that at an early stage of AI adoption, people are likely to be motivated more by sociotropic concerns than by material self-interest.

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 (actual or perceived) risk and how these risks, in turn, shape preferences.

First, this work examines the relationship between exposure to novel technology and individuals’ risks of job loss, changes in job quality, or income prospects. Analysts measure risks in varying ways. Some studies focus on subjective perceptions of risk, such as self-assessed unemployment risk, whereas others employ ‘objective’ measures of exposure to job replacing technologies, such as routine task intensity (RTI) scores at the occupation or sector level.

Second, this work theorizes that the relationship between technology exposure and preferences is mediated by this risk. Studies examine both the demand for compensatory policies (such as unemployment benefits or a universal basic income) and activation measures (for example, retraining opportunities and active labour market policies) (Bürgisser, 2023). The underlying logic in both cases is that people demand insurance against labour market risks posed by technology in line with their degree of exposure. We refer to this approach as the ‘risk-insurance framework.’

A substantial body of observational work shows that people who perceive higher levels of technologically driven risk tend to favor compensatory

policies such as unemployment benefits (Busemeyer & Tober, 2023; Dermont & Weisstanner, 2020; Kurer & Häusermann, 2022; Sacchi *et al.*, 2020; Thewissen & Rueda, 2019), although some also report null findings (Gallego *et al.*, 2022). Only a handful of studies find a positive association between technological risk and support for activating policies (Im, 2021; Im & Komp-Leukkuinen, 2021), while others report a null or even negative relationship (Busemeyer & Sahm, 2022; Busemeyer & Tober, 2023; Kurer & Häusermann, 2022). There are no clear differences in these observed patterns across whether studies use subjective or objective risk measures.

Whether these results demonstrate a causal relationship between technology, risk, and preferences, however, is less clear. While the risk-insurance framework builds on well-developed theoretical approaches that model the demand for insurance as a function of job-related economic risks (Rehm, 2011), recent wide-ranging scholarly reviews by both Gallego and Kurer (2022) and Weisstanner (2023) show that there is limited empirical consensus on the explanatory power of the risk-insurance framework. A small number of studies have attempted to experimentally identify the effects of technology exposure on social policy preferences, producing more modest evidence for the risk-insurance framework.

Gallego *et al.* (2022) use a priming experiment and find that exposure to information about technological risks does not increase support for compensatory policies, but instead for measures to slow down technological change, a form 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 – does not affect policy preferences, even when it changed factual beliefs about automation risks. Jeffrey (2021) shows, in an experimental study exposing respondents to varying rhetoric about automation, that compared to a treatment highlighting only risks, framing automation as also unfair leads to increased support for social policy. Overall, the evidence from these experimental studies shows only limited support for the risk-insurance framework. Nonetheless, 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 *et al.*, 2021).

Thus, existing research provides some evidence that technological exposure leads to greater support for passive, compensatory policies but not for activating ones, and that contextual factors such as existing welfare state institutions or education systems may play a moderating role (Busemeyer & Sahm, 2022; Gingrich & Ansell, 2012; Haslberger *et al.*, 2023). However, empirical questions remain about the extent to which people's political reactions to technological change correspond to the predictions of the risk-insurance framework.

Application to generative AI

The preceding discussion thus leaves us with two key questions: does actual exposure to generative AI increase respondents' perceptions of individual risk? And do these risk perceptions shape respondents' preferences towards compensatory or regulatory policies?

The logic of the risk-insurance framework underpins much of the current academic debate about the political implications of generative AI tools. Economic research focuses on identifying the potentially large risks of AI for employment – particularly among groups of workers insulated from past rounds of automation (Autor, 2022; Frey & Osborne, 2023; Gmyrek *et al.*, 2023). There is evidence that people in white collar jobs, who have until now largely benefited from technological change, might be at increased risk of replacement from generative AI (Felten *et al.*, 2023a; Hui *et al.*, 2024), reconfiguring the future winners and losers from technological progress (Autor, 2024; Knotz, 2025). These debates, however, remain relatively removed from the wider public – existing research shows that AI governance is not yet highly politicized (Margalit & Raviv, 2024; Schiff *et al.*, 2024).

If the risk-insurance framework holds, we would thus expect direct exposure to powerful new AI technology in common work tasks to provide novel information about its prospective effects on the future of work. If this novel information increases people's perception that AI tools could threaten their job stability or long-run income prospects – i.e., increase their subjective perception of risk – then we would expect rational, self-interested individuals to adopt a generally negative view of the technology and expand their support for policies to alleviate these risks. Research distinguishes between compensation, activation or investment, and steering policies (Bicchi *et al.*, 2024; Bürgisser, 2023). While economists typically advocate skill-focussed policies (Felten *et al.*, 2023b), we have seen above that compensatory measures are likely to be more popular with the public. Gallego *et al.* (2022) show that policies to regulate and slow down the adoption of new technologies might also be popular with electorates. Based on this reading of the extant literature, we therefore pre-registered the following hypotheses:¹

- H1.1: Treated respondents are *more likely* to fear job loss due to technology.
- H1.2: Treated respondents express *more negative* attitudes towards AI.
- H1.3: Treated respondents exhibit *more progressive* social policy preferences.

In short, this set of hypotheses predicts that people's response to encountering generative AI should fundamentally mirror previous instances of automation – we should expect that on average, it will increase subjective risk and

thereby trigger skepticism towards AI and strengthen demand for social policy – particularly of the compensatory kind.

An alternative perspective based on sociotropic preferences

However, there are reasons for hesitation around these expectations. First, AI may *not* mirror previous instances of technological change. Second, as the literature review showed, work on other forms of technological change does not unambiguously support the risk-insurance framework. This suggests that, in some cases, individuals may draw on alternative forms of reasoning. We address each point in turn.

Exposure to generative AI may reveal information that reduces, rather than increases, subjective risk, setting in motion what we might call a reverse risk-insurance framework. Busemeyer et al. (2023) find a weak tendency for workers who use ICT more frequently to express lower subjective technology risk. If exposure to generative AI similarly decreases (or at least does not increase) our participants' perceived employment risk, we would expect it to instead instill confidence that they will personally benefit from generative AI. In our case, the impressive but still-limited capabilities of early generative AI models may have temporarily alleviated fears of job loss, even as they increased productivity, in particular against the background of an alarmist discourse predicting imminent job destruction. This should apply *a fortiori* to the participants in our study, who by definition are early adopters and therefore prone to picture themselves among the prospective winners – after all, they are ahead of the curve with this groundbreaking technology. Thus, the productivity boost from our treatment might in fact reduce subjective risk and lead to more positive attitudes towards AI. Based on a rational calculation of material self-interest, this should then lead to a reduction in support for social policies which our respondents would not expect to benefit from. Effectively, this would amount to an inverse formulation of the standard risk-insurance framework.

However, if we relax the assumption that preferences about policy primarily follow from risk perceptions, we might expect an alternative response. In particular, an emerging body of work suggests that fairness considerations matter for how people approach social policy. 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). Several studies find that people exhibit greater support for social insurance when macroeconomic conditions are difficult – even if they are not themselves unemployed (Funk, 2000; Kam & Nam, 2008). Compton and Lipsmeyer (2019) qualify that individual insecurity is a stronger motivator of support for social policy, but 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.

These considerations direct us towards what we might call a sociotropic model of the politics of generative AI. Respondents exposed to AI receive information not just about their own performance, but also, albeit less explicitly, about implications for others. It is easy to see why they might anticipate negative consequences for some workers. Crucially, the speed at which generative AI is threatening to reshape the world of work (see, e.g., Bick *et al.*, 2024; Humlum & Vestergaard, 2025) and the fundamental uncertainty over who the eventual winners and losers of this process are going to be, makes it harder to blame workers for failing to adapt. This uncertainty over distributional implications arguably represents the most fundamental departure from earlier, more gradual automation processes (Acemoglu & Johnson, 2023), and makes it more likely that losers are seen as ‘deserving’ by respondents who do not consider themselves at risk (Cavaillé, 2023). This perspective suggests that even where respondents observe their own economic prospects improve with AI, it is not clear that they will interpret these processes as unequivocally socially beneficial. They could see (a) risks that other workers will be negatively affected by AI through no fault of their own while at the same time (b) anticipating that AI will bring productivity gains large enough to financially compensate losers. If such a sociotropic logic holds, following Jeffrey (2021) and others, we would expect people who anticipate negative consequences for others to transcend narrow self-interest and support progressive social policies – especially activating ones such as retraining programs.

It is worth noting that people may also express ostensibly sociotropic preferences for self-interested reasons. 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. However, this on its own is unlikely to overturn the effect predicted by the risk-insurance framework without genuine sociotropic motivations. Taken together, this set of arguments leads to a conflicting set of theoretical predictions which were, however, not pre-registered:

- H2.1: Treated respondents are *no more likely* to fear job loss due to technology.
- H2.2: Treated respondents express *more positive* attitudes towards AI.

- H2.3: Treated respondents exhibit *more progressive* social policy preferences.

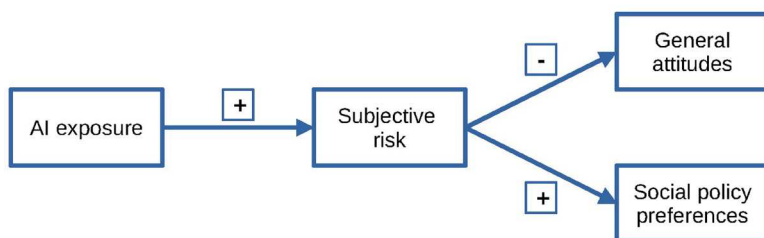
Note that H2.3 is identical to H1.3, as the prediction for social policy preferences is the same; merely the causal path is different: sociotropic concerns rather than material self-interest. We schematize these two contrasting sets of expectations in Figure 1. In scenario A, which corresponds to the standard narrative in political economy, AI exposure increases people's subjective risk, based on which they adopt more negative attitudes towards AI technology and become more supportive of social policies. In scenario B, AI exposure 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 exposure activates sociotropic concerns for others who might be at risk, which then positively affects support for social policy.

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

A: Risk-Insurance Framework



B: Sociotropic Model

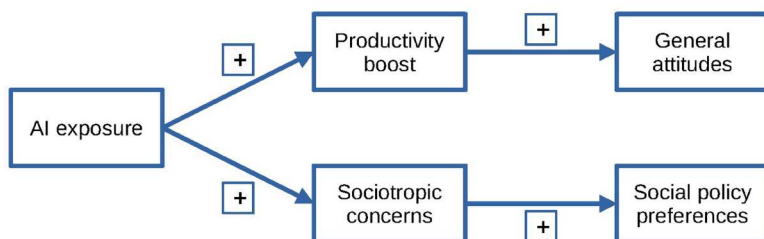


Figure 1. Two sets of theoretical expectations.

Committee. It adheres to APSA's Principles and Guidance for Human Subjects Research.² Through the survey company YouGov, we recruited a sample of 1041 respondents who had ChatGPT accounts and who are largely representative of the UK working-age population.³ This makes our study the first (to our knowledge) 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. Random allocation ensured balance on key socio-demographic characteristics (see Table B2). Appendix C reports a range of diagnostic analyses which indicate that our treatment worked and the vast majority of respondents complied with the instructions. 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.

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 shows screenshots of the treatment instructions and tasks and 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. Figure C1 shows that most respondents perform similar tasks at work at least sometimes. Therefore, compared to niche tasks such as coding, the tasks allowed respondents to relate the capabilities of AI to their everyday work, increasing the generalizability of our results.⁴ Moreover, 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. This design choice addresses some of the limitations to the external validity of information experiments (Barabas & Jerit, 2010).

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

Analytical procedure

Post-treatment, respondents encountered a range of attitudinal items. With this set of questions, we intend to provide an overview of people’s concerns and preferences that can serve as a starting point for future research into a broader range of distributive trade-offs and longer-term attitude changes. Table 2 lists the four question blocks. We first asked about automation-related subjective job risk, followed by questions eliciting general expectations and attitudes regarding AI. We next asked about social policy preferences. The order of these questions was randomized within question blocks. 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. YouGov provided demographic information, including age, sex, education, occupation, income, and other variables.

We dichotomize the quantitative outcome variables and estimate linear probability models which are straightforward to interpret as the probability of agreeing with a statement. For the main analyses, we regress the outcome on a treatment dummy. That is, we estimate intention-to-treat (ITT) effects, as we do not have perfect control over compliance with the treatment instructions. This is theoretically indicated, as our argument, following the automation literature, focuses on technology *exposure* rather than use.⁵ To analyze potential heterogeneous effects of AI exposure, we

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. • AI will increase surveillance in the workplace. • 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 reasonable standard of living for the unemployed? • ...ensure a job for everyone who wants one? • ...ensure training opportunities for everyone who wants to upgrade their skills?	5-point scale, dichotomized
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

interact the treatment dummy with a dummy for important socio-demographic characteristics (sex, occupation, age, education).

Given our research design, the effect sizes we estimate are likely to be conservative: since the sample is recruited from people with a ChatGPT account, even the control group has at least some prior awareness of generative AI. This was necessary since at the time it was not possible to embed ChatGPT directly in the survey for data protection reasons. Insofar as there is a 'wow' effect for those who are newly exposed to AI technology, we should therefore find it somewhat subdued. However, with a view to the increasing adoption of AI in the workplace, this design strikes us as more realistic than a scenario assuming zero prior awareness of generative AI. By the time it reaches their organization, most people would have a level of awareness of AI at least commensurate to that of the respondents in our study. Arguably, this turns an apparent weakness of our study design into a strength.

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 such as sentiment analysis and structural topic modelling can then be used to elicit first-order concerns regarding policy issues, akin to their widespread use on non-survey data (Ferrario & Stantcheva, 2022). 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

Overall, our analyses support 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. This supports H2.1 over H1.1 and thus rules out the standard formulation of the risk-insurance framework, for if the treatment does not increase subjective risk, a positive effect on social policy preferences cannot operate through risk perceptions. Since a clear majority of AI users at the same time rated it as useful (see Figure C2), this indicates that people were not simply underwhelmed by the capabilities of ChatGPT. Yet, they did not worry about possible substituting effects. This is surprising as the questions explicitly ask respondents to consider possible risks associated with AI over a 5-year time horizon. One plausible explanation is that people may actually welcome text-based tasks such as those the respondents

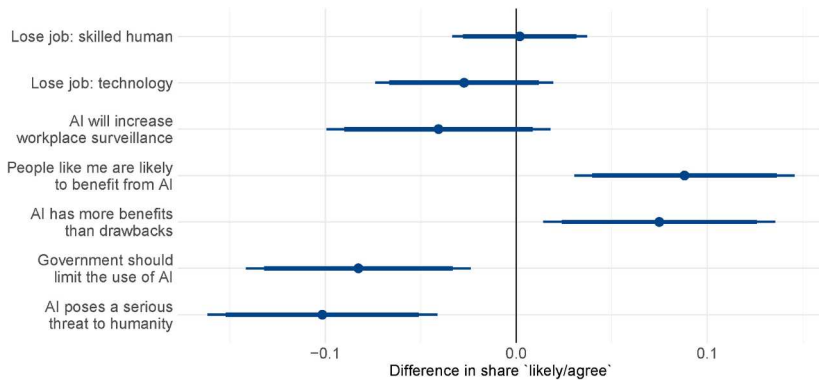


Figure 2. Effect of AI exposure on expectations and attitudes.

Note: Treatment effects based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). $N = 1041$ in all models. Full output in Table D1.

performed in our study being taken over by AI. Still, this would reveal a static conception of AI capabilities and a lack of forward-looking thinking about future AI capabilities. Hence, our finding suggests that, if AI will really eliminate vast numbers of jobs, the soon-to-be-jobless are oblivious of what is in store for them.

Moreover, H2.2 is supported over H1.2, as people in the treatment group express more positive attitudes towards AI. Treated respondents are 8 to 9 percentage points more likely to see AI as a positive force for themselves and in general, and 8 to 10 percentage points less likely to see AI as a threat which should be restricted by the government. There is no significant difference in the share of respondents who expect workplace surveillance to increase. That treated respondents are more likely to say that they expect to benefit personally from using AI further indicates that they are unconcerned about the potential second-order effects of a productivity-increasing technology that is potentially capable of replacing a vast number of workers. 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 survey respondents in Spain advocated for policies slowing down technological change. However, they look at concern about technology rather than individual technology exposure, which, as we have seen, need not be correlated. The absence of an effect of AI exposure on subjective risk and the positive effect on overall attitudes contradicts the first building blocks of the standard formulation of the risk-insurance framework, whereby exposure increases risk, technology skepticism, and ultimately support for social policy. We must now look to the effect on social policy preferences to adjudicate between an inverse formulation of the risk-insurance framework (where the negative/null effect on risk and the positive effect

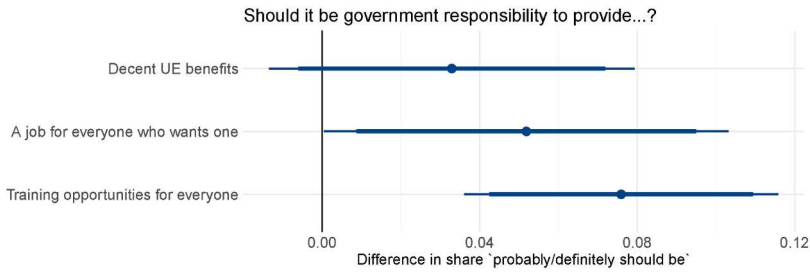


Figure 3. Effect of exposure to AI on social policy preferences.

Note: Treatment effects based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). $N = 1041$ in all models. Full output in Table D2.

on overall attitudes of AI exposure is accompanied by a negative effect on social policy preferences) and the sociotropic model.

Figure 3 shows that the effect of exposure to AI on social policy preferences is in line with the theoretical predictions of the sociotropic model (H2.3). Respondents in the treatment group are slightly more likely to advocate for the government to provide for the unemployed, but this effect does not reach statistical significance. However, treated respondents are about 5 percentage points ($p < 0.05$) more likely to agree that it should be the government's responsibility to provide a job for everyone who wants one and express almost 8 percentage points higher support for government-provided training opportunities ($p < 0.001$). 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 help them adapt (Weisstanner, 2023). However, since the respondents in our study mostly do not think of these policies as intended to benefit them personally, it is unsurprising that they support activating measures over compensation. This is exactly what one would expect from respondents who are motivated by sociotropic concerns about the capacity of others to adapt to AI. The results therefore paint a consistent picture in line with the sociotropic model. It appears that individuals exposed to AI in our study consider themselves to be insulated from potential negative employment effects, but after experiencing the capabilities of generative AI nevertheless acknowledge a need for redistributive state action aimed at helping workers who might be negatively affected by AI to adapt. Our findings suggest that there may be a window of opportunity for public policy during which the (anticipated) winners from the proliferation of generative AI support social policies that are presented as cushioning the blow for the losers from AI. However, with the continuing expansion of AI capabilities, the more self-interested reasoning predicted by the risk-insurance model may reassert itself once the threat to jobs by AI becomes more tangible. A corollary of this is that if firms

prioritized worker-complementing applications of AI, this could help delay the resurgence of self-interested thinking.

One important caveat is that by virtue of having a ChatGPT account, our respondents might be highly selected from a technological avant-garde. Note, however, that the majority of study participants report using ChatGPT rarely (47%) or never/don't know (24%). Only 30% stated that they use ChatGPT at least weekly, which is very similar to what Bick *et al.* (2024) report based on a nationally representative survey fielded in the US in August 2024. In other words, assuming similar adoption rates in the UK and US, only a year later the general population was as familiar with generative AI as our sample of summer 2023 ChatGPT users. Moreover, 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 Tables G2 and G3. This dispels concerns that our findings might be driven by self-selection: we observe the same patterns in individuals who have had minimal exposure to generative AI before the study and therefore, despite having a ChatGPT account, are unlikely to differ substantially from the general population. Thus, while we acknowledge the selection issue, we see little evidence that it undermines the external validity of our findings.

Additionally, the models shown here 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). Nevertheless, we show in Tables G4 and G5 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 (Tables G6 and G7). When we calculate the local average treatment effect (LATE) by instrumenting compliance with the treatment instructions with treatment assignment, we obtain similar results (Tables G8, G9 and G10). Hence, we are confident in the robustness and generalizability of our main results.

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, occupation, age, or education are some of the most persistent features of modern labour markets and wider societies (Blau & Kahn, 2017; Goldin, 2014; Haslberger, 2021). Exposure to, familiarity with, and attitudes towards novel technologies also vary along these lines (Borwein *et al.*, 2024; Carvajal *et al.*, 2024; Haslberger *et al.*, 2025; OECD, 2018). It is therefore plausible that exposure to AI affects the views of members of various groups differently. The next set of analyses explores these possibilities by interacting our treatment dummy with the respective

demographic indicators. For reasons of space, we focus on gender and occupation differences in the main text, and provide results for age and education in Appendix F.

Gender differences

There is mounting evidence that women are more hesitant to use AI (Bick *et al.*, 2024; Haslberger *et al.*, 2025; Humlum & Vestergaard, 2025) and may consequently lag behind men in building AI-related skills (Carvajal *et al.*, 2024), replicating earlier patterns of gender differences in technology use (OECD, 2018). Recent work also suggests that women are generally more skeptical of technology use in the workplace (Borwein *et al.*, 2024), and face different risk environments with regard to technology in general (Gallego & Kurer, 2022; Gingrich & Kuo, 2022) and AI technology in particular (Felten *et al.*, 2023b; Gmyrek *et al.*, 2023). We would therefore expect that women are more likely than men to conform to the risk-insurance framework. Hence, we would expect them to experience higher subjective risk and develop more negative views towards AI alongside stronger preferences for progressive social policy. However, we find few significant heterogeneous effects, and those we do find point in the opposite direction.

Panel A of Figure 4 shows treatment effects by sex for the subjective risk questions. While in the control group, women are 5 percentage points more likely than men to fear replacement by someone with better technology skills, this relationship is reversed in the treatment group, where men are 3 percentage points more likely to fear job loss – a statistically significant interaction. Thus, the precise null finding in Figure 2 is the result of two borderline significant opposing effects canceling each other out. This finding suggests that direct exposure to AI alleviates fears about a lack of skill in women, narrowing some pre-existing gender gaps (OECD, 2018). By contrast, we find neither treatment nor gender effects on people's expectation that they will lose their job to technology directly.

Panel B shows results for the expectations and attitudes questions. As expected, 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. The other baseline differences are insignificant but directionally also point towards greater AI-skepticism among women. While the treatment leads to significantly more positive views in men in all cases except for the surveillance question, we find no significant interaction effects. This indicates that the treatment affects women and men similarly positively, even as women tend to exhibit higher baseline skepticism regarding AI and most of the gaps narrow somewhat.

The results for social policy preferences align with the findings for subjective risk (see Panel C). The effect on support for training programs is larger for

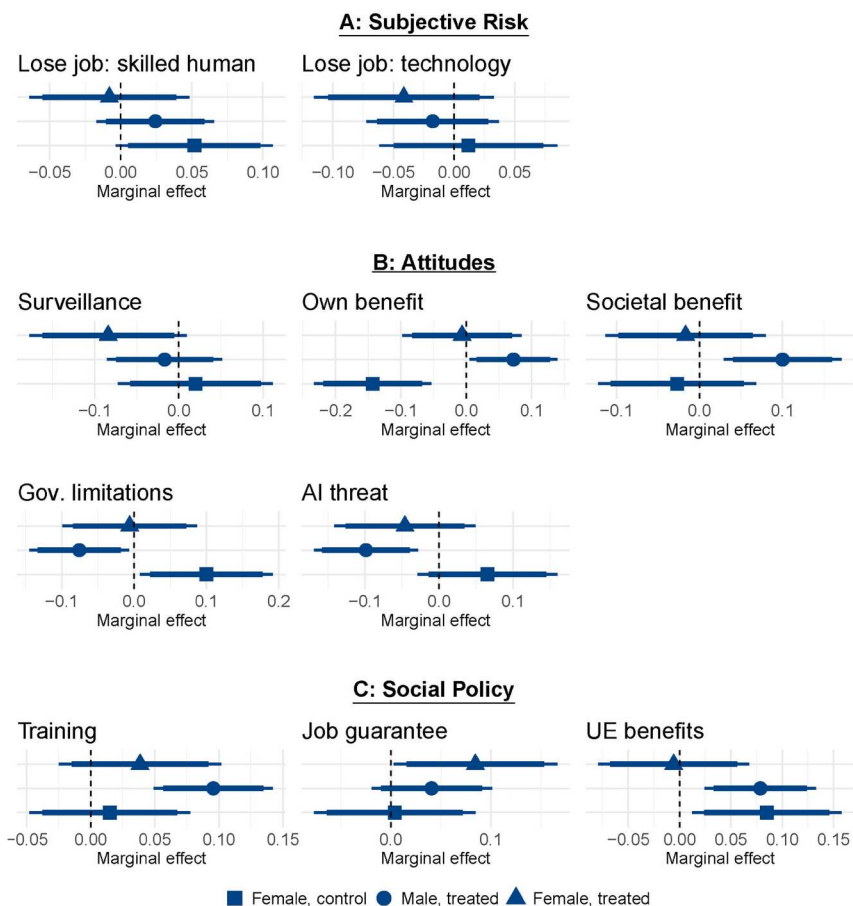


Figure 4. Heterogeneity in treatment effects by sex.

Note: Treatment effects by sex, relative to non-treated men. Estimates based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). $N = 1041$ in all models. Full output in Tables D3 and D4.

men and the effect on support for job guarantees is larger for women, but the interaction terms do not reach statistical significance. However, men in the control group are about 8 percentage points less supportive of unemployment benefits than women and in the treatment group this is exactly reversed. This is in line with the finding that women become less worried about job loss due to a lack of technology skills after being exposed to AI, while for men, if anything, the opposite is true. 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. An alternative possibility is that women and men make

different self-interested calculations upon encountering generative AI. However, here the inferred rationales contradict estimates which suggest that women are more exposed to disruption from generative AI (Felten *et al.*, 2023b; Gmyrek *et al.*, 2023). This makes clear that gender dynamics are an important aspect of the emerging political economy of AI that requires further research.

Occupational differences

We now consider heterogeneous effects by occupation. We distinguish between professional and managerial occupations (ISCO groups 1 and 2, hereafter simply ‘professionals’) and the remaining occupational groups.⁶ Existing research argues that white collar workers are likely to be most exposed to generative AI (Felten *et al.*, 2023a), threatening an occupational group that has so far been on the winning side of technological change. Hui *et al.* (2024) show that online freelance work declined both in quantity and remuneration after the introduction of ChatGPT, a finding echoed by Qiao *et al.* (2025) who argue that for every occupation an inflection point in AI capabilities exists beyond which AI replaces human workers. At the same time, workers in these potentially threatened occupations are most likely to use AI (Bick *et al.*, 2024; Haslberger *et al.*, 2025; Humlum & Vestergaard, 2025). Thus, professionals who used to be among the winners in recent automation processes might now be cross-pressured: aware of tremendous productivity benefits that people like them might enjoy, and a looming near- to medium-term threat to their jobs. Therefore, it is not obvious how the treatment should have affected them relative to non-professionals. Indeed, we find little evidence for systematic heterogeneity in the effects of AI exposure.

Panel A of Figure 5 shows that treated professionals see an insignificant reduction in subjective risk and non-professional workers experience a small increase. The interaction is marginally statistically significant ($p < 0.1$) in the case of losing one’s job to technology. In Panel B, we find no significant interactions with regard to expectations and attitudes, indicating that, even though professionals have somewhat more positive attitudes towards AI, exposure has similarly positive effects on non-professionals. When it comes to social policy preferences (Panel C), both baseline differences and interaction effects are minor, indicating that occupational background does not condition how AI affects views on social policy.

Thus, we find no evidence that professionals, who are most familiar with the types of tasks that study participants performed, were systematically differently affected by exposure to our AI treatment. However, we caution that this may be a momentary phenomenon. AI capabilities are evolving rapidly; as occupation-specific applications are developed and threats and opportunities become more clearly delineated, occupational differences

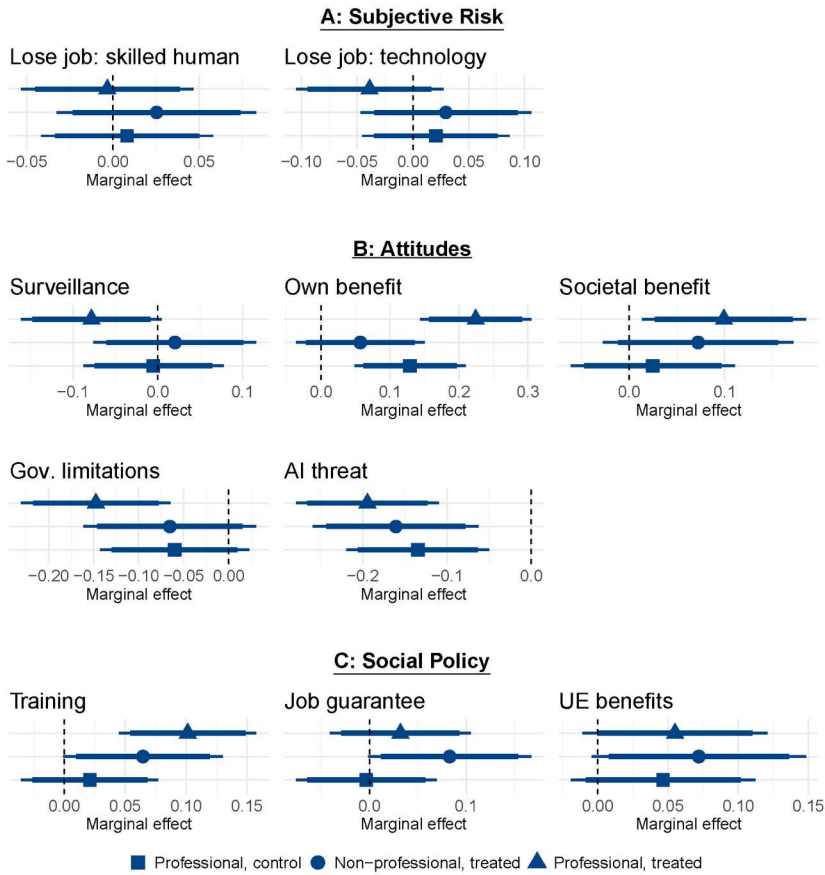


Figure 5. Heterogeneity in treatment effects by occupational group.

Note: Treatment effects by occupational group, relative to non-treated workers in non-professional occupations. Estimates based on linear probability models with weights, with 90% and 95% confidence intervals (thick and thin lines). $N = 1027$ in all models. Full output in Tables D5 and D6.

may emerge. The intuition that white collar workers are likely to be most exposed to generative AI is sound, and in light of the potentially grave political consequences of a powerful group of voters turning from winners to losers of technological change (Gallego *et al.*, 2022; Knotz, 2025), it remains imperative to monitor how people's occupations structure their thinking around AI and social policy.

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.⁷ 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 labour 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 think that AI has more benefits than drawbacks while few fear for their job. Sentiment is virtually indistinguishable in the control group and the treatment group, as shown in Figure 6. 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 regarding the implications of generative AI for 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 labour market and how they could be addressed. Other than directing their focus to the labour market, we did not guide the respondents in any way. This makes the question particularly suitable for structural topic modelling (STM). We implemented an unsupervised approach using the *stm* package 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 of ideas. The relatively small number of topics reflects the brevity and similarity of many of the open-ended responses.⁸ Panel A of Figure 7 shows the prevalence of the three topics. Topic 1, with an average topic loading of 0.32,

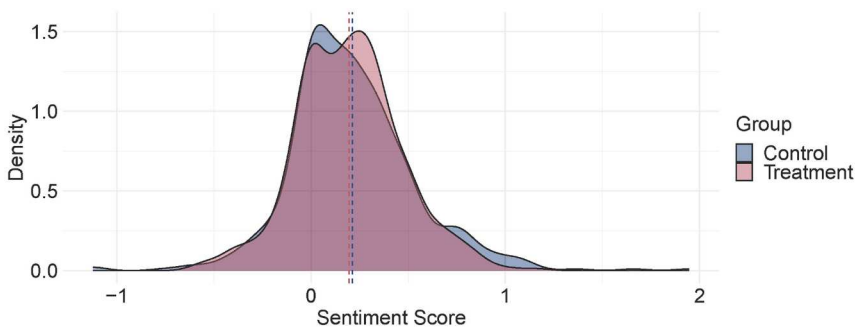


Figure 6. Sentiment scores by treatment group.

Note: Sentiment scores calculated using the *sentimentr* package. Higher scores indicate more positive sentiment.

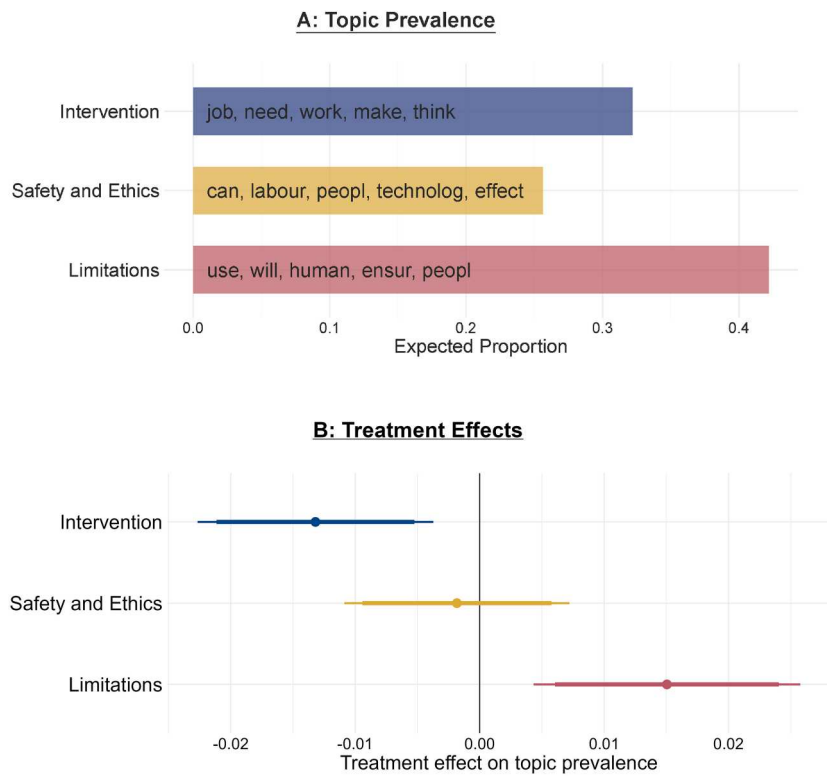


Figure 7. Structural topic model.
Note: Structural topic model estimated with the *stm* package. Full output to Panel B in Table D7.

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 0.26. 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 skepticism at its capability to replace humans on a large scale. Table E1 shows 3 exemplary responses for each of the topics.

Panel B of Figure 7 shows how the treatment affects the salience of the different topics. Compared to the control group, responses in the treatment group load significantly less on topic 1 and significantly more on topic 3, with no difference in the loadings on topic 2. Thus, treated respondents tend to talk less about the need for government intervention, for example through policies that promote skill acquisition or income support, and more about the current limitations of AI. It is at first surprising that the treatment leads to lower loadings on topic 1 – which revolves around government

intervention through social policy – given that the treatment had a positive effect on social policy preferences. However, we expect that this is due to the fact that topic loadings for each response must sum up to 1, so that for people in the control group, who did not directly interact with AI in the experiment, the need for intervention was likely relatively more salient due to its prominence in public discourse. In a model including key socio-demographic covariates (age, sex, education, occupation), we obtain essentially the same topics and estimates of the treatment effects (see Table E2). Older respondents' answers tend to focus more on safety and ethics and less on current limitations of AI, while the answers of professionals load less heavily on intervention and more strongly on limitations.

Overall, the text analysis provides complementary insights into people's reasoning about the possible consequences of generative AI. The sentiments and underlying themes we discovered by and large reinforce our conclusions from the inferential analysis. In particular, the STM analysis indicates that the treatment highlighted some of the limitations of AI and thereby alleviated concerns about job replacement, as we speculated above. This explains why treated respondents engage relatively less with topic 1 despite advocating various social policies and supports our interpretation that first-hand experience with AI 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 exposure to generative AI for subjective automation risk and political preferences. Based on a large sample of working-age adults in the UK, we document that respondents who were randomly assigned to use generative AI in three short text-based tasks did not experience greater subjective technology risk – they were no more likely to worry about losing their job to someone with better technology skills or to technology outright than the control group. At the same time, treated respondents became more positively disposed towards AI technology and more supportive of activating social policy measures. We argue that this contradicts the risk-insurance framework which underpins much of the literature on the politics of technological change and instead supports a model in which exposure to AI technology activates sociotropic preferences.

Additionally, we found meaningful heterogeneity along socio-demographic lines in baseline subjective risk and attitudes, as well as some differential treatment effects, highlighting the importance of dimensions such as gender and occupation in the emerging political economy of AI. Analyzing text data from an open-ended question, we found that amidst overall positive sentiment, treated respondents were less likely to mention job losses and the

need for government intervention and more likely to allude to the limitations of current generative AI technology, reinforcing the conclusion that AI exposure dispelled, rather than accentuated, subjective risk.

Overall, we posit that the support for our model of sociotropic preferences over the self-interest-based risk-insurance framework likely reflects the limited capabilities of AI at the time of the study together with a sense that people who are negatively affected by AI are not personally at fault, as well as uncertainty about who the relative winners and losers of widespread adoption of AI will be. We therefore reiterate that we do not interpret our findings as invalidating the risk-insurance framework as such. What we do show is that at an early stage of AI adoption, exposure to AI did not increase subjective risk as would be predicted by the risk-insurance framework. Yet, it still led to an increase in support for social policy, indicating at least some level of sociotropic motivation. However, as AI adoption continues apace, self-interest may again come to dominate people's reasoning.

In this process, the role of firms will be paramount. Firm characteristics shape how managers believe AI should be regulated and dealt with through social policy (Weisstanner & Van Kersbergen, 2025), and firms ultimately decide how to deploy AI within a given regulatory framework (Acemoglu & Johnson, 2023). The workplace experience so created reflects power relations as well as task and organizational structures, and shapes workers' political preferences (Beane & Anthony, 2024; Gingrich *et al.*, *Forthcoming*). For example, if firms prioritize the substitution of junior roles with AI as some scholars warn (Beane, 2024; Mollick, 2024), young workers in particular might demand compensation for their dwindling labour market opportunities. Thus, how AI is deployed in practice may be at least as important as the raw capabilities of frontier models.

While this article provides important and novel evidence, it has a number of limitations. First, 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. Furthermore, due to time limitations, our survey included a rather limited set of policy questions. Future work will benefit from a wider range of questions – in particular, questions that capture trade-offs (Busemeyer & Tober, 2023; Magistro *et al.*, 2024). Another possible objection concerns the nature of our experimental treatment itself. Our intention 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. The advantage of this setup is that it is more realistic and is likely to leave a stronger impression on respondents. On the negative side, it means that we have less control over what people take away from the intervention – in this case, it appears that participants found AI useful but did not make the connection to

potential job risks. While this might be construed as a shortcoming of the study design, we contend that it is itself an important insight: at current levels of capability, direct exposure to generative AI does not spark widespread concern about medium-term job replacement. Finally, we want to highlight again that generative AI is a rapidly evolving technology and hence our study illustrates effects at a particular point in time. For deeper insights, longitudinal studies which track changes in popular attitudes alongside changes in AI capabilities and deployment would be particularly useful.

To conclude, scholars interested in labour markets, redistribution and the welfare state, or voting behaviour 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 evolving capabilities and uses of AI. Generative AI use may increasingly cause subjective risk as the technology matures. Second, as long as uncertainty about the economic winners and losers of labour market changes is pervasive, theories that rely on material self-interest to predict political consequences are likely to provide limited insights. At the same time, it would be erroneous to conclude based on our study that self-interest as in the risk-insurance framework will play no role in the politics of AI. Depending on how firms choose to implement the technology, it may well regain its explanatory power as the relative winners and losers become clearer. 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 theory-building. Furthermore, our study highlights the advantages and disadvantages 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.

Notes

1. The pre-analysis plan is available under https://osf.io/dgh2u/?view_only=f30331b363c64ab883d9f76ae0565c68.
2. For details, see Appendix A.
3. 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 *et al.* 2025). Yet, when we use weights to account for such imbalances, our results remain unaffected.

4. On the other hand, the general nature of the tasks entails that they are unlikely to replicate the activities that give people meaning at work. It is possible that using AI in such tasks would evoke different reactions. However, this would have been largely outside the practical capabilities of generative AI in mid-2023 and would have required an occupationally targeted sampling strategy.
5. We report local average treatment effects (LATE) of *AI use*, obtained through a 2-stage least squares (2SLS) procedure, in Tables G9 and G10.
6. Professionals are overrepresented in our sample, reflecting early patterns of AI adoption.
7. 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. See Appendix E for details on the selection process.
8. Additional details from the STM analysis are reported in Appendix E.

Acknowledgments

We thank two anonymous reviewers and Noah Bacine, Patrick Emmenegger, Lukas Paleckis, Simone Tonelli, Bilyana Petrova, and Silja Häusermann for helpful comments on the project. Previous versions of the paper were presented at the GOVPET Advisory Board Meeting 2024, the Zurich Political Economy Seminar Series, the Freiburg Economics Research Seminar, the VPET-ECON Conference 2024, SASE 2024, CES 2024, the LCSSP AI Conference 2024, SPSSA 2025, ECPR Joint Sessions 2025, and EPSA 2025. All remaining errors are our own.

Data availability

Replication files for this article can be found at <https://doi.org/10.7910/DVN/YOSGEL>.

Generative AI

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.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

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.

Notes on contributor(s)

Matthias Haslberger is a Postdoctoral Researcher in Political Economy at the University of St. Gallen and an Associate Member of Nuffield College, Oxford.

Jane Gingrich is a Professor of Social Policy at the University of Oxford.

Jasmine Bhatia is a Senior Lecturer in Politics at Birkbeck, University of London.

References

- Acemoglu, D., & Johnson, S. (2023). *Power and progress: Our thousand-year struggle over technology and prosperity*. PublicAffairs.
- Autor, D. (2022). *The labor market impacts of technological change: From unbridled enthusiasm to qualified optimism to vast uncertainty*. NBER Working Paper Series (30074).
- Autor, D. (2024). *Applying AI to rebuild middle class jobs*. NBER Working Paper Series (32140).
- Barabas, J., & Jerit, J. (2010). Are survey experiments externally valid? *American Political Science Review*, 104(2), 226–242. <https://doi.org/10.1017/S0003055410000092>
- Beane, M. (2024). *The skill code: How to save human ability in an age of intelligent machines*. Harper Business.
- Beane, M., & Anthony, C. (2024). Inverted apprenticeship: How senior occupational members develop practical expertise and preserve their position when new technologies arrive. *Organization Science*, 35(2), 405–431. <https://doi.org/10.1287/orsc.2023.1688>
- Bicchi, N., Kuo, A., & Gallego, A. (2024). Unpacking technological risks: Different sources of concern and policy preferences. *Political Studies*, 73(3), 1054–1077. REMOVE THIS FIELD - I CANNOT DO IT.
- Bick, A., Blandin, A., & Deming, D. (2024). *The rapid adoption of generative AI*. NBER Working Paper Series (32966).
- Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3), 789–865. <https://doi.org/10.1257/jel.20160995>
- Borwein, S., Magistro, B., Loewen, P., Bonikowski, B., & Lee-Whiting, B. (2024). The gender gap in attitudes toward workplace technological change. *Socio-Economic Review*, 22(3), 993–1017. <https://doi.org/10.1093/ser/mwae004>
- Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics*, 140(2), 889–942. <https://doi.org/10.1093/qje/qjae044>
- Bürgisser, R. (2023). *Policy responses to technological change in the workplace*. <https://doi.org/10.31235/osf.io/kwxn2>
- Bussemeyer, M. R., Gandenberger, M., Knotz, C., & Tober, T. (2023). Preferred policy responses to technological change: Survey evidence from OECD countries. *Socio-Economic Review*, 21(1), 593–615. <https://doi.org/10.1093/ser/mwac015>
- Bussemeyer, M. R., & Sahm, A. H. J. (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. <https://doi.org/10.1017/S0047279421000519>
- Bussemeyer, M. R., & Tober, T. (2023). Dealing with technological change: Social policy preferences and institutional context. *Comparative Political Studies*, 56(7), 968–999. <https://doi.org/10.1177/00104140221139381>

- Carvajal, D., Franco, C., & Isaksson, S. (2024). Will artificial intelligence get in the way of achieving gender equality? *SSRN Electronic Journal*. <http://dx.doi.org/10.2139/ssrn.4759218>
- Cavaillé, C. (2023). *Fair enough? Support for redistribution in the age of inequality*. Cambridge University Press.
- Ciani, E., Freget, L., & Manfredi, T. (2021). *Learning about inequality and demand for redistribution: A meta-analysis of in-survey informational experiments*. OECD Papers on Well-being and Inequalities 02.
- Compton, M. E., & Lipsmeyer, C. S. (2019). Everybody hurts sometimes: How personal and collective insecurities shape policy preferences. *The Journal of Politics*, 81(2), 539–551. <https://doi.org/10.1086/701721>
- Dell’Acqua, F., McFowland Iii, E., Mollick, E., Lifshitz-Assaf, H., Kellogg, K. C., Rajendran, S., Kraye, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. *SSRN Electronic Journal*. <http://dx.doi.org/10.2139/ssrn.4573321>
- Dermont, C., & Weisstanner, D. (2020). Automation and the future of the welfare state: Basic income as a response to technological change? *Political Research Exchange*, 2(1), 1757387. <https://doi.org/10.1080/2474736X.2020.1757387>
- Felten, E. W., Raj, M., & Seamans, R. (2023a). How will language modelers like chatGPT affect occupations and Industries? *SSRN Electronic Journal*. <http://dx.doi.org/10.2139/ssrn.4375268>
- Felten, E. W., Raj, M., & Seamans, R. (2023b). Occupational heterogeneity in exposure to generative AI. *SSRN Electronic Journal*. <http://dx.doi.org/10.2139/ssrn.4414065>
- Ferrario, B., & Stantcheva, S. (2022). Eliciting people’s first-order concerns: Text analysis of open-ended survey questions. *AEA Papers and Proceedings*, 112, 163–169. <https://doi.org/10.1257/pandp.20221071>
- Frey, C. B., & Osborne, M. (2023). Generative AI and the future of work: A reappraisal. *Brown Journal of World Affairs*, XXX(1), 161–177.
- Funk, C. L. (2000). The dual influence of self-interest and societal interest in public opinion. *Political Research Quarterly*, 53(1), 37–62. <https://doi.org/10.1177/106591290005300102>
- Gallego, A., Kuo, A., Manzano, D., & Fernández-Albertos, J. (2022). Technological risk and policy preferences. *Comparative Political Studies*, 55(1), 60–92. <https://doi.org/10.1177/00104140211024290>
- Gallego, A., & Kurer, T. (2022). Automation, digitalization, and artificial intelligence in the workplace: Implications for political behavior. *Annual Review of Political Science*, 25(1), 463–484. <https://doi.org/10.1146/polisci.2022.25.issue-1>
- Gallego, A., Kurer, T., & Schöll, N. (2022). Neither left behind nor superstar: Ordinary winners of digitalization at the ballot box. *The Journal of Politics*, 84(1), 418–436. <https://doi.org/10.1086/714920>
- Gingrich, J., & Ansell, B. (2012). Preferences in context: Micro preferences, macro contexts, and the demand for social policy. *Comparative Political Studies*, 45(12), 1624–1654. <https://doi.org/10.1177/0010414012463904>
- Gingrich, J., & Kuo, A. (2022). Gender, technological risk, and political preferences. In M. R. Busemeyer, A. Kemmerling, P. Marx, & K. van Kersbergen (Eds.), *Digitalization and the welfare state* (pp. 157–173). Oxford University Press.
- Gingrich, J., Wu, N., & Zhang, B. (Forthcoming). Workplace governance and labor perceptions of technological risks and benefits. *Industrial and Corporate Change*.

- Gmyrek, P., Berg, J., & Bescond, D. (2023). *Generative AI and jobs: A global analysis of potential effects on job quantity and quality*. ILO Working Paper 96.
- Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4), 1091–1119. <https://doi.org/10.1257/aer.104.4.1091>
- Hakim, C. (1982). The social consequences of high unemployment. *Journal of Social Policy*, 11(4), 433–467. <https://doi.org/10.1017/S0047279400022522>
- Haslberger, M. (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. <https://doi.org/10.1016/j.rssm.2021.100623>
- Haslberger, M., Emmenegger, P., & Durazzi, N. (2023). The missing link: Technological change, dual VET, and social policy preferences. In *EPSA 2023* (Glasgow).
- Haslberger, M., Gingrich, J., & Bhatia, J. (2025). No great equalizer: Experimental evidence on productivity effects of generative AI use in the UK labor market. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4594466>
- Hui, X., Reshef, O., & Zhou, L. (2024). The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market. *Organization Science*, 35(6), 1977–1989.
- Humlum, A., & Vestergaard, E. (2025). The unequal adoption of ChatGPT exacerbates existing inequalities among workers. *Proceedings of the National Academy of Sciences*, 122(1). <https://doi.org/10.1073/pnas.2414972121>
- Im, Z. J. (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. <https://doi.org/10.1017/ics.2020.22>
- Im, Z. J., & Komp-Leukkunen, K. (2021). Automation and public support for workfare. *Journal of European Social Policy*, 31(4), 457–472. <https://doi.org/10.1177/09589287211002432>
- Jeffrey, K. (2021). Automation and the future of work: How rhetoric shapes the response in policy preferences. *Journal of Economic Behavior & Organization*, 192, 417–433. <https://doi.org/10.1016/j.jebo.2021.10.019>
- Kam, C. D., & Nam, Y. (2008). Reaching out or pulling back: Macroeconomic conditions and public support for social welfare spending. *Political Behavior*, 30(2), 223–258. <https://doi.org/10.1007/s11109-007-9048-3>
- Knotz, C. M. (2025). Technological vulnerability among the higher-educated: Implications for party preferences. *European Political Science Review*. <https://doi.org/10.1017/S1755773925000013>
- Kurer, T., & Häusermann, S. (2022). Automation risk, social policy preferences, and political participation. In M. R. Busemeyer, A. Kemmerling, P. Marx, & K. van Kersbergen (Eds.), *Digitalization and the welfare state* (pp. 139–156). Oxford University Press.
- Magistro, B., Loewen, P., Bonikowski, B., Borwein, S., & Lee-Whiting, B. (2024). Attitudes toward automation and the demand for policies addressing job loss: The effects of information about trade-offs. *Political Science Research and Methods*, 12(4), 783–798.
- Margalit, Y., & Raviv, S. (2024). The Politics of Using AI in Policy Implementation: Evidence from a Field Experiment. *SSRN Electronic Journal*. <http://dx.doi.org/10.2139/ssrn.4573250>
- Mollick, E. (2024). *Co-intelligence: Living and working with AI*. Portfolio/Penguin.
- Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187–192. <https://doi.org/10.1126/science.adh2586>

- OECD (2018). *Bridging the digital gender divide -- include, upskill, innovate*. OECD Publishing.
- Qiao, D., Rui, H., & Xiong, Q. (2025). AI and freelancers: Has the inflection point arrived? In *Proceedings of the 58th Hawaii International Conference on System Sciences*.
- Rehm, P. (2011). Social policy by popular demand. *World Politics*, 63(2), 271–299. <https://doi.org/10.1017/S0043887111000037>
- Rueda, D., & Stegmueller, D. (2016). The externalities of inequality: Fear of crime and preferences for redistribution in Western Europe. *American Journal of Political Science*, 60(2), 472–489. <https://doi.org/10.1111/ajps.2016.60.issue-2>
- Sacchi, S., Guarascio, D., & Vannutelli, S. (2020). Risk of technological unemployment and support for redistributive policies. In R. Careja, P. Emmenegger, & N. Giger (Eds.), *The European social model under pressure* (pp. 277–295). Springer VS Wiesbaden.
- Schiff, D. S., Jackson Schiff, K., & Jacobson, R. (2024). The politics of AI: Will bipartisan-ship last or is polarization inevitable? In *MPSA 2024* (Chicago).
- Thewissen, S., & Rueda, D. (2019). Automation and the welfare state: Technological change as a determinant of redistribution preferences. *Comparative Political Studies*, 52(2), 171–208. <https://doi.org/10.1177/0010414017740600>
- Weisstanner, D. (2023). Technological change and labour market policy preferences. In D. Clegg & N. Durazzi (Eds.), *Handbook of labour market policy in advanced democracies* (pp. 132–147). Edward Elgar Publishing.
- Weisstanner, D., & Van Kersbergen, K. (2025). Managing the artificial intelligence revolution: Perceived risks and social policy preferences among firm-level decision makers. *Social Policy & Administration*. <https://doi.org/10.1111/spol.13148>
- Zhang, B. (2022). No rage against the machines: Threat of automation does not change policy preferences. In *Proceedings of the 2022 AAI/ACM Conference on AI, Ethics, and Society* (pp. 856–866). ACM.