WORKING PAPER SERIES
ISSN 2785-1079

Number 32
JUNE 2026

WHAT IF YOU SEE IT? GENERATIVE AI EXPOSURE AND WORKERS' DIVERGENT POLICY PREFERENCES

Massimo Anelli

Italo Colantone

Aina Gallego

Piero Stanig

Università Bocconi
GREEN
Centre for Geography, Resources, Environment, Energy and Networks

What If You See It? Generative AI Exposure and Workers' Divergent Policy Preferences

Massimo Anelli* Italo Colantone† Aina Gallego‡ Piero Stanig§

June 14, 2026

Abstract

Generative AI is reshaping the world of labor, producing distributional consequences that may create a new structural divide between winners and losers. We study the implications of this phenomenon for workers' policy and political preferences. Specifically, we conduct a survey experiment on around 6,000 individuals employed in 98 occupations directly exposed to Generative AI, in Germany, Italy, and the US. Treated respondents are shown a video of ChatGPT performing the most frequent core task in their occupation. On average, they become more optimistic about the technology's impact, yet a substantial share becomes more pessimistic. This leads to divergent policy preferences: pessimistic workers increase their support for AI regulation, automation taxes, redistribution schemes, employment protection, and trade protectionism. Optimistic individuals move in the opposite direction. Such divergent effects are stronger for AI-related "new policies" than for traditional welfare state instruments. Pessimistic workers also express warmer attitudes toward "backlash" political parties.

*Bocconi University, Baffi Research Centre, Dondena Research Centre, and IGIER Institute for Economic Research. Email: massimo.anelli@unibocconi.it

†Bocconi University, Baffi Research Center, GREEN Research Centre, Institute for European Policy (IEP), CESIfo, and FEEM. Email: italo.colantone@unibocconi.it

‡University of Barcelona, Institut Barcelona d'Estudis Internacionals (IBEI). Email: agallego@ibei.org

§Bocconi University, Dondena Research Centre, and GREEN Research Centre. Email: piero.stanig@unibocconi.it

A large literature on routine-biased technological change shows that exposure to technological disruptions has important political consequences. Workers exposed to industrial robots vote for radical-right and populist parties (Frey, Berger and Chen, 2018; Anelli, Colantone and Stanig, 2021; Kurer, 2020a); they demand more redistribution or compensation (Thewissen and Rueda, 2019), but notably do *not* support active labor market policies (Gallego, Kuo, Manzano and Fernández-Albertos, 2022a); they oppose immigration (Wu, 2022) and demand protection from technological change (Gallego *et al.*, 2022a; Bicchi, Kuo and Gallego, 2025). At the same time, highly educated workers who have so far benefited from complementarity with Information and Communication Technologies (ICTs) are more likely to vote for mainstream political parties (Gallego, Kurer and Schöll, 2022b). This finding carries important implications for democratic stability: as Iversen and Soskice (2019) argue, democracy remains stable so long as its median voter is among the beneficiaries of economic change.

GenAI is now emerging as a general-purpose technology (GPT)—a technology characterized by broad applicability, multi-purpose capabilities, and the ability to spur complementary innovations (Bresnahan and Trajtenberg, 1995)—much like the steam engine, electricity, and semiconductors. Like previous general purpose technologies, GenAI is expected to generate large labor market transformations (Brynjolfsson, Chandar and Chen, 2025a; Brynjolfsson, Korinek and Agrawal, 2025b), with potentially profound implications for political alignments. These transformations, however, will unfold over years: we are still in the early stages of adoption, and workers are only beginning to learn how GenAI will affect their jobs. At this stage, how people *appraise* AI —as an opportunity or as a threat—likely matters more for politics than its as-yet-unrealized economic effects. A large literature shows that both perceived economic risks and aspirations of upward mobility shape policy preferences (Iversen and Soskice, 2001; Rehm, 2009; Rehm, Hacker and Schlesinger, 2012; Kurer and Van Staalduinen, 2022; Cox, 2024; Häusermann, Kurer and Zollinger, 2023; Scheiring, Serrano-Alarcón, Moise, McNamara and Stuckler, 2024; Hübscher and

Sattler, 2026; Scheve and Slaughter, 2001), suggesting that GenAI-related perceptions may similarly shape political attitudes.

Unlike ICTs, GenAI is set to automate cognitive tasks and reshape the jobs of workers with relatively high levels of education and income (Eloundou, Manning, Mishkin and Rock, 2024; Pizzinelli, Panton, Mendes Tavares, Cazzaniga and Li, 2024). This extends the reach of automation to the educated, aspirational middle classes that earlier technological waves largely rewarded. How these groups respond politically is particularly consequential given their traditional role as a stabilizing force in democratic capitalism (Iversen and Soskice, 2019; Häusermann, Pinggera, Ares and Enggist, 2022; Boix, 2019). To date, however, only limited evidence exists on how exposed workers reason about GenAI. A first strand of literature has collected descriptive data on attitudes toward the labor market impacts of GenAI (e.g., Zhang, 2024), and correlated measures of risk exposure or risk perception with policy and political preferences (e.g., Weisstanner and van Kersbergen, 2025; Busemeyer, Stutzmann and Tober, 2025; Green, Grant, Evans and Inglese, 2025; Borwein, Magistro, Alvarez, Bonikowski and Loewen, forthcoming). A notable early finding comes from Armstrong, Chen, Cuellar, Forsey-Smerek and Shah (2024): drawing on a large survey across nine advanced democracies, they report that more workers anticipate benefits from new technologies such as robots and GenAI, in terms of safety, comfort, pay, and autonomy, than anticipate costs. These observational studies offer valuable insights, but they cannot identify causal effects: the relationship between GenAI perceptions and policy preferences is likely confounded by third variables or subject to reverse causation, since socio-political orientations plausibly shape how GenAI is perceived in the first place. A second strand has begun to tackle this identification problem experimentally, exposing respondents to information about GenAI (e.g., Haslberger, Gingrich and Bhatia, 2025; Magistro, Borwein, Alvarez, Bonikowski and Loewen, 2025). These studies, however, have so far relied on generic informational stimuli, which are unlikely to shift workers' risk perceptions in their own job in a meaningful way.

In this paper, we contribute theoretical insights and empirical evidence to the emerging political economy literature on GenAI. We argue that occupational exposure to GenAI does not mechanically translate into a uniform political response: the same technology can be interpreted as either a threat or an opportunity, depending on how workers assess its implications for their jobs. We argue that workers who have pessimistic expectations react politically by demanding more redistribution and production-stage policies (e.g., AI regulation, automation taxes, codetermination), including those aimed at steering — or slowing down — technology adoption. Conversely, workers who have optimistic expectations demand more fiscally conservative and laissez-faire, technology-accelerationist policies.

To study our hypotheses, we field a pre-registered survey experiment on roughly 6,000 workers in 98 occupations directly exposed to GenAI in Germany, Italy, and the United States. Treated respondents watch an occupation-specific video of ChatGPT performing the most frequent core task in their own occupation, while control respondents receive either a placebo video or no video. The occupation-specific videos were produced via a standardized workflow: the aim was to offer respondents a very realistic experience of an interaction with ChatGPT in their work domain.[1] The treatment is intentionally non-directional: it is designed to reveal capability rather than to tell respondents what to think, so that it can push perceptions in either a pessimistic or an optimistic direction. This design allows us to identify how targeted exposure to AI capabilities affects perceived occupational impact and, most importantly, how this perception shapes policy preferences downstream. We capture the heterogeneity in reactions in a principal stratification framework (Frangakis and Rubin, 2002), classifying respondents into strata that are respectively pushed to pessimistic, pushed to optimistic, or unaffected by the treatment.

The paper makes three contributions. First, it provides experimental evidence that being exposed to information about GenAI capabilities in specific occupations causally shifts

[1] Our design and predictions were pre-registered. The anonymized PAP is included with the submission.

perceptions about workplace consequences and policy preferences across multiple domains, encompassing production-stage policies, redistribution, globalization, and political party evaluations. Our occupation-specific video treatment identifies the causal effect of GenAI perceptions on downstream preferences.

Second, the paper documents how GenAI perceptions induce divergence in policy preferences. The existing literature focuses overwhelmingly on the "losers" of technological change (Kurer, 2020b; Frey *et al.*, 2018; Anelli *et al.*, 2021), or finds convergence between perceived winners and losers on redistribution (Green *et al.*, 2025). We are able to characterize the preferences of both winners and losers, showing that experimentally-induced changes in perceptions generate divergent movements in policy preferences, with pessimistic and optimistic workers pulling in opposite directions. In the conclusions, we discuss how this pattern might complicate the formation of political coalitions for governing the distributional consequences of AI.

Third, the cross-country design, unlike single-country studies (Gallego *et al.*, 2022a; Green *et al.*, 2025; Jacobs, 2024; Haslberger *et al.*, 2025), allows us to assess to what extent AI-driven preference divergence is a general phenomenon or whether it is contingent on institutional context. The three countries in the study represent distinct welfare state regimes (Esping-Andersen, 1990; Ferrera, 1996) and varieties of capitalism (Hall and Soskice, 2001; Amable, 2003): commonality of results across these different institutional settings speaks to the generalizability of our findings to advanced capitalist democracies in general.

Theoretical considerations

Our theoretical starting point is the political economy of labor-market risk. Workers who expect greater exposure to negative economic shocks tend to support more insurance and intervention. This insight underlies a broad literature on redistribution and risk that links labor-market insecurity to policy demands (Iversen and Soskice, 2001; Rehm, 2009; Cusack, Iversen and Rehm, 2006). Recent work extends this framework to automation, where workers at risk of routine-biased technological change support redistribution more (Thewissen and Rueda, 2019) and those concerned about technology demand protection from technological change via policies that slow adoption, which they prefer over compensatory transfers (Gallego *et al.*, 2022a; Bicchi *et al.*, 2025). Throughout the paper, we distinguish two families of policies that can be adopted to respond to technological change. Redistributive policies compensate those who lose from technological shifts through transfers and welfare-state instruments — e.g., unemployment benefits, general redistribution, or the minimum wage. Production-stage policies intervene directly on the adoption of the technology and on its labor-market consequences — e.g., regulation of the technology itself, automation taxes, worker codetermination, and employment protection. Our experiment measures preferences over both families of policies.

On the side of the winners of structural change, it might be fully rational for individuals to oppose redistribution if they expect to be in a high-income group in the future, as per the "prospects of upward mobility" hypothesis (Benabou and Ok, 2001). Analogous considerations apply to preferences for production-stage policies —policies that accelerate or slow down technology adoption— rather than purely redistributive policies: workers can be expected to oppose regulation that fetters developments they expect to benefit from. Recent work on trade preferences documents how employees of firms that stand to benefit from international trade are much more supportive of trade openness and the liberal international order (Lee and Liou, 2022; Kim, 2025), expanding on an older literature suggesting that sectoral considerations shape workers' trade preferences (Hiscox, 2002; Beaulieu, 2002; Mayda and Rodrik, 2005; Blonigen, 2008; Beaulieu, Benarroch and Gaisford, 2011). Parties like the Conservatives in the UK, that talk about digitalization by highlighting business opportunities and prosperity, are better positioned to attract the winners of ICT introduction in the workplace (Gallego *et al.*, 2022b), and experimental evidence documents that pro-trade messages emphasizing potential job gains increase support for international trade (Rodríguez Chatruc, Stein, Vlaicu and Zuluaga, 2025).

We extend this framework to GenAI by focusing on perceived occupational threat and opportunity. Workers whose core occupational tasks can be performed by GenAI might infer that the returns to their education, expertise, or experience may fall. This should increase demand for policies that cushion or slow down the transition (Busemeyer, Gandenberger, Knotz and Tober, 2023; Bicchi *et al.*, 2025). At the same time, the new technology might be perceived as a complement to, rather than a substitute for, skilled cognitive labor, and some workers in these occupations might infer that the new technology, as it is complementary to their skills, is therefore productivity- or career-enhancing (Armstrong *et al.*, 2024). In summary, we expect preferences for redistributive policies and production-stage policies to be shaped by occupational self-interest.

Unlike earlier technological waves, GenAI targets tasks performed by highly educated workers in cognitive occupations, not just low-skill or routine labor. That means the relevant political mechanism is not simply class-based displacement, but a perceived loss in the value of specialized human capital. Importantly, a purely redistributive approach is often insufficient to address the consequences of structural economic change: for instance, addressing the long-term economic disruption caused by globalization would have required much more than compensatory transfers, involving instead rethinking the entire economy of trade-exposed regions (Colantone, Ottaviano and Stanig, 2022). It is therefore plausible to expect that the priority for workers who are pessimistic about the occupational impact of GenAI is not (or not only) temporary compensation or smoothing out a transition. Temporary buffers like income support would not address the long-term depletion of the returns to valuable skills and, indirectly, also the loss of social status and sense of self-worth obtained from being employed in a cognitive occupation. In addition, traditional compensation in the form of transfers may not be very credible: the "failure of compensation" that obtained with trade globalization in advanced economies is instructive of this point

(Frieden, 2019). One implication is that perceived losers of GenAI may prioritize production-stage policies over redistributive policies.

An additional relevant point is that one should not assume a symmetry in the responses of winners and losers, as suggested by prospect theory (Kahneman and Tversky, 1979). It is not implausible for the reactions to diverge, but it is also possible that perceived winners do not react, while losers become more supportive of redistribution and production-stage policies. The mechanisms shaping the policy preferences of the pessimistic and the optimistic may work differently and may need somewhat different analytical lenses to be understood: but we consider this first and foremost an empirical question.

The political-economic logic outlined above yields two central expectations. First, when workers learn that GenAI can perform the central task in their own occupation, some will update toward greater threat perception, and a pessimistic outlook, while others will update toward greater optimism. Second, changes in expectations about the impact of GenAI on their job will also change preferences for state intervention in the economy in general, encompassing both redistributive and production-stage policies. Workers who develop negative expectations about the impact of GenAI in their job will demand more compensation and policies that slow down technological development. Conversely, workers who develop more positive expectations should move in the opposite direction: if GenAI is seen as productivity- or career-enhancing, then policy interventions appear less necessary and may even be considered costly or distortionary.

To evaluate these expectations, we manipulate perceptions with an informational intervention in order to study how perceptions affect policy stances. The discussion of potential divergent reactions has important implications for the design of our experiment. The same informational shock can pull workers apart rather than pushing them uniformly against AI or in favor of regulation or redistribution, generating a divergence in policy preferences. Our design is explicitly non-directional, and the optimistic group's reaction is equally important as the reaction of perceived losers. That is a central departure from many studies of technology politics, which, with few exceptions (Gallego *et al.*, 2022b; Schöll and Kurer, 2024), tend to focus mainly on the losers of technological change. The treatment is therefore not meant to impose a single interpretation of the consequences of GenAI in the workplace, but to activate self-relevant reasoning about whether this technology complements or substitutes one's labor. This design choice is important: the experimental literature on trade and globalization shows that informational frames often work by making self-interest salient, rather than by simply transmitting facts (Hiscox, 2006; Ardanaz, Murillo and Pinto, 2013; Rho and Tomz, 2017).

An additional set of expectations concerns the type of policy that should be most affected. AI-specific regulation, taxes on automation, worker codetermination in technology adoption decisions, and employment protection are policies that directly address the risks created by technological substitution. We expect stronger effects on new, AI-specific policies than on older welfare-state instruments such as unemployment benefits or general redistribution. This, for two reasons, that are easier to separate analytically than empirically. First, as discussed above, compensation might be perceived as insufficient to fully address the consequences of labor market restructuring following the spread of GenAI. In addition, traditional welfare policies are more institutionalized and ideologically crystallized (Beramendi, Häusermann, Kitschelt and Kriesi, 2015; Gingrich and Häusermann, 2015), so short-run informational shocks may have less room to move them. By contrast, AI regulation, automation taxes, and codetermination are newer and less settled, making them more responsive to changes in perceived threat.

A final set of expectations concerns political spillovers. Previous work linking technological change to backlash voting and support for populist parties shows that technological losers may become more open to economic nationalism or radical-right appeals. We therefore expect that workers pushed toward pessimism may become more supportive of trade protectionism and immigration restrictions, and warmer toward parties that promise protection, while those pushed toward optimism may become less supportive of such policies and parties. As AI-exposed workers are often more educated and higher-income than the constituencies affected by earlier automation shocks, however, these reactions may differ in magnitude and direction from the classic "left behind" pattern (Knotz, Ugarte Montero, Lavanchy and Wagner, 2024).

In summary, our hypotheses are informed by an approach that emphasizes workers' *expectations* about the impacts of GenAI — a perspective we believe is especially relevant at the early stages of technology adoption. GenAI does not merely create winners and losers in objective economic terms; it also produces interpretive winners and losers, depending on how workers read the technology's implications for their own occupations. Some workers may infer that GenAI threatens their autonomy, job security, or career prospects; others may anticipate complementarity, productivity gains, or improved work quality. These divergent perceptions should translate into opposing preferences over production-stage policies and redistribution, with likely spillovers on broader orientations — including trade protectionism — and on political alignments more generally. The central theoretical claim of the paper is therefore that, at these initial stages, individual exposure to GenAI generates divergence among workers in AI-exposed occupations rather than a uniform leftward or rightward shift. What we may be observing, then, is the emergence of a new structural divide — one that, if perceptions harden as the technology matures and political entrepreneurs mobilize, could consolidate into a durable political cleavage.

Research Design

We carry out a pre-registered survey experiment on around 6,000 individuals working in GenAI-exposed occupations in three countries: Germany, Italy, and the United States. The experiment is designed to shift workers' perceptions of GenAI's occupational impact through occupation-specific video demonstrations, and then trace how the resulting changes in perceptions translate into policy and political preferences. By manipulating workers' perceptions in an experimental setting, we aim to detect causal effects of perceived GenAI-related occupational material interests on policy and political preferences. In this section, we first explain how GenAI-exposed occupations have been selected, and provide descriptive information on the resulting sample of workers. Next, we explain how we measure GenAI's occupational impact perceptions, and present some descriptive evidence on these perceptions at baseline. We then illustrate how the experimental video stimuli were conceived and produced. Finally, we present the experimental design.

Sample

Our sample includes workers employed in 98 narrowly-defined occupations that are identified as "directly exposed" to GenAI. Specifically, these occupations are selected at the 6-digit level of disaggregation of the Standard Occupational Classification (SOC). For the selection, we follow three steps.[2] First, for all 3-digit SOC categories, we obtain occupation-specific scores of exposure to AI from three earlier studies: Webb (2019), Tolan, Pesole, Martínez-Plumed, Fernández-Macías, Hernández-Orallo and Gómez (2021), and Felten, Raj and Seamans (2021). We then conduct a k-means cluster analysis with $k = 2$, on each measure separately, to classify occupations into two groups: exposed or not exposed to AI. We thus obtain three different binary classifications of occupations. We select all the 3-digit occupations classified as exposed to AI according to at least two out of three classifications. This step results in the selection of 39 3-digit SOC occupations (out of a total of 94). They are reported in the pre-analysis plan and in the replication materials.

As a second step, we focus on all the 369 6-digit SOC occupations within the selected 39 3-digit occupations. Among them, we identify as directly exposed to GenAI those where *the most frequent core task* is automatable through ChatGPT. The most frequent core task for each occupation is identified from O*NET data, that are based on a combination of expert evaluations and worker surveys. The evaluation as to whether a given task can

[2] These are described in full details in SI Section S1.

be directly automated through ChatGPT is performed by ChatGPT itself, based on the prompt developed by Eloundou *et al.* (2024). Specifically, direct exposure requires that ChatGPT can reduce the time it takes to complete a task with equivalent quality by at least 50%. This step leaves us with a total of 176 occupations classified as directly exposed to GenAI. These account for around 16.3% of total employment in the US, based on 2022 BLS data. We finally select 98 out of these 176 occupations by including all the 82 occupations representing at least 0.03% of US employment, plus 16 occupations that were needed in order to ensure representativeness also in the German and Italian panels of the survey company (YouGov). The list of all 6-digit SOC occupations included in the sample is provided in the pre-analysis plan in and in the replication materials.

We select into the sample only respondents who are currently employed in one of the selected occupations at the time of the survey. Descriptive information on the sample is provided in SI Section S4. Consistent with the nature of GenAI-exposed occupations, our sample consists of workers who tend to be younger, higher-income, and substantially more educated than the overall population (see Figure S1). The political profile of the sample mirrors this demographic skew. As shown in Tables S3-S5, relative to the general population, our respondents are more likely to turn out in national elections; they are less likely to support radical-right parties (e.g., FdI and Lega in Italy, Republicans in the US, though not AfD in Germany); and they are more likely to support progressive alternatives (e.g., Greens and Die Linke in Germany, PD in Italy, and Democrats in the US).

Measuring GenAI's impact perceptions

We measure the perceived occupational impact of GenAI via a composite index, AIP, that is based on the answers to several items asked both to treated and control respondents. Higher values indicate more optimistic perceptions; lower values indicate more pessimistic perceptions. Specifically, the AIP index is the simple average of three items. The first is an index of perceptions of overall work impact, itself obtained as a simple average over five sub-items: salary and career opportunities, job security, autonomy, enjoyment of daily tasks, and job pride. Answers are given on a 5-point scale, from "mostly negative" (-2) to "mostly positive" ($+2$). The average in the control group is 0.28. The second item gives respondents three options regarding how GenAI will impact workers like them: make them redundant (-2); have no impact (0); or make them more productive and better paid ($+2$). In the control group, 29% select more productive and 23% redundant, with the remainder—a plurality—choosing no impact, a notable stance given that the sample targets directly exposed occupations. The third item asks respondents how they perceive GenAI on a 5-point scale that goes from "a large threat" (-2) to "a large opportunity" ($+2$). In the control group, 35% of respondents perceive it as an opportunity, 29% as a threat, while the remaining plurality chooses the neutral option.[3]

We begin by documenting how baseline perceptions vary across workers, regressing the AIP index on pre-treatment covariates among control respondents (SI Tables S6–S9). Standard demographics are largely uninformative: baseline AIP is uncorrelated with age, gender, and education. What matters more is economic position and exposure to the technology—higher-income workers, and those more aware of and experienced with GenAI, are systematically more optimistic, while the less familiar are more pessimistic. Occupational experience and content also shape perceptions: workers who have suffered job loss or unemployment spells, and those in jobs with little interpersonal contact, are more pessimistic, whereas those whose work demands logic and problem-solving are more optimistic. Workers in coding- and data-intensive occupations are among the most optimistic, and perceptions vary little across industries. Personality is also predictive: risk tolerance, agreeableness, conscientiousness, and emotional independence are associated with greater optimism, whereas neuroticism and extraversion predict greater pessimism.

[3] For presentational consistency throughout the paper, the AIP index and all its components are coded so that positive values indicate optimistic perceptions; this reverses the sign convention used in the pre-registered analysis plan but leaves all estimated effects, standard errors, and inferences unchanged.

Experimental Stimuli

The video stimuli aim at illustrating the potential of GenAI in a given job context. For each of the selected occupations, we produced, via a standardized workflow, a video showing ChatGPT performing the most frequent core task specific to the occupation. For each occupation we configured a custom GPT ("My GPT"), trained with extensive information about the content and nature of the job (e.g., from O*NET), and we asked it to conceive a hypothetical scenario in which ChatGPT would perform the most frequent core task of that occupation. We allowed ChatGPT to choose which capabilities to activate when performing the task—i.e., Web Search, Image Generation, or Code Interpreter and Data Analysis. We then asked ChatGPT to formulate an optimal prompt, intended for submission to ChatGPT itself; we submitted it to a different instance; and recorded the session. A step-by-step description of the whole procedure is provided in SI Section S2. Videos were created in English and translated into German and Italian, with synthesized voice-overs (randomizing male and female voices). Each video lasts approximately 1.5 minutes. For instance, following up on the previous examples, the videos shown to "Accountants and Auditors" and to "Credit Analysts" are available at the following links: Accountants and Auditors; Credit Analysts.[4]

By design, the treatment videos are occupation-specific: each video shows ChatGPT performing the most frequent core task in the respondent's own occupation. This tailoring is a key feature of the experiment, since it ensures that each respondent sees a stimulus relevant to their own job. At the same time, randomization ensures that we can estimate the average treatment effects of being exposed to a video created following the standardized procedure we described. Yet, one could be concerned that the videos might not be comparable across occupations—for instance, respondents might perform the shown task more frequently in some occupations than in others, or find it more or less enjoyable or

[4]All videos are stored in an on-line folder available at this link.

important. SI Table S2 shows that occupation fixed effects explain little of the variation in respondents' evaluations of the task shown (frequency, enjoyment, importance) or of the video's effectiveness, supporting the interpretation that the videos function as equivalent stimuli.

Experimental Design

Our design is block-randomized where the 6-digit occupations are the blocks: within each of the 98 occupations, respondents are assigned to the treatment condition—i.e., they watch a video displaying ChatGPT performing a task—with 50% probability. Within the control condition, respondents are randomly assigned, with equal probability, to one of two control arms: in one, respondents do not receive any video stimulus; in the other, respondents watch a video (common to all respondents in this arm) of approximately the same length, referring to technologies for building bridges. This control video is neutral and discusses a technology with no direct connections with the labor market. It is available at this link.

Our preferred empirical approach relies on pooling all controls into a single control group. Following a pre-registered procedure, we ensure that there are no statistically significant differences across the two control arms in terms of perceptions of GenAI's occupational impact by estimating the ATE of receiving the control video (vs. no video) on AIP, based on the specification outlined in Equation 1. The results are reported in Table S1. In the absence of consistent statistically detectable effects of the control video (compared to no video), we proceed throughout the analysis without differentiating the two control arms.

Estimation

Treatment effects on GenAI's occupational impact perceptions

Our non-directional informational treatment is designed to elicit naturalistic reactions regarding the effects of GenAI, that could go in both optimistic and pessimistic directions. As a preliminary piece of evidence, Figure 1 shows estimates of the average treatment effect on the perceived occupational impact of GenAI (AIP). To estimate the average treatment effect on the perceived occupational impact of GenAI, we estimate models of the form

$$Y_i = \alpha + \tau T_i + \sum_j \beta_j B_{j(i)} + \sum_j \gamma_j T_i (B_{j(i)} - \overline{B}_j) \tag{1}$$

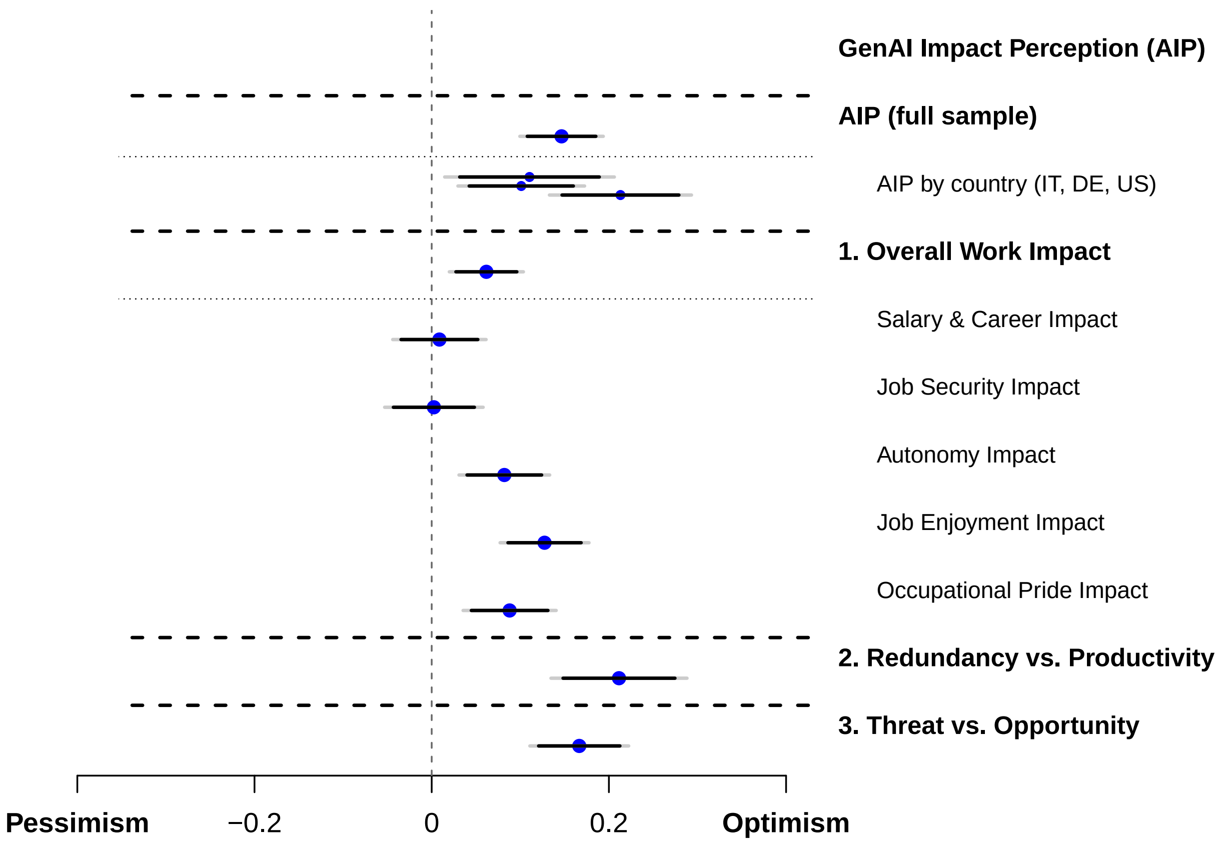
where T_i is the treatment indicator, B_j is an indicator for belonging to block j, and $\overline{B}_j$ is the mean of the block indicator—i.e., the proportion of observations in the block.[5] Such specification includes block fixed effects; by re-centering the interactions, the model conveniently retrieves the familiar difference-in-means within each block, averaged across blocks by block size, through the single parameter τ (Imbens and Wooldridge, 2009; Imbens, 2011).

The average treatment effect on the overall AIP index, estimated on the full sample, is positive and statistically different from zero: the treatment improves perceptions of the occupational impact of GenAI. The magnitude is about one-seventh of a point on the -2 to $+2$ scale (roughly one-sixth of a standard deviation). The effect is positive and significant also when estimated separately for each of the three sample countries. The point estimate is slightly larger in the US than in Germany and Italy.

The remaining coefficients come from regressions where the outcome variable is either one of the three items that make up the overall AIP index, or one of the five sub-items

[5] We follow pre-registered rules to aggregate the small number of observations in insufficiently numerous blocks into a residual block. More detail is provided in SI subsection S5.

Figure 1: **Treatment Effects on GenAI Impact Perceptions.** Tables reporting the complete set of estimates underlying this plot can be found in SI Section S9.



of the overall work impact item. Importantly, the analysis yields positive and significant effects for each of the three main items, with a stronger effect for the job redundancy item than for the others. As for the five dimensions of overall work impact, the improvement in perceptions is driven by stances on work autonomy, job enjoyment, and pride—for which we find positive and significant effects—while no statistically discernible effects are found on salary, career and job security impact perceptions. In SI Fig. S3, we show that the main result on the overall AIP index is stable across three groups of respondents defined by prior use of Large Language Models (LLMs) such as ChatGPT: frequent, occasional, none. Some degree of heterogeneity across groups emerges only when considering the estimated effects on specific items.

Impact perceptions, policy and political preferences

Stratification

As we mentioned above, our treatment is deliberately non-directional, allowing for responses in the direction of increased optimism, or increased pessimism, regarding the consequences of GenAI adoption in the workplace. The overall positive average treatment effect on perceptions reported in the previous subsection masks any heterogeneity between respondents that became more optimistic, respondents that became more pessimistic, and those whose perceptions were unaffected by the treatment. Accounting for such heterogeneity is key as we move to the central research question of the paper: how perceptions of the occupational impact of GenAI influence policy and political preferences. To address this question, we work in a principal stratification framework (Frangakis and Rubin, 2002). Specifically, we classify respondents in one of four strata: the "pushed to pessimistic" stratum, i.e., those whose perceptions are worsened; the "pushed to optimistic" stratum, i.e., those whose perceptions are improved; the "always pessimistic" and the "always optimistic" strata, that include respondents whose perceptions remain roughly unchanged. The hypotheses are stratum-specific, i.e., the expected effects are conditional on the change in GenAI impact perception induced by the treatment. In practice, as per the pre-registration document, we estimate stratum-specific treatment effects.

Assigning individuals to a stratum requires two pieces of information: their perceived occupational impact of GenAI (AIP) in the absence of the treatment, and their AIP after receiving the treatment. Yet, only one of these is observed for any given respondent. Given random assignment, treated and control units are identical in expectation. We can therefore fit separate prediction models—based only on pre-determined variables—in the control and treatment groups, and use the predictions as prognostic scores of the potential outcomes of AIP. In practice, we predict AIP under control for all respondents based on a random forest estimated only on control respondents, and AIP under treatment for all respondents based on a random forest estimated only on treated respondents.[6] Relying exclusively on predicted AIP values, we then classify as "pushed to pessimistic" individuals predicted to be above the median AIP under control but below the median under treatment, and as "pushed to optimistic" those predicted to be below the median under control but above the median under treatment. Respondents predicted to stay below (or above) the median in both cases are placed in the two "unchanged" (disassociative) strata.[7] Full details on the random forest implementation and robustness checks incorporating uncertainty regarding stratum membership are reported in the SI section S7.

The stratification we propose is always based on prognostic scores constructed exclusively from pre-treatment covariates. It is important to stress that conditioning on this approximation to stratum membership defined by these prognostic scores does not introduce post-treatment bias. In fact, post-treatment bias arises when one conditions on the

[6]Technically, we estimate approximations to two conditional expectation functions for the potential outcomes of GenAI impact perception (AIP): $\mathbb{E}[\mathrm{AIP}(0) \mid X]$ under control and $\mathbb{E}[\mathrm{AIP}(1) \mid X]$ under treatment. Both models use only pre-determined covariates X—such as occupation, income, education, age, gender, and psychological traits (risk aversion, attachment style, big-5). Hence the predicted GenAI impact perceptions are prognostic scores that depend exclusively on pre-treatment variables.

[7]Summary statistics of pre-treatment covariates by stratum are reported in SI Table S10.

observed value of a mediator (AIP, in our case). By contrast, our approach conditions on (an approximation to) the pair of potential outcomes for AIP—that is, on stratum membership itself. These potential outcomes, and thus stratum membership, are not affected by treatment. Therefore, any comparison within a principal stratum (i.e., conditional on stratum membership) has a causal interpretation when, as in our case, treatment assignment is random and thus orthogonal to the potential outcomes (Frangakis and Rubin, 2002).

Effects on policy preferences

After classifying respondents in one of four strata, we estimate equations of the form:

$$Y_i = \alpha_{s(i)} + \tau_{s(i)} T_i + \sum_j \beta_j B_{j(i)} + \sum_j \gamma_{s(i)j} T_i (B_{j(i)} - \overline{B}_j), \tag{2}$$

where the average treatment effect for stratum s is given by τ_s.

Figure 2 displays the stratum-specific average treatment effects across the full set of policy preferences we examine, for the strata "pushed to optimistic" and "pushed to pessimistic" (optimistic and pessimistic, henceforth). These effects are estimated through regression models with occupation fixed effects (Equation 2). For presentational purposes, all outcomes are standardized to have standard deviation equal to one.

The first set of items concerns traditional welfare state policy instruments. Specifically, we focus on support for: government responsibility to reduce income differences through taxation and redistribution[8]; specific redistributive instruments such as early retirement, unemployment benefits, and active labor market policies[9]; a higher minimum wage[10]; and a tax on the "super-rich"[11]. We expected a positive treatment effect on support for

[8] In the control group, the modal response to the general redistribution item is "somewhat agree", indicating a generally progressive outlook.

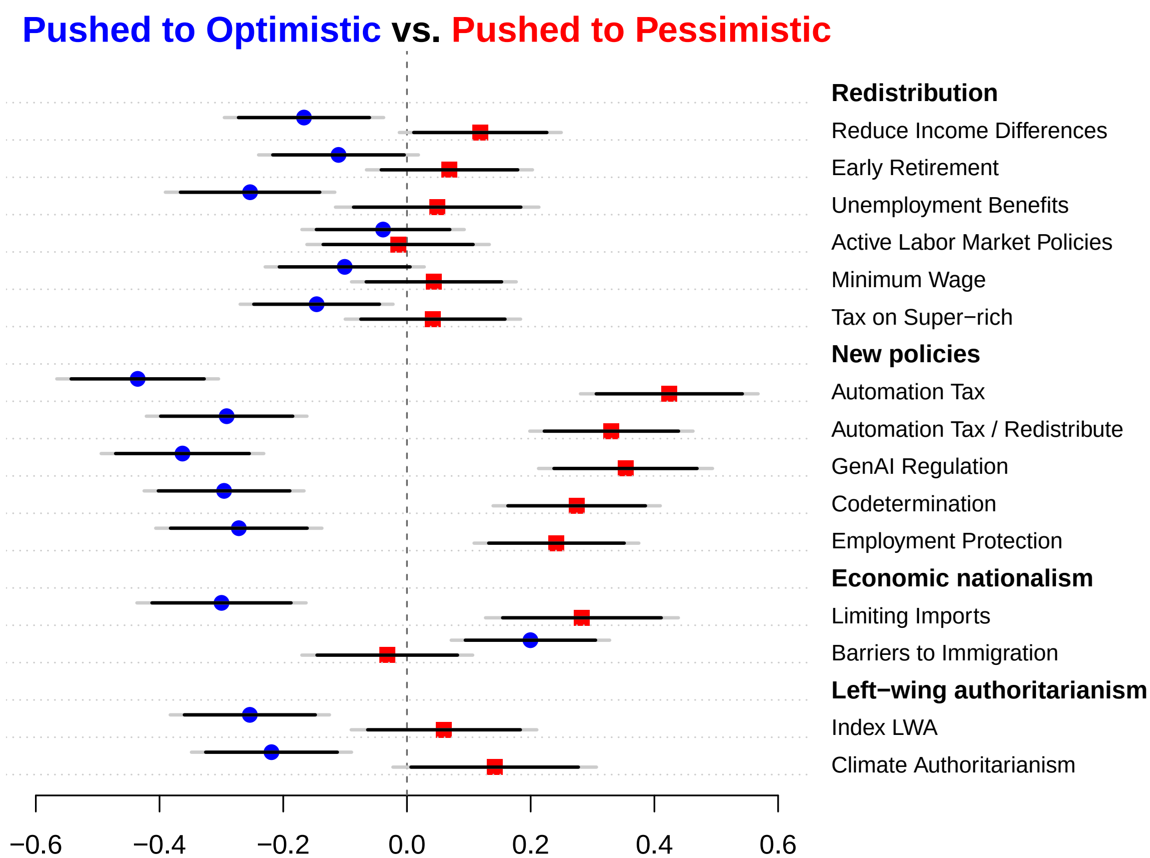
[9] Survey questions on these instruments are introduced by the warning that "higher taxes may be needed to pay for them".

[10] We tailor the minimum wage item to each national context.

[11] This item randomizes: the level at which the tax is implemented (national vs. international); the tax rate (2% vs. 10%); and the wealth threshold for identifying the super-rich (10 million vs. 1 billion €/$ wealth). We pool across all versions of the question, regardless of the exact text seen by the respondent. Average

redistribution items in the pessimistic stratum, and a negative effect in the optimistic stratum. In line with the pre-registered hypotheses, the optimistic become less supportive of government role in reducing income differences, of unemployment benefits, and of a tax on the super-rich. In addition, there is weak evidence ($p < 0.1$) in support of the hypothesis that the pessimistic become more supportive of government's general role in taxation and redistribution, while no significant effects are detected for the specific policies we ask about.[12]

Figure 2: **Treatment Effects on Policy Preferences by Stratum**. Standardized coefficients. Tables reporting the complete set of unstandardized estimates for this plot can be found in SI Section S9.



support in the control group is 7.37 on a 0-10 scale.

[12]The pre-registered heterogeneity analyses by income, education, and age yield no statistically discernible results. At the same time, there is some heterogeneity driven by ideology and past party choice. For instance, the optimistic who supported a left-wing party do not become more opposed to the tax on the super-rich when treated. All results from the heterogeneity analysis are reported in SI Section S10.

Moving down the plot, we examine support for a number of "new policies" that are more directly related to managing the distributional consequences of technological progress (Abbott and Bogenschneider, 2018; Bradford, 2023; Johnson and Acemoglu, 2023). First, we assess support for increasing taxes on firms that adopt robots and algorithms to replace workers.[13] Second, we consider support for a policy bundle that links an automation tax to specific redistribution measures, such as retraining displaced workers.[14] Third, we elicit views on government regulation of GenAI, on a scale from "not intervene" (-2) to "regulate most strictly" ($+2$). Fourth, we measure agreement with codetermination legislation granting workers' representatives a voice in company decisions concerning the introduction of new technologies. Finally, we gauge preferences regarding employment protection for workers who become redundant, ranging from "businesses should have freedom on lay-off decisions" (0) to "governments should find ways to prohibit such lay-offs" (10).[15] In general, we expect pessimistic workers to show increased support for these policies, whereas optimistic workers are expected to become less supportive. The results are consistent with the pre-registered hypotheses: all the estimated treatment effects for the pessimistic stratum are positive and significant, while they are all negative and significant for the optimistic stratum. These findings indicate that policy preferences diverge based on the perceived occupational impact of GenAI: the pessimistic demand more state intervention, codetermination, and protection of workers, while the optimistic move in the opposite direction.

The results are visibly stronger for new policies than for the traditional welfare state items considered above. This is probably unsurprising—especially for the type of workers we study—given the long-standing salience and institutional inertia of welfare state policies.

[13] In the control group, the modal category is weak support.

[14] This item randomizes both the revenue collection strategy and the spending destination. We pool across all versions of the question, regardless of the exact text seen by the respondent. Average support in the control group is 6.44 on a 0–10 scale.

[15] This item randomizes the cause of lay-offs, including workers' substitution with robots and AI. We pool across all versions of the question, regardless of the exact text seen by the respondent. Average support in the control group is 5.70 on a 0-10 scale.

Conversely, since new policies are directly linked to technological change—and preferences toward them are arguably not yet crystallized—our informational treatment may be particularly effective at shaping preferences in these increasingly salient policy domains. Interestingly, in the heterogeneity analysis reported in SI Section S10, we find that in the pessimistic stratum the positive treatment effect is similar across left and right identifiers. Conversely, in the optimistic stratum the negative effect is stronger for respondents who do not identify as left wing, and for men.

Next, we examine treatment effects on attitudes toward international trade and immigration. Specifically, we measure support for placing new limits on imports to protect domestic jobs (explicitly noting that such limits could raise consumer prices and hurt domestic exports)[16], and we gauge preferences over whether the number of immigrants allowed in one's country should be raised, kept constant, or reduced.[17] Losers of structural economic changes—such as globalization and robotization—tend to shift their support toward parties and candidates advocating economic nationalist platforms (e.g. Colantone and Stanig, 2018) that combine import tariffs and trade restrictions with an exclusionary approach to immigration, and economic grievances rooted in import competition and automation tilt citizens' attitudes in an authoritarian, nationalist, and nativist direction (Agnolin, Colantone and Stanig, 2025; Anelli *et al.*, 2021; Ballard-Rosa, Malik, Rickard and Scheve, 2021; Ballard-Rosa, Jensen and Scheve, 2022; Carreras, Carreras and Bowler, 2019; Ferrara, 2022; Hays, Lim and Spoon, 2019; Gamez-Djokic and Waytz, 2020). This "cultural" shift has played an important role in pushing the constituencies of the losers of structural changes closer to economic nationalist parties, especially of the radical right (Ferrari, Franzese, Jackson, Wai, Wu, Kim, Kim, Pollack and Sanders, 2021; Franzese, 2019). We hypothesize that the treatment increases support for protectionism and barriers to

[16] In the control group, the median response is neutral in all three countries; support for restrictions is higher in the United States (about 42% somewhat/strongly support) than in Germany and Italy (both below 30%).

[17] In the control group, the median response is 0 on a scale from -2 ("greatly increased") to 2 ("greatly decreased").

immigration among respondents in the pessimistic stratum.

We find support for the first hypothesis, but not for the second. This may be explained by the fact that workers exposed to GenAI tend to be highly educated individuals employed in cognitive or "socio-cultural" occupations, who tend to hold more progressive views (Simón and Claveria, 2024). Their profile differs from that of the losers of globalization and earlier technological shifts, such as robotization. These losers more often belonged to less educated, blue-collar constituencies potentially more prone to exclusionary nationalism and nativism. This difference in profiles, already emphasized by Eloundou *et al.* (2024), provides additional motivation for our investigation in this domain. In the optimistic stratum, treated respondents become less supportive of trade protectionism but, at the same time, more supportive of restrictions to immigration. In the heterogeneity analysis reported in SI Section S10, we show that the latter finding is driven by right-leaning respondents and by those with higher social dominance orientation. Ideological heterogeneity is pronounced also when considering international trade preferences: in both strata, the effect on protectionism is stronger in absolute value and in terms of statistical significance among voters of conservative parties and those who do not self-identify as left of center.

Given the higher-education profile of workers exposed to GenAI, it is worth investigating also whether GenAI perceptions are related to preferences regarding left-wing authoritarian policies. Specifically, we consider an index of left-wing authoritarianism computed as the average of three items: (1) how much authority should the government have to enforce strict environmental regulations, even if it means overriding individual freedoms and business interests[18]; (2) support for laws that allow surveilling, restricting, or banning political parties that put forward anti-democratic ideologies[19]; and (3) support for laws that require schools to teach children about one's country history with racism, sexism and

[18] Measured on a 4-point scale from "minimal" to "full" authority, then rescaled to vary between 0 and 1. The median response in the control group is "significant authority".

[19] Measured on a 5-point scale from "strongly oppose" to "strongly support", then rescaled to vary between 0 and 1. The median response in the control group is 0.75.

homophobia.[20] The estimated treatment effect is negative and significant for the optimistic stratum, and positive but not significant for the pessimistic stratum. The latter result indicates that worsening perceptions of GenAI's occupational impact do not translate into more left-wing authoritarian attitudes. A positive and marginally significant ($p < 0.1$) effect among the pessimistic emerges only when the climate item is considered in isolation, as shown at the bottom of the plot. Yet this latter result is not robust to incorporating uncertainty regarding stratum membership—whether by removing observations very close to stratification thresholds or by bootstrapping the random forest predictions. These robustness checks are conducted for all results presented in Figure 2, and are reported in SI Fig. S2.

As reported in SI Section S9, we also find some modest but significant effects for the unchanged strata. Specifically, in the "always pessimistic" stratum, the treatment increases support for GenAI regulation and workers' codetermination, whereas in the "always optimistic" stratum it reduces support for redistribution and climate authoritarianism. These results—albeit small in magnitude and limited to these items—indicate that the videos may influence policy preferences even without a large change in perceptions of GenAI's occupational impact, for instance by raising the salience of this technological shift.

Effects on political preferences

Finally, we investigate treatment effects on political preferences. We use respondents' ratings of their warmth toward specific political parties, measured using feeling thermometers ranging from 0 ("as cold and negative as possible") to 100 ("as warm and positive as possible"). We focus on the average warmth toward two groups of parties: "backlash" and "non-backlash". Backlash parties are either radical right or radical and populist left parties; non-backlash parties are all the others. Prominent examples of backlash parties

[20] Measured on a 5-point scale from "strongly oppose" to "strongly support", then rescaled to vary between 0 and 1. Overall, the control-group mean of the left-wing authoritarianism index is approximately 0.6.

are: Afd and Die Linke in Germany; Brothers of Italy, Lega, and the Five Star Movement in Italy; and the Republican party in the US.[21] These parties share anti-globalization and anti-establishment positions, combined with domestic economic stances that span both the left and right of the political spectrum.

Earlier research shows that the losers of automation tend to increase their support for both radical right and radical left backlash parties (Anelli *et al.*, 2021; Milner, 2021). In line with that, we make two hypotheses: (1) the treatment increases warmth towards backlash parties among pessimistic respondents who were voters of backlash parties in the last election; and (2) the treatment decreases warmth toward non-backlash parties, and increases warmth toward backlash parties, among pessimistic respondents who were voters of non-backlash parties in the last election. Given that we split respondents based on the party grouping they supported in the last election, instead of stratifying by s, we account for heterogeneity via a (pre-registered) model with the continuous interaction between the treatment indicator and $\hat{\mu}_i$, the predicted value of the answer to the direct question about the reaction to the video. The procedure used to predict $\hat{\mu}_i$ is analogous to the one adopted for the approximations to the potential outcomes of AIP and is described in SI S7.

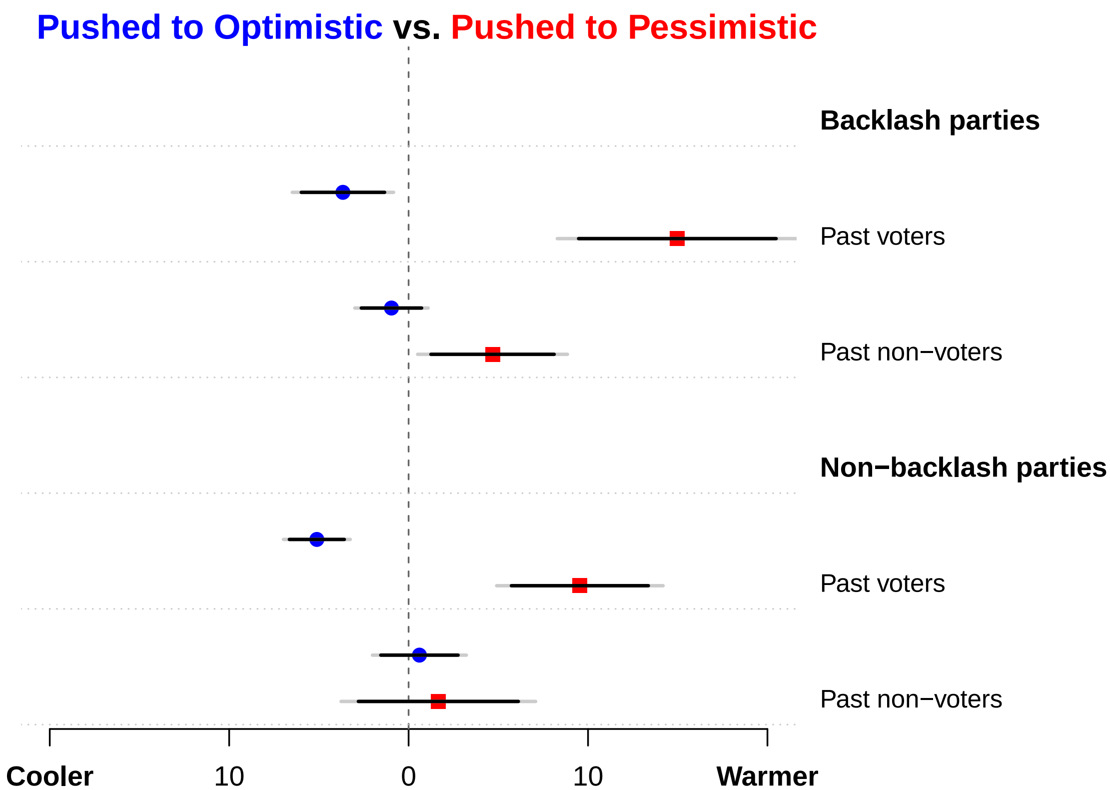
We estimate equations of the form

$$\begin{aligned} Y_i = \alpha + \tau T_i + \delta \hat{\mu}_i T_i + \sum_j \beta_j B_{j(i)} + \sum_j \gamma_j T_i (B_{j(i)} - \overline{B}_j) \\ + \sum_j \xi_j \hat{\mu}_i T_i (B_{j(i)} - \overline{B}_j) + \theta \hat{\mu}_i \end{aligned} \tag{3}$$

The estimated treatment effects shown in Figure 3 are evaluated as $\tau + \delta \hat{\mu}_\phi$ at values $\hat{\mu}_\phi \in \{-1, 1\}$. The scale of $\hat{\mu}$ is the same as the survey item based on which it is predicted, ranging from -2 ("much more concerned") to $+2$ ("much more optimistic"). Hence evaluated effects with $\hat{\mu}_\phi = 1$ are illustrative of individuals made optimistic by the treatment and those with $\hat{\mu}_\phi = -1$ of the pessimistic.

[21] The (pre-registered) complete list includes also Alleanza Verdi e Sinistra, Unione Popolare, and Italexit in Italy.

Figure 3: **Treatment Effects on Warmth towards Backlash and Non-Backlash Parties, by Stratum and Past Vote**. A table reporting the complete set of estimates underlying this plot can be found in SI Section S11. The estimated treatment effect is evaluated at value $\hat{\mu}_\phi = -1$ for individuals pushed to pessimistic, and at value $\hat{\mu}_\phi = 1$ for individuals pushed to optimistic.



Results are reported in Figure 3. In line with the pre-registered hypotheses, among the pessimistic the treatment raises warmth towards backlash parties, both among past supporters (by around 15 points on the 0-100 scale) and among other voters (by around 5 points). In contrast with expectations, when treated, past voters of non-backlash parties among the pessimistic also increase their warmth toward the party grouping they chose in the last election. In general, optimistic respondents slightly reduce their warmth toward the party grouping they supported in the last elections. In sum, these results suggest that GenAI's impact perceptions may have meaningful consequences for voting behavior.

Conclusion

Our experiment is the first to expose workers to an occupation-specific demonstration of GenAI performing their most frequent core task, providing causal evidence on how perceptions of this technology shape policy and political preferences. Our survey experiment shows that, when confronted with concrete demonstrations of GenAI capabilities applied to their own jobs, workers become on average more optimistic about the occupational impact of this technology. Yet there is substantial heterogeneity in reactions, which translates into diverging policy preferences. Pessimistic workers demand stronger government intervention through regulation and targeted redistribution measures, worker protection and codetermination, and trade protectionism. Optimistic workers move in the opposite, laissez-faire, direction.

The divergence of exposed workers between self-perceived "winners" and "losers" of GenAI has important social and political implications. In particular, it implies that the distributional consequences of GenAI are unlikely to generate a straightforward majority in favor of any coherent policy bundle that combines technology diffusion with the redistribution of its welfare gains. In other words, this divergence undermines the possibility of a broad social contract for governing the structural consequences of GenAI. Acemoglu and Restrepo (2019) and Johnson and Acemoglu (2023) argue that a promising response to rising inequalities between winners and losers of technology lies in institutional arrangements that redirect technological change toward complementing workers and generating broad shared gains, rather than prioritizing labor-cost savings and task automation. Such arrangements, however, require the construction of countervailing powers—through regulation, taxation, and workers' codetermination. Our evidence highlights why building these institutions may prove politically arduous: precisely those who feel they may benefit from GenAI become less supportive of policies that could discipline or reshape the trajectory of adoption, while those who feel pessimistic demand stronger intervention. As a result, there is little prospect for broad political coalitions around comprehensive policy packages designed to harness the technology's potential while mitigating its risks and spreading its benefits.

Even within each group of exposed workers—i.e., the optimistic or the pessimistic—the reactions triggered by GenAI in terms of policy preferences give rise to complex combinations of political demands that are not obviously met by the policy bundles currently supplied by political parties. For instance, optimistic workers demand at the same time less government intervention, lower taxation and redistribution, and fewer restrictions on imports. Yet, on the supply side of politics, the parties more in favor of deregulation and a thin state—such as the Republicans in the US—tend to be also strongly protectionist. Their anti-globalization stances actually place them closer to the demands of workers pessimistic about GenAI, yet these workers also call for more government intervention and redistribution. Consistent with these complexities, we find that pessimistic workers get warmer toward backlash parties, but also toward mainstream parties. Overall, the political repercussions of the structural divides driven by GenAI are still unclear. There is an open space for political entrepreneurs to fine-tune coherent platforms that may be appealing to broader constituencies.

Taken together, the results of this paper suggest that GenAI may fragment rather than unify worker preferences, increasing the difficulty of designing and implementing policies that reduce inequality and sustain democratic support. The very technologies that expand productivity and autonomy at work may, perversely, erode the political foundations for collective responses to their disruptive effects. Understanding how to bridge this emerging divide—by building narratives and institutions that can mobilize and bring together both pessimistic and optimistic workers—emerges as a central challenge for democratic societies facing the AI revolution.

References

ABBOTT, R. and BOGENSCHNEIDER, B. (2018). Should robots pay taxes: Tax policy in the age of automation. *Harv. L. & Pol'y Rev.*, **12**, 145.

ACEMOGLU, D. and RESTREPO, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of economic perspectives*, **33** (2), 3–30.

AGNOLIN, P., COLANTONE, I. and STANIG, P. (2025). In search of the causes of the globalization backlash: Methodological considerations on post-treatment bias. *Comparative Political Studies*, **58** (8), 1603–1635.

AMABLE, B. (2003). *The diversity of modern capitalism*. Oxford University Press.

ANELLI, M., COLANTONE, I. and STANIG, P. (2021). Individual vulnerability to industrial robot adoption increases support for the radical right. *Proceedings of the National Academy of Sciences*, **118** (47), e2111611118.

ARDANAZ, M., MURILLO, M. V. and PINTO, P. M. (2013). Sensitivity to issue framing on trade policy preferences: evidence from a survey experiment. *International Organization*, **67** (2), 411–437.

ARMSTRONG, B., CHEN, V. K., CUELLAR, A., FORSEY-SMEREK, A. and SHAH, J. A. (2024). Automation from the worker's perspective. *arXiv preprint arXiv:2409.20387*.

BALLARD-ROSA, C., JENSEN, A. and SCHEVE, K. (2022). Economic decline, social identity, and authoritarian values in the United States. *International Studies Quarterly*, **66** (1).

—, MALIK, M. A., RICKARD, S. J. and SCHEVE, K. (2021). The economic origins of authoritarian values: Evidence from local trade shocks in the United Kingdom. *Comparative Political Studies*, **54** (13), 2321–53.

BEAULIEU, E. (2002). Factor or industry cleavages in trade policy? an empirical analysis of the stolper–samuelson theorem. *Economics & Politics*, **14** (2), 99–131.

—, BENARROCH, M. and GAISFORD, J. D. (2011). Intra-industry trade liberalization: Why skilled workers are more likely to support free trade. *Review of International Economics*, **19** (3), 579–594.

BENABOU, R. and OK, E. A. (2001). Social mobility and the demand for redistribution: The poum hypothesis. *The Quarterly journal of economics*, **116** (2), 447–487.

BERAMENDI, P., HÄUSERMANN, S., KITSCHELT, H. and KRIESI, H. (2015). *The politics of advanced capitalism*. Cambridge University Press.

BICCHI, N., KUO, A. and GALLEGO, A. (2025). Unpacking technological risks: Different sources of concern and policy preferences. *Political Studies*, **73** (3), 1054–1077.

BLONIGEN, B. (2008). *New evidence on the formation of trade policy preferences*. Tech. rep., National Bureau of Economic Research.

BOIX, C. (2019). Democratic capitalism at the crossroads: Technological change and the future of politics.

BORWEIN, S., MAGISTRO, B., ALVAREZ, R. M., BONIKOWSKI, B. and LOEWEN, P. J. (forthcoming). The potential for political backlash against ai. *Public Opinion Quarterly*.

BRADFORD, A. (2023). *Digital empires: The global battle to regulate technology*. Oxford University Press.

BREIMAN, L. (1996). Bagging predictors. *Machine Learning*, **24** (2), 123–140.

— (2001). Random forests. *Machine Learning*, **45** (1), 5–32.

BRESNAHAN, T. F. and TRAJTENBERG, M. (1995). General purpose technologies 'engines of growth'? *Journal of Econometrics*, **65** (1), 83–108.

BRYNJOLFSSON, E., CHANDAR, B. and CHEN, R. (2025a). *Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence*. Tech. rep., Working paper. Latest version available at https://digitaleconomy. stanford ....

—, KORINEK, A. and AGRAWAL, A. K. (2025b). A research agenda for the economics of transformative ai.

BUSEMEYER, M. R., GANDENBERGER, M., KNOTZ, C. and TOBER, T. (2023). Preferred policy responses to technological change: Survey evidence from oecd countries. *Socio-Economic Review*, **21** (1), 593–615.

—, STUTZMANN, S. and TOBER, T. (2025). Digitalization and the green transition: Different challenges, same policy responses? *Regulation & Governance*, **19** (2), 422–447.

CARRERAS, M., CARRERAS, Y. I. and BOWLER, S. (2019). Long-term economic distress, cultural backlash, and support for Brexit. *Comparative Political Studies*, **52** (9), 1396–1424.

COLANTONE, I., OTTAVIANO, G. and STANIG, P. (2022). The backlash of globalization. In *Handbook of international economics*, vol. 5, Elsevier, pp. 405–477.

— and STANIG, P. (2018). The trade origins of economic nationalism: Import competition and voting behavior in western Europe. *American Journal of Political Science*, **62** (4), 936–953.

COX, L. (2024). Great expectations: The effect of unmet labor market expectations after higher education on ideology. *American Journal of Political Science*, **68** (4), 1416–1430.

CUSACK, T., IVERSEN, T. and REHM, P. (2006). Risks at work: The demand and supply sides of government redistribution. *Oxford review of economic policy*, **22** (3), 365–389.

ELOUNDOU, T., MANNING, S., MISHKIN, P. and ROCK, D. (2024). Gpts are gpts: Labor market impact potential of llms. *Science*, **384** (6702), 1306–1308.

ESPING-ANDERSEN, G. (1990). *The Three Worlds of Welfare Capitalism*. Princeton, NJ: Princeton University Press.

FELTEN, E., RAJ, M. and SEAMANS, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, **42** (12), 2195–2217.

FERRARA, F. M. (2022). Why does import competition favor Republicans? Localized trade shocks and cultural backlash in the US. *Review of International Political Economy*, **115**, 1–24.

FERRARI, D., FRANZESE, R., JACKSON, H., WAI, R., WU, P., KIM, B., KIM, W., POLLACK, E. and SANDERS, H. (2021). How socioeconomic malaise & decline fuel xenophobic nationalist extremism, mimeo, University of Michigan.

FERRERA, M. (1996). The 'southern model' of welfare in social europe. *Journal of European Social Policy*, **6** (1), 17–37.

FRANGAKIS, C. E. and RUBIN, D. B. (2002). Principal stratification in causal inference. *Biometrics*, **58** (1), 21–29.

FRANZESE, R. J. (2019). The comparative and international political economy of anti-globalization populism. *Oxford Research Encyclopedia of Politics*.

FREY, C. B., BERGER, T. and CHEN, C. (2018). Political machinery: Did robots swing the 2016 US presidential election? *Oxford Review of Economic Policy*, **34** (3), 418–442.

FRIEDEN, J. (2019). The political economy of the globalization backlash: Sources and implications. *Policies to Make Trade Work for All, Princeton and Oxford*, pp. 181–196.

GALLEGO, A., KUO, A., MANZANO, D. and FERNÁNDEZ-ALBERTOS, J. (2022a). Technological risk and policy preferences. *Comparative Political Studies*, **55** (1), 60–92.

—, KURER, T. and SCHÖLL, N. (2022b). Neither left behind nor superstar: ordinary winners of digitalization at the ballot box. *The Journal of Politics*, **84** (1), 418–436.

GAMEZ-DJOKIC, M. and WAYTZ, A. (2020). Concerns about automation and negative sentiment toward immigration. *Psychological Science*, **31** (8), 987–1000.

GINGRICH, J. and HÄUSERMANN, S. (2015). The decline of the working-class vote, the reconfiguration of the welfare support coalition and consequences for the welfare state. *Journal of European Social Policy*, **25** (1), 50–75.

GREEN, J., GRANT, Z., EVANS, G. and INGLESE, G. (2025). Linking artificial intelligence job exposure to expectations: Understanding ai losers, winners, and their political preferences. *Research & Politics*, **12** (2), 20531680251337897.

HALL, P. A. and SOSKICE, D. (eds.) (2001). *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford: Oxford University Press.

HASLBERGER, M., GINGRICH, J. and BHATIA, J. (2025). Rage against the machine? generative ai exposure, subjective risk, and policy preferences. *Journal of European Public Policy*, pp. 1–28.

HÄUSERMANN, S., KURER, T. and ZOLLINGER, D. (2023). Aspiration versus apprehension: Economic opportunities and electoral preferences. *British Journal of Political Science*, **53** (4), 1230–1251.

—, PINGGERA, M., ARES, M. and ENGGIST, M. (2022). Class and social policy in the knowledge economy. *European Journal of Political Research*, **61** (2), 462–484.

HAYS, J., LIM, J. and SPOON, J.-J. (2019). The path from trade to right-wing populism in europe. *Electoral Studies*, **60**, 102038.

HISCOX, M. J. (2002). Commerce, coalitions, and factor mobility: Evidence from congressional votes on trade legislation. *American Political Science Review*, **96** (3), 593–608.

— (2006). Through a glass and darkly: Attitudes toward international trade and the curious effects of issue framing. *International Organization*, **60** (3), 755–780.

HÜBSCHER, E. and SATTLER, T. (2026). Austerity and populism. *Annual Review of Political Science*, **29**.

IMBENS, G. W. (2011). Experimental design for unit and cluster randomid trials. *International Initiative for Impact Evaluation Paper*.

— and WOOLDRIDGE, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, **47** (1), 5–86.

IVERSEN, T. and SOSKICE, D. (2001). An asset theory of social policy preferences. *American political science review*, **95** (4), 875–893.

— and — (2019). *Democracy and Prosperity: Reinventing Capitalism through a Turbulent Century*. Princeton University Press.

JACOBS, J. (2024). The artificial intelligence shock and socio-political polarization. *Technological Forecasting and Social Change*, **199**, 123006.

JOHNSON, S. and ACEMOGLU, D. (2023). *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity| Winners of the 2024 Nobel Prize for Economics*. Hachette UK.

KAHNEMAN, D. and TVERSKY, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47** (2), 263–291.

KIM, H. (2025). Firm heterogeneity and individual attitudes toward the liberal international order in south korea. *International Area Studies Review*, **28** (3), 334–359.

KNOTZ, C., UGARTE MONTERO, A., LAVANCHY, M. and WAGNER, J. (2024). Bankers are afraid of technology now: explaining perceived vulnerability to technological change among the higher-educated. *Political Research Exchange*, **6** (1), 2389910.

KURER, T. (2020a). The declining middle: Occupational change, social status, and the populist right. *Comparative Political Studies*, **53** (10-11), 1798–1835.

— (2020b). The declining middle: Occupational change, social status, and the populist right. *Comparative Political Studies*, **53** (10–11), 1798–1835.

— and VAN STAALDUINEN, B. (2022). Disappointed expectations: Downward mobility and electoral change. *American Political Science Review*, **116** (4), 1340–1356.

LEE, H. N.-K. and LIOU, Y.-M. (2022). Where you work is where you stand: A firm-based framework for understanding trade opinion. *International Organization*, **76** (3), 713–740.

MAGISTRO, B., BORWEIN, S., ALVAREZ, R. M., BONIKOWSKI, B. and LOEWEN, P. J. (2025). Attitudes toward artificial intelligence (ai) and globalization: Common microfoundations and political implications. *American Journal of Political Science*.

MAYDA, A. M. and RODRIK, D. (2005). Why are some people (and countries) more protectionist than others? *European Economic Review*, **49** (6), 1393–1430.

MILNER, H. V. (2021). Voting for populism in europe: Globalization, technological change, and the extreme right. *Comparative Political Studies*, **54** (13), 2286–2320.

NATIONAL CENTER FOR O*NET DEVELOPMENT (2025). O*NET OnLine. https://www.onetonline.org/.

PIZZINELLI, C., PANTON, A. J., MENDES TAVARES, M., CAZZANIGA, M. and LI, L. (2024). *Labor Market Exposure to AI: Cross-Country Differences and Distributional Implications*. Working Paper IMF Working Paper 2024, International Monetary Fund.

REHM, P. (2009). Risks and redistribution: An individual-level analysis. *Comparative political studies*, **42** (7), 855–881.

—, HACKER, J. S. and SCHLESINGER, M. (2012). Insecure alliances: Risk, inequality, and support for the welfare state. *American political science review*, **106** (2), 386–406.

RHO, S. and TOMZ, M. (2017). Why don't trade preferences reflect economic self-interest? *International Organization*, **71** (S1), S85–S108.

RODRÍGUEZ CHATRUC, M., STEIN, E., VLAICU, R. and ZULUAGA, V. (2025). *How employment framing affects trade preferences: Evidence from survey experiments*. Tech. rep., IDB Working Paper Series.

SCHEIRING, G., SERRANO-ALARCÓN, M., MOISE, A., MCNAMARA, C. and STUCKLER, D. (2024). The populist backlash against globalization: A meta-analysis of the causal evidence. *British Journal of Political Science*, **54** (3), 892–916.

SCHEVE, K. F. and SLAUGHTER, M. J. (2001). What determines individual trade-policy preferences? *Journal of International Economics*, **54** (2), 267–292.

SCHÖLL, N. and KURER, T. (2024). How technological change affects regional voting patterns. *Political Science Research and Methods*, **12** (1), 94–112.

SIMÓN, P. and CLAVERIA, S. (2024). Social class and the support for environmental policies. *International Journal of Public Opinion Research*, **37** (1).

THEWISSEN, S. and RUEDA, D. (2019). Automation and the welfare state: Technological change as a determinant of redistribution preferences. *Comparative Political Studies*, **52** (2), 171–208.

TOLAN, S., PESOLE, A., MARTÍNEZ-PLUMED, F., FERNÁNDEZ-MACÍAS, E., HERNÁNDEZ-ORALLO, J. and GÓMEZ, E. (2021). Measuring the occupational impact of ai: tasks, cognitive abilities and ai benchmarks. *Journal of Artificial Intelligence Research*, **71**, 191–236.

WEBB, M. (2019). The impact of artificial intelligence on the labor market. *Available at SSRN 3482150*.

WEISSTANNER, D. and VAN KERSBERGEN, K. (2025). Managing the artificial intelligence revolution: Perceived risks and social policy preferences among firm-level decision makers. *Social Policy & Administration*.

WU, N. (2022). Misattributed blame? attitudes toward globalization in the age of automation. *Political Science Research and Methods*, **10** (3), 470–487.

ZHANG, B. (2024). Public opinion toward artificial intelligence. In J. B. B. et al. (ed.), *The Oxford Handbook of AI Governance*, Oxford: Oxford University Press.

Supporting Information (SI) for:

What If You See It?
Generative AI Exposure and Workers' Divergent Policy Preferences

Massimo Anelli, Italo Colantone, Aina Gallego*,
Piero Stanig

*Corresponding author. Email: agallego@ibei.org

For online publication only

Table of contents:

S1 Selection of occupations

Our sample includes workers employed in 98 6-digit SOC occupations that are identified as "directly exposed" to GenAI. In order to select these occupations, we follow the following steps. First, we source occupation-specific scores of exposure to AI from three existing studies: Felten *et al.* (2021), Tolan *et al.* (2021), and Webb (2019).[22] Scores by Felten *et al.* (2021) and Webb (2019) are available at the 6-digit level of disaggregation, whereas those by Tolan *et al.* (2021) are only available at the more aggregated 3-digit level. We aggregate the 6-digit scores in Felten *et al.* (2021) and Webb (2019) at the 3-digit level using US employment weights from the Bureau of Labor Statistics (BLS). Specifically, weights are based on May 2022 estimates for the total number of employees within each 6-digit occupation in the US. After this aggregation, we have three distinct automatability scores for each 3-digit SOC occupation.

We conduct a k-means cluster analysis with $k = 2$ across occupations on each measure, separately, to classify occupations into two groups: exposed or not exposed to AI. We thus obtain three different binary classifications of occupations. We select all the 3-digit occupations classified as exposed to AI according to at least two out of three classifications. This step results in the selection of 39 3-digit SOC occupations (out of a total of 94). They are reported in the pre-analysis plan and in the replication materials.

We then focus on all the 369 6-digit SOC occupations within the selected 39 3-digit occupations. Among them, we identify as directly exposed to GenAI those where *the most frequent core task* is automatable through ChatGPT. The most frequent core task for each occupation is identified based on O*NET data. In fact, O*NET provides a list of performed tasks for each occupation, along with data on how frequently each given task is performed by people employed in a given occupation.[23] In order to evaluate the automatability of each of the 369 occupation-specific tasks, we interrogate the OpenAI API (model gpt-4 at temperature 0). The assessment is based on the classification system and the exposure rubric in Eloundou *et al.* (2024). Specifically, a task can be classified by ChatGPT in one of four ways: non exposed to LLMs (coded E0); directly exposed to LLMs (coded E1); exposed by LLM-powered applications (coded E2); and exposed though image capabilities (coded E3). In the Eloundou et al. rubric, a task qualifies as "directly exposed to LLMs" (E1) when access to ChatGPT alone—without additional tooling—can reduce the time required to complete the task with equivalent quality by at least 50%. This is the operational meaning of the 50% time-reduction criterion stated in the main paper.

Our prompt text reads:

> Please rely on the following rubric to classify different combinations of occupa-

[22]These measures are obtained through different, and complementary, approaches. In particular, scores by Felten *et al.* (2021) and Tolan *et al.* (2021) are based on progress in AI capabilities that are relevant to specific abilities and occupations. Scores by Webb (2019) are based on the development of patents that are relevant for specific tasks, that are in turn linked to occupations. We have downloaded data for the three scores from the replication package of Tolan *et al.* (2021), available at `https://repositori.upf.edu/handle/10230/55770`.

[23]The frequency of each task is computed based on survey data among workers within each occupation. In almost all cases there is one core task that is strictly more frequent than the others. Only in a handful of cases there are two core tasks that are equally frequent. In such situations we focus on the most important one, based on O*NET importance scores. Most 6-digit SOC occupations correspond to a single '6-digit SOC.00' code in O*NET. In cases where multiple correspondences exist (including '.00'), we prioritize the '.00' code, ignoring others like '.01', '.02', etc. If there is no '.00' code, but only codes like '.01', etc., we select the most frequent core task among these O*NET codes. For occupations where the last digit before the decimal in the code changes (referred to as the 'prefix'), such as 13-1020 evolving into 13-1021.00, 13-1022.00, 13-1023.00, we follow the same approach as above but consider all codes with the new prefix.

> tions and job tasks. It is crucial that you do not simply match keywords between the task descriptions and the rubric. Instead, please apply a thorough analysis, considering the overall context, required skills, and complexity of the tasks. Ensure to utilize your general knowledge about work tasks and automation possibilities in various industries. I provide the rubric below, and after that I will give you the list of occupations and tasks to be classified. Please provide the output [with] [without] explanation in the following csv format separated by semi-colons: classification; task ID; [explanation].

We then pass the classification rubric sourced from Eloundou *et al.* (2024) and ask ChatGPT to classify each occupation-specific task 5 times. We select only those occupations where the answer is unanimously "directly exposed to LLMs". This leaves us with a total of 176 6-digit SOC occupations classified as directly exposed to GenAI. These account for around 16.3% of total employment in the US, based on 2022 BLS data. We finally select 98 out of these 176 occupations by including all the 82 occupations representing at least 0.03% of US employment, plus 16 occupations that were needed in order to ensure representativeness also in the German and Italian panels of the survey company (YouGov). The list of all 6-digit SOC occupations included in the sample is provided in the pre-analysis plan.

We select into the sample only respondents who are currently employed in one of the 98 selected occupations. Respondents report their current occupation at the very beginning of the survey through an interactive interface. The reference 6-digit SOC occupation titles are in English. The translation into German and Italian is done through the official cross-walk from O*NET to ISCO codes made available by the EU Commission.[24] There are some cases in which the same SOC title corresponds to multiple ISCO codes, with their corresponding official denominations in German and Italian. In such cases, all the corresponding ISCO codes are included in the sample. The full list of included ISCO codes, with their denominations in German and Italian, is reported in the pre-analysis plan and will be included in the replication materials.

S2 Production of the videos

The occupation-specific treatment videos are produced through the steps described in what follows.

Step 1. For each occupation, we create a GPT extension. This is trained on information provided by O*NET (National Center for O*NET Development, 2025), and on the detailed occupation description produced by CareerOneStop, which is the U.S. Department of Labor's flagship career, training, and job-search website, providing free tools, data, and guidance for job seekers, students, businesses, and workforce professionals. Sponsored by Department Of Labor's Employment and Training Administration, it connects users to labor-market information, skills assessments, and local services through the American Job Center network.

For each occupation, we pass to ChatGPT the following instructions:

[24] It can be found at this link: `https://esco.ec.europa.eu/en/use-esco/other-crosswalks`.

> Read carefully the information available at the following links:
>
> https://www.onetonline.org/link/summary/[O*NET-SOC Code]
>
> and
>
> https://www.onetonline.org/link/details/[O*NET-SOC Code]
>
> Your job description is as follows: "[O*NET-SOC Description]"
>
> In order to give you a better understanding of the job I am pasting the transcript of a video explanation of the job:
>
> [Insert here the transcript from Career One Stop][25]

We also upload on the GPT extension all the occupation-specific O*NET Excel files with information on: tasks performed; (detailed) work activities; skills, abilities and knowledge required; tools used; work styles and interests.

Step 2. We ask the GPT extension to conceive a hypothetical scenario in which ChatGPT would perform the most frequent core task of the given occupation, through the following prompt:

> Imagine that you are tasked with the following: "[O*NET Task statement associated with an occupation]"
>
> Your mission is to demonstrate how you would complete this task in a hypothetical real-life scenario, for example while preparing graphs, data analysis reports, essays, etc. Please provide an illustration of you completing the task – for example, creating graphs, data analysis reports, writing essays, summarizing, translating, coding, making internet research, and so on, as applicable to the scenario. Show me the output of your work. Instead of a list of actions, please provide a concrete example of you performing the task. Before creating the hypothetical scenario, ensure you leverage the information from the provided links. Explain your choice based on this information.

We allow for some tailoring of our interaction with ChatGPT. This includes: (1) selecting the scenario or the task example we deem most aligned with the specified task in cases where ChatGPT presents multiple options; and (2) requesting ChatGPT to provide materials, such as Excel files or PowerPoint presentations, used to carry out the task.

Step 3. We guide ChatGPT to formulate an optimal prompt for the selected scenario, by asking "How would you describe this hypothetical scenario to ChatGPT to ensure it effectively carries out the task?". The resulting prompt is designed to accurately depict the hypothetical scenario, thereby enabling ChatGPT to execute the task effectively.

[25] For example, if the occupation code is 29-1141.00 (registered nurse), the summary link would be https://www.onetonline.org/link/summary/29-1141.00; the details link would be https://www.onetonline.org/link/details/29-1141.00; and the video link would be https://www.careeronestop.org/Toolkit/Careers/Occupations/occupation-profile.aspx?keyword=291141.00.

Step 4. We submit to ChatGPT the prompt of Step 3, along with any file attachments produced at Step 2. We include specific instructions, such as “please limit your response to no more than 150 words”. Once we receive a response, we request ChatGPT to translate both the prompt and the response into German and Italian, to be used for survey participants in Germany and Italy.

Step 5. For each language, we perform the minimal editing needed to have a consistent video format and length and to standardize the pace at which the conversation appears on screen, without altering the substance of the conversation. After these adjustments, we record the unfolding of both the prompt and the response through OBS Studio.

Step 6. Synthesized voiceovers (male and female for each language) read out the entire conversation. For English, we use OpenAI’s API TTS-1 model with the male voice “Echo” and the female voice “Nova”. For German and Italian, we use Google’s Text-to-Speech APIs. For German, we use the male voice “de-DE-Wavenet-B” and the female voice “de-DE-Wavenet-C”. For Italian, we use the male voice “it-IT-Wavenet-C” and the female voice “it-IT-Wavenet-B”. We randomize between male and female voices in the experiment to control for potential implicit biases related to gender. The same voice is used for both the prompt and the answer. A voice of the opposite gender is used between the end of the prompt and the beginning of the answer to introduce ChatGPT’s answer, to emphasize the distinction between the prompt and the answer and to help survey participants better follow and understand ChatGPT’s writing flow. Synchronization between the voice and the on-screen typing is done through DaVinci Resolve Studio, a professional video editing software. Where needed, key inputs of the task, such as Excel and PowerPoint files, are displayed. All videos are available at the following link.

S3 Validation of the experimental stimuli

Table S1: **Differences in GenAI impact perceptions between the two control conditions.**

	(1) Overall Work Impact	(2) Productivity vs. Redundancy	(3) Threat vs. Opportunity	(4) AIP
Control video	0.0211	-0.0352	0.00986	-0.00141
	(0.0297)	(0.0518)	(0.0388)	(0.0319)
Observations	3,057	3,057	3,057	3,057
R-squared	0.114	0.094	0.095	0.101
Mean Dep. Var.	0.310	0.217	0.126	0.218
S.D. Dep. Var	0.834	1.510	1.101	0.933

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Note: As a preliminary assessment, we ensure that there are no statistically significant differences across the two control arms in terms of perceptions of GenAI’s occupational impact by estimating the ATE of receiving the control video (vs. no video) on AIP and its main components. This is done based on the specification outlined in Equation 1. As per the pre-registered criterion, in the absence of consistent statistically detectable effects of the control video, we have proceeded in the analysis without differentiating the two control arms.

Table S2: **Variation across occupations by respondents' evaluations of the task shown in the video and effectiveness of the video.**

	(1) Task Frequency	(2) Task Enjoyment	(3) Task Feeling	(4) Video Effectiveness
Adjusted R^2	0.0701	0.0620	0.0827	0.0497
Observations	2,926	2,926	2,926	2,926
Occupation (Block) FE	Yes	Yes	Yes	Yes

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note: The outcome variables are respondents' assessments of: (1) the frequency at which they perform the task; (2) the enjoyment they obtain from performing the task; (3) the feeling about the importance of the task compared to others; and (4) the effectiveness of the video at demonstrating ChatGPT's general potential to assist with tasks in their work environment. These outcome variables are regressed on the full set of occupation fixed effects. The adjusted R^2 values are consistently low across the board, pointing to limited between-occupation systematic variation.

S4 Sample descriptives

The survey was fielded by YouGov between 29 January and 3 March 2025 in Germany and Italy, and between 15 January and 2 March 2025 in the United States. Because we recruit only respondents employed in occupations directly exposed to GenAI, our sample is not representative of the overall national population. The plots in SI Figure S1 compare our sample against the general population of Germany, Italy, and the United States on age, gender, education, and income. Across all three countries, our respondents are younger, more educated, and earn higher incomes than the general population, consistent with the cognitive-occupation focus of the design. SI Tables S3, S4, and S5 extend the comparison to political behavior, reporting self-reported turnout in the most recent national election and vote shares for the major parties, together with the corresponding figures for the general population. The vote-share comparisons reveal a parallel pattern: our sample under-represents radical-right parties (FdI and Lega in Italy, and Republicans in the US, albeit not AfD in Germany) and over-represents progressive alternatives (Greens and Linke in Germany, PD in Italy, Democrats in the US), consistent with the demographic-educational skew documented in the figures.

Table S3: Self-reported turnout and vote shares: **Germany**

	AfD	Greens	CDU	CSU	Die Linke	FDP	SPD	Other	PNA	Turnout
Sample	11.2%	18.1%	18.2%	4.6%	6.0%	9.9%	22.2%	3.8%	6.2%	85.7%
Population	10.4%	14.7%	19.0%	5.2%	4.9%	11.4%	25.7%	8.7%	n/a	76.4%

Note: Vote shares refer to the 2021 German federal election. In our sample, the category "Other" reads as "other party or candidate only." "PNA" means "prefer not to answer."

Figure S1: Sample versus population comparisons by country. Rows correspond to age, gender, education, and income. Columns correspond to Germany, Italy, and the United States.

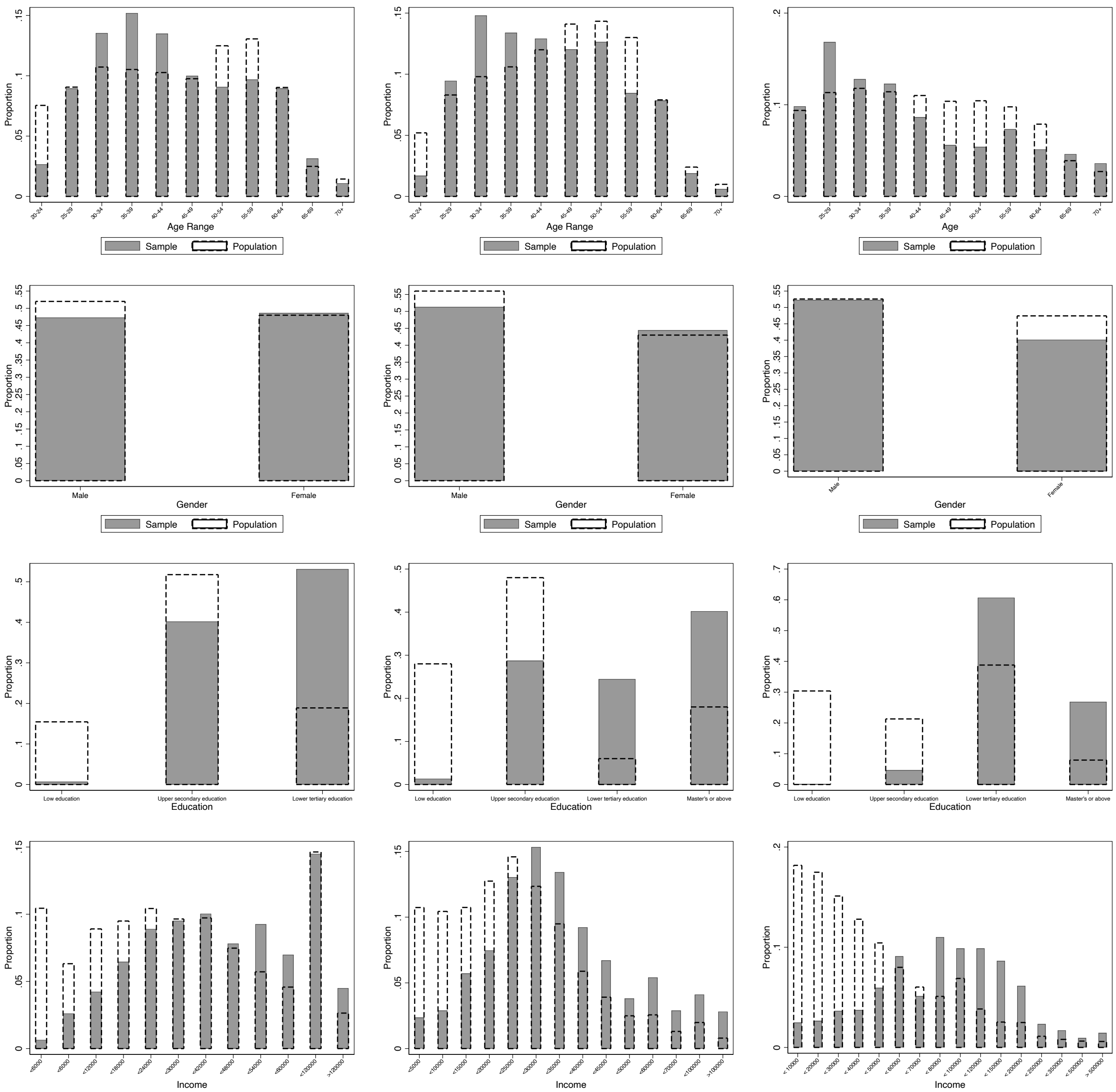


Table S4: Self-reported turnout and vote shares: **Italy**

	FI	FdI	Lega	M5S	PD	Other	PNA	Turnout
Sample	3.9%	15.9%	4.3%	13.7%	21.2%	23.5%	17.5%	84.1%
Population	8.1%	26.0%	8.8%	15.4%	19.0%	22.7%	n/a	63.8%

Note: Vote shares refer to the 2022 Italian legislative election. In our sample, the category "Other" includes the following: +Europa, Alleanza Verdi e Sinistra, Azione - Italia Viva – Calenda, Italexit per l'Italia, Unione Popolare, and Other party or candidate. "PNA" means "prefer not to answer."

Table S5: Self-reported turnout and vote shares: **United States**

	Republican	Democrat	Other	PNA	Turnout
Sample	36.1%	59.3%	2.3%	2.3%	93.6%
Population	49.8%	48.3%	1.9%	n/a	64.1%

Note: Vote shares refer to the 2024 U.S. presidential election. In our sample, the category "Other" includes the following: Green Party, Libertarian Party, and Other party or candidate. "PNA" means "prefer not to answer."

S5 Dealing with insufficiently numerous blocks

Having at least one treated and one control unit is required for the treatment effect by block to be well defined. In addition, in order to correctly estimate the sampling variance of the estimate as explained by Imbens (2011), we need at least two treated and two controls within each block. Blocks with at least two treated and two controls are "well-formed", whereas we call blocks without at least two treated and two controls "insufficiently numerous blocks" (INBs). As it turns out, only 13 observations belong to INBs. These are found in six different 3-digit SOC occupations.[26] Following the pre-registered procedure, these observations are pooled in a residual block. After this aggregation, the analysis includes 92 blocks.

[26] Specifically, they belong to the following 6-digit SOC categories: 19-2042 "Geoscientists, Except Hydrologists and Geographers"; 21-1015 "Rehabilitation Counselors; 29-1126 "Respiratory Therapists"; 49-3011 "Aircraft Mechanics and Service Technicians; 51-8012 "Power Distributors and Dispatchers"; 51-8093 "Petroleum Pump System Operators, Refinery Operators, and Gaugers"; 53-2012 "Commercial Pilots".

S6 Baseline correlates of AI perceptions

Table S6: Baseline AIP and socio-demographic correlates

Socio-demographic covariates batch	(1) AIP (control)
Age dummy	-0.037
	(0.035)
Female	-0.008
	(0.033)
ISCED 3–4: Upper secondary	0.365**
	(0.183)
ISCED 5–6: Lower tertiary	0.369**
	(0.182)
ISCED 7–8: Master's or above	0.456**
	(0.183)
Education missing	0.482*
	(0.264)
High income (within country)	0.155***
	(0.034)
AI awareness: Somewhat	-0.206***
	(0.042)
AI awareness: A little	-0.246***
	(0.054)
AI awareness: Not at all	-0.458***
	(0.112)
LLM use: Yes, occasionally	-0.133***
	(0.047)
LLM use: No	-0.494***
	(0.053)
Constant	0.061
	(0.185)
Observations	2,714
R-squared	0.119

Robust standard errors in parentheses. Country fixed effects absorbed.
Higher values of the AIP index indicate more optimistic perceptions.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

S7 Stratification approach

In the pre-analysis plan, we noted that the potential effect of the informational treatment could be heterogeneous across respondents. ChatGPT can complement workers and make them more productive, but it can also substitute workers in many of the tasks they perform

Table S7: Baseline AIP and psychological correlates

Psychological covariates batch	(1) AIP (control)
Risk tolerance	0.060***
	(0.007)
Extraversion	-0.018*
	(0.010)
Agreeableness	0.064***
	(0.011)
Conscientiousness	0.018*
	(0.011)
Neuroticism	-0.048***
	(0.010)
Openness / imagination	0.012
	(0.010)
Attachment: depend on others	0.044***
	(0.016)
Attachment: fear abandonment	0.005
	(0.015)
Attachment: handle problems alone	-0.018
	(0.017)
Attachment: afraid of others	0.036**
	(0.016)
Constant	-0.700***
	(0.170)
Observations	3,057
R-squared	0.073

Robust standard errors in parentheses. Country fixed effects absorbed.
For the attachment items, higher values mean stronger disagreement with the target label.
Higher values of the AIP index indicate more optimistic perceptions.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table S8: Baseline AIP, ideology, and psychological correlates

Ideological & psychological covariates batch	(1) AIP (control)
Ideological self-placement	-0.023***
	(0.009)
Social dominance orientation	-0.023
	(0.023)
Right-wing authoritarianism	0.128***
	(0.024)
Left-wing authoritarianism	-0.062***
	(0.022)
Risk tolerance	0.062***
	(0.008)
Extraversion	-0.021**
	(0.011)
Agreeableness	0.057***
	(0.012)
Conscientiousness	0.017
	(0.011)
Neuroticism	-0.050***
	(0.011)
Openness / imagination	0.006
	(0.010)
Attachment: depend on others	0.053***
	(0.017)
Attachment: fear abandonment	0.001
	(0.016)
Attachment: handle problems alone	-0.007
	(0.018)
Attachment: afraid of others	0.027
	(0.017)
Constant	-0.504***
	(0.189)
Observations	2,739
R-squared	0.088

Robust standard errors in parentheses. Country fixed effects absorbed.
For the attachment items, higher values mean stronger disagreement with the target label.
Higher values of the AIP index indicate more optimistic perceptions.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table S9: Baseline AIP and occupational correlates

Occupational covariates batch	(1) AIP (control)
Work trauma	-0.270***
	(0.050)
Interpersonal relations importance	-0.069***
	(0.019)
Writing tasks	-0.035
	(0.024)
Coding tasks	-0.090***
	(0.022)
Data analysis tasks	-0.092***
	(0.024)
Customer service tasks	0.059**
	(0.024)
Active listening	-0.007
	(0.029)
Critical thinking	-0.080***
	(0.030)
Public sector	-0.156***
	(0.042)
Hours: 10–19	-0.002
	(0.130)
Hours: 20–29	0.060
	(0.117)
Hours: 30–39	0.190*
	(0.112)
Hours: 40–49	0.126
	(0.113)
Hours: 50 or more	0.184
	(0.140)
Permanent contract	-0.039
	(0.040)
Union member	-0.008
	(0.044)
Constant	0.686***
	(0.140)
Observations	3,057
R-squared	0.063

Robust standard errors in parentheses. Country FE absorbed. Industry FE suppressed.
For interpersonal relations and work tasks, higher values mean lower importance.
Higher values of the AIP index indicate more optimistic perceptions.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table S10: Summary statistics of pre-treatment characteristics, by stratum.

	Strata									
	Always Optimistic		Pushed to Pessimistic		Always Pessimistic		Pushed to Optimistic		Total	
Female	0.38	(0.48)	0.47	(0.50)	0.55	(0.50)	0.47	(0.50)	0.47	(0.50)
Age	38.93	(12.30)	41.98	(12.95)	46.27	(13.14)	42.99	(13.20)	42.57	(13.19)
Income	96,561.24	(100,829.75)	72,966.24	(73,911.98)	60,864.18	(39,213.09)	70,768.06	(70,235.93)	76,949.01	(77,193.25)
Education	2.09	(0.65)	1.92	(0.68)	1.83	(0.72)	1.95	(0.72)	1.96	(0.70)
Extraversion [2; 10]	6.09	(1.77)	5.90	(1.72)	5.83	(1.81)	6.05	(1.80)	5.96	(1.78)
Agreeableness [2; 10]	7.22	(1.66)	6.53	(1.54)	6.38	(1.58)	6.83	(1.52)	6.76	(1.64)
Conscientiousness [2; 10]	7.87	(1.62)	7.32	(1.67)	7.47	(1.63)	7.54	(1.64)	7.60	(1.65)
Neuroticism [2; 10]	4.85	(1.81)	5.61	(1.71)	5.70	(1.86)	5.48	(1.80)	5.35	(1.85)
Openness [2; 10]	7.09	(1.60)	6.90	(1.71)	6.83	(1.70)	7.09	(1.67)	6.97	(1.66)
Risk aversion [0; 10]	6.62	(2.22)	5.87	(2.22)	5.16	(2.42)	5.89	(2.35)	5.89	(2.39)
Dependence [1; 5]	3.42	(1.11)	3.34	(1.04)	3.22	(1.05)	3.45	(1.06)	3.34	(1.07)
Abandonment [1; 5]	2.98	(1.26)	3.14	(1.24)	3.08	(1.29)	3.07	(1.28)	3.05	(1.27)
Self-reliance [1; 5]	3.79	(1.00)	3.71	(0.97)	3.73	(0.97)	3.70	(1.06)	3.74	(1.00)
Attachment anxiety[1; 5]	2.76	(1.22)	3.00	(1.13)	2.89	(1.19)	2.98	(1.22)	2.88	(1.20)
AI Awareness [1; 4]	1.40	(0.56)	1.74	(0.71)	2.21	(0.75)	1.83	(0.71)	1.80	(0.76)
Previous LLM Use [1; 3]	1.59	(0.54)	1.98	(0.73)	2.66	(0.61)	2.09	(0.74)	2.10	(0.77)
Ideology [0;10]	5.18	(2.61)	5.12	(2.40)	5.08	(2.39)	5.09	(2.47)	5.13	(2.49)
Observation	2100 (35.1%)		891 (14.9%)		2101 (35.1%)		891 (14.9%)		5983 (100.0%)	

Notes: (1) Income is adjusted in terms of purchasing power to be comparable across countries. Education is measured on a standardized scale, with values ranging from 0 ("ISCED 0-2: Low education") and 1 ("ISCED 3-4: Upper secondary education - broad") to 2 ("ISCED 5-6: Lower tertiary") and 3 ("ISCED 7-8: Master's or above").
(2) Extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience are big-5 personality traits, with values ranging from 2 to 10 (where 2 and 10 represent extremely low levels and respectively extremely high levels of the trait concerned). For risk aversion, the scale is inverted, such that lower values mean higher risk aversion, while higher values indicate a stronger risk appetite. For items related to the individual's comfort with dependence, fear of abandonment, self-reliance, and fear of other people ("attachment anxiety"), higher values represent a higher intensity of the characteristic concerned.
(3) "AI Awareness" is the response to the survey question "Prior to taking this survey, how much had you heard or read about generative AI?" Answers range from 1 ("A lot") to 4 ("Not at all"). Similarly, previous LLM use comes from the survey question "Have you used large language models like ChatGPT or Gemini for your work or other activities?", where answers range from 1 ("Yes, frequently") to 3 (No").
(4) Ideology is the numerical response to the survey question: "Thinking of politics, where would you place yourself on a scale where 0 is most to the left, and 10 is most to the right"?

on the job, and make them redundant in the limit if a large share of activities are substituted. Even after watching exactly the same video, different respondents employed in the same occupation may be tilted either in a positive or in a negative direction regarding their assessment of GenAI's impact on their occupation. This may depend on personality traits and other individual characteristics. In light of this, the reaction to the video treatment is a crucial moderator for our analysis of treatment effects on policy and political preferences. This is reflected in the way in which hypotheses are posited and tested.

The analysis is based on a principal stratification approach where hypotheses are stratum-specific (Frangakis and Rubin, 2002). That is, the hypothesized treatment effects are conditional on the change in GenAI impact perception induced by the treatment. We classify respondents in one of four strata: the "pushed to pessimistic" stratum, i.e., those whose perceptions are worsened; the "pushed to optimistic" stratum, i.e., those whose perceptions are improved; the "always pessimistic" and the "always optimistic" strata, including respondents whose perceptions remain essentially unchanged. We then estimate stratum-specific treatment effects to investigate how perceived occupational consequences shape policy and political preferences. Table S10 reports descriptive statistics of pre-treatment characteristics by stratum membership.

We assign individuals to a stratum based on two predicted values of GenAI Impact Perception (AIP), under treatment and under control. Denote by AIP_i the impact perception of individual i, with potential outcomes $AIP_i(T)$ as a function of the treatment $T \in \{0,1\}$. For each individual we predict two "prognostic" scores for the potential outcomes of the moderator: $A\hat{I}P_i(0)$ and $A\hat{I}P_i(1)$. We predict $A\hat{I}P_i(0)$ for all respondents based on a random forest estimated only on control respondents, and $A\hat{I}P_i(1)$ for all respondents based on a random forest estimated only on treated respondents, as detailed below. We then assign respondents to a stratum based on whether the predicted values $A\hat{I}P_i(0)$ and $A\hat{I}P_i(1)$ are above or below the median of the predicted values taken over the whole sample: $A\tilde{I}P(0)$ and $A\tilde{I}P(1)$. The allocation is outlined in Table S11. The last column of the table reports the predicted change in AIP, $\hat{\mu}_i$, where $\mu_i = AIP_i(1) - AIP_i(0)$. This is the individual causal effect of the treatment on AIP_i. We use a direct measure of $\hat{\mu}_i$ as an alternative approach to assess heterogeneity in the analysis.

Table S11: **Strata Assignment**

Stratum	$A\hat{I}P_i(0)$	$A\hat{I}P_i(1)$	$\hat{\mu}_i$
Always Optimistic (1)	$> A\tilde{I}P(0)$	$> A\tilde{I}P(1)$	≈ 0
Pushed to Pessimistic (2)	$> A\tilde{I}P(0)$	$< A\tilde{I}P(1)$	< 0
Always Pessimistic (3)	$< A\tilde{I}P(0)$	$< A\tilde{I}P(1)$	≈ 0
Pushed to Optimistic (4)	$< A\tilde{I}P(0)$	$> A\tilde{I}P(1)$	> 0

As inputs for the random forest exercises, we employ the following variables. They are all measured on five-point Likert scales unless otherwise specified. Likert scales are treated as numeric. We include variable-specific indicators for missing values in numerical variables. We treat "missing" as an additional category in categorical and binary variables. The list is:

- age (treated as numeric)
- gender (categorical)
- education (categorical, harmonized according to the ISCED scheme)
- individual income (ordered categories, harmonized across countries, and treated as numeric)
- ideological self-placement (0-10 scale treated as numerical with indicator for missing values)
- "big five" personality traits (ten items, included separately)
- risk aversion (0-10 scale, treated as numeric)
- importance of interpersonal relations at work
- traumatic events at work (three binary variables)

- attachment (four items)
- willingness to retrain (two items)
- occupation (six-digit SOC, categorical)[27]

∗ industry of occupation (NACE Rev 1.1 Sections)

∗ union membership (binary)

∗ private vs. public sector (binary)

∗ type of contract (categorical)

∗ hours worked (treated as numerical)

∗ foreign born (binary)

∗ ethnicity (categorical, with country-specific categories)

∗ AI awareness (four-point Likert, treated as numeric)

∗ prior use of Gen AI (three-category ordered scale, treated as numeric)

∗ work tasks (six items, each measured on a three-point Likert, treated as numeric)

All these variables are conceptually pre-treatment, even though some of the items—those marked by $*$ in the list above—are asked in the questionnaire after showing the video. We are comfortable using these items as pre-treatment inputs in the predictive exercise, based on the idea that respondents remain, for instance, private or public sector workers regardless of the treatment received.

To select the optimal random forest model for prediction, we tune the hyperparameters through cross-validation using the `caret` package in R. We tune two key hyperparameters: (1) the number of randomly selected features at each split (mtry) and (2) the minimum number of observations required in terminal nodes (nodesize) via a grid search over the following ranges: mtry values of 2, 5, 10, and 15, and nodesize values of 5, 10, 20, 50. A 5-fold cross-validation procedure is applied, yielding mtry=15 and nodesize=5 for each of the three response variables. For the final models, 500 trees (ntree = 500) are grown to ensure stability in predictions. We perform the random forest exercise twice: to estimate the conditional expectation function for $AIP(0)$—based only on the controls—and to estimate

[27]Occupation is classified using the BLS Standard Occupational Classification (SOC) at the 6-digit level. Whenever a 6-digit occupation has fewer than 60 observations in the sample (pooled across countries, i.e., 1% of the sample), we proceed as outlined in the pre-analysis plan. First, we group it with other 6-digit occupations within the same 3-digit occupation that are below the threshold of 60 observations. We form 9 groups at the 3-digit level at this step, involving a total of 1,409 observations. 4,233 observations retain their original 6-digit assignment. 341 observations are placed in a single residual category as it was not possible to form with them, at the 3-digit level, a group of size larger than 60. Notice that this strategy differs from the aggregation rules for causal estimation. There, we need at least two treated and two controls per occupation, while here the objective is estimating the predictive model and therefore avoid including sparse categories.

it for $AIP(1)$—based only on the treated. We then predict $AIP_i(0)$ and $AIP_i(1)$ for all observations from the respective random forest estimations. We implemented random forests using the `ranger` package in R, which by default employs bagging: each tree is trained on a bootstrap sample of the data (Breiman, 1996, 2001). Bagging reduces overfitting by averaging predictions across many trees trained on resampled datasets.

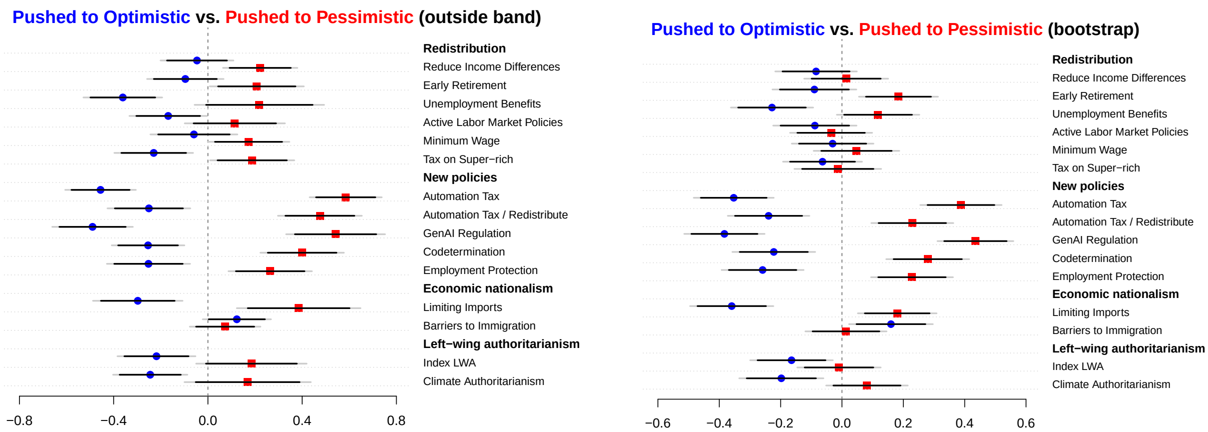
The in-sample predictive performance of the random forests is quite high. In the control arm, the predicted and observed values of $AIP(0)$ are correlated at 0.97, and a regression of the observed values on the predicted values yields an RMSE of .23, approximately one-quarter of a point where AIP is measured on the $[-2, 2]$ scale. In the treatment arm, the predicted and observed values of $AIP(1)$ are correlated at 0.92, and the RMSE of the regression of observed on predicted values yields an RMSE of 0.27.

For some analyses, we leverage also the directly-observed empirical counterpart of the treatment effect μ_i. We ask—only—treated respondents whether watching the video has made them more optimistic or more concerned than before about the overall impact of GenAI on jobs like theirs. Answers are on a five-point scale from “Much more concerned” to “Much more optimistic”. We use this as outcome in a random forest to estimate the distribution of the individual reaction to the video conditional only on pre-treatment covariates, i.e., $E(\mu|X)$, and then predict prognostic scores $\hat{\mu}_i$ for all respondents based on the observed values of their pre-treatment covariates. In the treatment arm, the observed and predicted values are correlated at 0.96, and the regression of observed on predicted values yields an RMSE of 0.27.

The stratification we propose is always based on prognostic scores constructed exclusively from pre-treatment covariates. It is important to stress that conditioning on stratum membership defined by these prognostic scores does not introduce post-treatment bias. In fact, post-treatment bias arises when one conditions on the observed value of a mediator (AIP, in our case). In that situation, comparing treated and control units conditional on the observed mediator involves averaging over different mixtures of principal strata. For example, consider a comparison between treated and control units who both display optimistic observed perceptions about GenAI. Among control units, this group includes those who are always optimistic, and those who would have been made pessimistic by treatment had they received it. Among treated units, it includes those who are always optimistic, and those who are made optimistic by the treatment. The resulting difference in means does not identify a causal quantity, even if treatment is randomly assigned, because the treated and control observations belong to different mixtures of strata. By contrast, our approach conditions on (an approximation to) the pair of potential outcomes for AIP—that is, on stratum membership itself. These potential outcomes, and thus stratum membership, are not affected by treatment. Therefore, any comparison within a principal stratum (i.e., conditional on stratum membership) has a causal interpretation when, as in our case, treatment assignment is random and thus orthogonal to the potential outcomes (Frangakis and Rubin, 2002).

S8 Robustness checks

Figure S2: **Stratum-specific treatment effects are robust to alternative treatments of uncertainty in stratum assignment.**



(a) Excluding observations near thresholds **(b) Bootstrap uncertainty**

Note: Panel (a) re-runs the analysis in Figure 2 after excluding observations whose predicted $\widehat{AIP}_i(0)$ or $\widehat{AIP}_i(1)$ lies within one-tenth of the interquartile range from the respective sample median used for stratum assignment. Panel (b) re-runs the analysis using strata defined from bootstrapped random-forest predictions. We draw $K = 100$ bootstrap samples of the 500 trees (with replacement), reclassify observations into strata for each draw, and re-estimate the stratum-specific treatment effects. The figure reports average treatment effects across replications. Standard errors incorporate both within-replication estimation uncertainty and between-replication variation.

S9 Estimates underlying the main figures

Table S12: Treatment Effects on GenAI Impact Perceptions (by country splits and main indicators)

	(1) AIP (full)	(2) AIP (US)	(3) AIP (DE)	(4) AIP (IT)	(5) Overall Work Impact	(6) Productivity vs. Redundancy	(7) Threat vs. Opportunity
Treated	0.147***	0.213***	0.101***	0.110**	0.0617***	0.211***	0.167***
	(0.0235)	(0.0401)	(0.0357)	(0.0479)	(0.0209)	(0.0384)	(0.0279)
Observations	5,983	2,375	2,192	1,416	5,983	5,983	5,983
R-squared	0.085	0.070	0.070	0.071	0.090	0.071	0.075

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table S13: Treatment Effects on GenAI Impact Perceptions (items of)

	(1) Salary and Career	(2) Job security	(3) Autonomy	(4) Task Enjoyment	(5) Occupational Pride
Treated	0.00871	0.00248	0.0820***	0.127***	0.0880***
	(0.0263)	(0.0278)	(0.0256)	(0.0251)	(0.0263)
Observations	5,983	5,983	5,983	5,983	5,983
R-squared	0.075	0.058	0.086	0.078	0.069

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table S14: Treatment Effects for Redistribution Outcomes

	(1) Reduce Income Differences	(2) Early Retirement	(3) Unemployment Benefits	(4) Active Labour Market Policies	(5) Minimum Wage	(6) Tax on Super-rich
Treated	-0.114**	-0.0159	0.122**	0.0488	-0.00506	-0.179
	(0.0564)	(0.0510)	(0.0523)	(0.0412)	(0.0534)	(0.131)
Treated # Pessimistic	0.262***	0.0909	-0.0706	-0.0588	0.0651	0.308
	(0.0961)	(0.0896)	(0.108)	(0.0799)	(0.0947)	(0.234)
Treated # Always Pessimistic	0.107	0.0975	-0.150*	-0.0422	-0.00903	0.225
	(0.0835)	(0.0741)	(0.0768)	(0.0652)	(0.0835)	(0.204)
Treated # Optimistic	-0.0865	-0.105	-0.410***	-0.0848	-0.114	-0.232
	(0.0957)	(0.0871)	(0.0946)	(0.0736)	(0.0920)	(0.213)
Observations	5,983	5,983	5,983	5,983	5,983	5,983
R-squared	0.141	0.138	0.151	0.156	0.140	0.126

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table S15: Treatment Effects for New Policies

	(1) Automation Tax	(2) Automation Tax / Redistribute	(3) GenAI Regulation	(4) Codetermination	(5) Employment Protection
Treated	-0.00906	-0.240*	-0.0743	-0.0829*	-0.0480
	(0.0594)	(0.125)	(0.0530)	(0.0479)	(0.126)
Treated # Pessimistic	0.512***	1.136***	0.451***	0.363***	0.714***
	(0.107)	(0.219)	(0.0969)	(0.0833)	(0.221)
Treated # Always Pessimistic	-0.0139	0.313*	0.248***	0.260***	0.0704
	(0.0820)	(0.185)	(0.0724)	(0.0692)	(0.189)
Treated # Optimistic	-0.503***	-0.541**	-0.309***	-0.218***	-0.689***
	(0.100)	(0.214)	(0.0909)	(0.0812)	(0.221)
Observations	5,983	5,983	5,983	5,983	5,983
R-squared	0.128	0.142	0.176	0.131	0.134

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

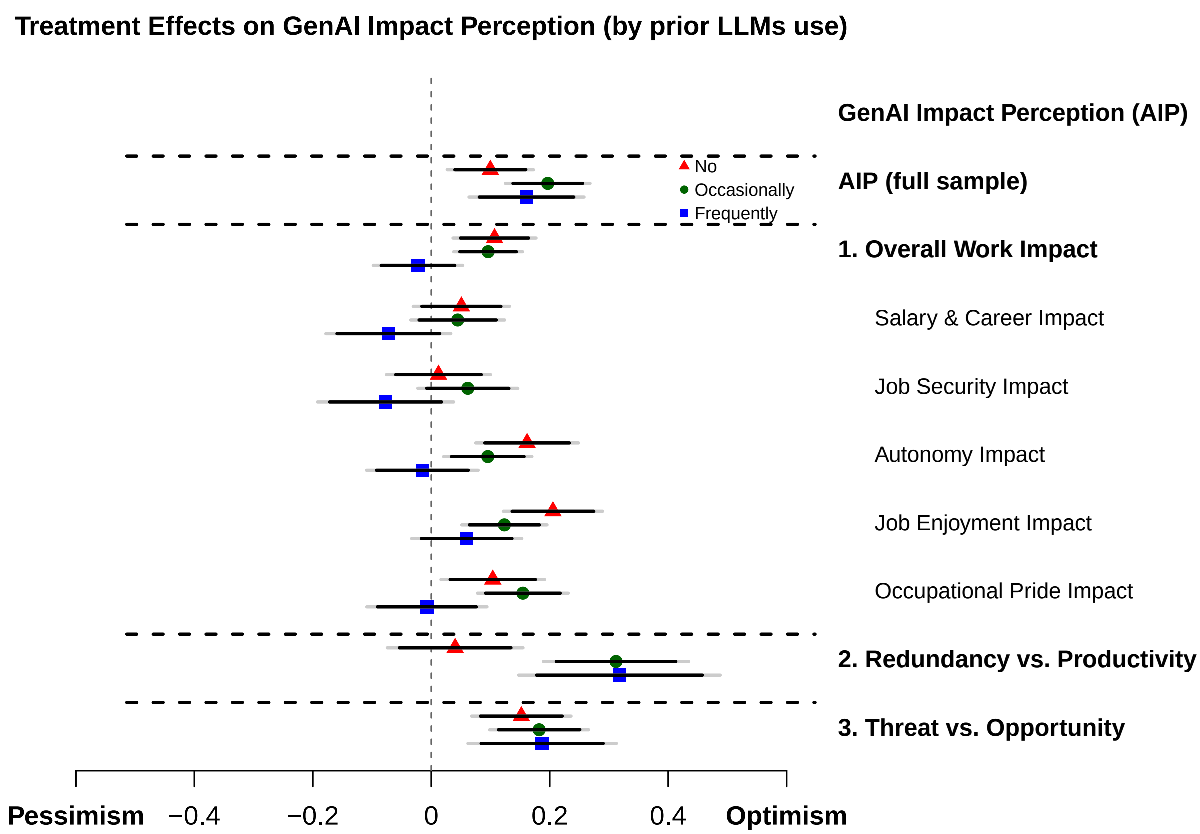
Table S16: Treatment Effects for Economic Nationalism and Left-Wing Authoritarianism

	(1) Limiting Imports	(2) Barriers to Immigration	(3) Index LWA	(4) Climate Authoritarianism
Treated	-0.0914*	-0.0392	-0.0170	-0.0418***
	(0.0549)	(0.0563)	(0.0103)	(0.0141)
Treated # Pessimistic	0.403***	0.00240	0.0331	0.0846***
	(0.108)	(0.1000)	(0.0206)	(0.0285)
Treated # Always Pessimistic	0.153**	0.112	0.0310*	0.0710***
	(0.0746)	(0.0807)	(0.0162)	(0.0209)
Treated # Optimistic	-0.244**	0.279***	-0.0436**	-0.0222
	(0.0976)	(0.0945)	(0.0184)	(0.0242)
Observations	5,983	5,983	5,983	5,983
R-squared	0.142	0.163	0.151	0.141

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

S10 Heterogeneity analysis

Figure S3: **Treatment effects across three groups of respondents defined by prior use of Large Language Models (LLMs) such as ChatGPT.**



Note: We re-estimate the main specification in Figure 1 including interactions between the treatment indicator and the discrete item `llmuse` capturing whether respondents have used large language models like ChatGPT for their work or other activities, with three categories: "Yes, frequently"; "Yes, occasionally"; "No". The main result on the overall AIP index is stable across the three groups. Some degree of heterogeneity across groups emerges when considering the estimated effects on specific items.

Table S17: Treatment Effect Heterogeneity for Redistribution Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome: ⟶	Reduce Income Differences	Reduce Income Differences	Early Retirement	Unemployment Benefits	Unemployment Benefits	Minimum Wage	Tax on Super-rich	Tax on Super-rich
Heterogeneity by variable HV: ⟶	*income*	*education*	*age*	*income*	*education*	*income*	*ideological self-placement*	*left party vote*
HV	0.223	0.126	-0.0180***	-0.0892	-0.463	0.175	1.631	0.762
	(0.51)	(0.34)	(-6.09)	(-0.23)	(-1.52)	(0.50)	(1.46)	(1.05)
Treated	-0.114	-0.00790	0.0854	0.0919	0.0537	0.0353	-0.0735	-0.240
	(-1.20)	(-0.11)	(0.55)	(1.05)	(0.83)	(0.46)	(-0.34)	(-1.69)
Treated # HV	0.0353	-0.128	-0.00191	-0.0847	-0.0538	-0.000842	-0.0803	0.375
	(0.30)	(-1.23)	(-0.48)	(-0.79)	(-0.56)	(-0.01)	(-0.31)	(1.54)
Pessimistic	0.0414	-0.1000	-0.455*	-0.154	-0.202*	-0.180	-0.321	-0.166
	(0.38)	(-1.07)	(-2.27)	(-1.48)	(-2.34)	(-1.78)	(-1.00)	(-0.84)
Always Pessimistic	-0.0219	-0.0226	-0.533**	-0.169*	-0.179**	-0.262***	-0.335	-0.136
	(-0.24)	(-0.30)	(-3.25)	(-2.06)	(-2.69)	(-3.30)	(-1.40)	(-0.91)
Optimistic	0.224*	0.273**	-0.430*	0.0328	0.0704	-0.0379	0.713**	0.617***
	(2.04)	(3.00)	(-2.10)	(0.31)	(0.81)	(-0.43)	(2.83)	(3.65)
Pessimistic # HV	-0.223	0.0656	0.0106*	-0.117	0.0469	0.0544	0.338	0.192
	(-1.53)	(0.49)	(2.23)	(-0.85)	(0.36)	(0.40)	(0.89)	(0.60)
Always Pessimistic # HV	-0.105	-0.0852	0.0121**	-0.00997	0.0418	-0.00913	0.537	0.460
	(-0.86)	(-0.76)	(3.15)	(-0.09)	(0.41)	(-0.08)	(1.89)	(1.80)
Optimistic # HV	-0.0553	-0.178	0.0146**	0.106	0.0421	0.0468	-0.340	-0.691*
	(-0.39)	(-1.36)	(2.91)	(0.78)	(0.33)	(0.38)	(-1.09)	(-2.27)
Treated # Pessimistic	0.0650	0.102	-0.0654	0.0871	0.0466	0.0406	0.501	0.480
	(0.42)	(0.80)	(-0.23)	(0.59)	(0.40)	(0.28)	(1.18)	(1.77)
Always Treated # Pessimistic	0.143	0.00238	-0.0155	-0.0912	-0.0643	-0.0137	-0.0201	0.264
	(1.17)	(0.02)	(-0.07)	(-0.81)	(-0.72)	(-0.13)	(-0.06)	(1.27)
Treated # Optimistic	-0.0808	-0.217	0.0116	-0.227	-0.234	-0.217	-0.790	-0.587*
	(-0.54)	(-1.68)	(0.04)	(-1.56)	(-1.91)	(-1.63)	(-1.90)	(-2.19)
Treated # Pessimistic # HV	0.0879	-0.0131	0.00349	0.0645	0.105	-0.0641	-0.214	-0.537
	(0.43)	(-0.07)	(0.52)	(0.33)	(0.57)	(-0.33)	(-0.42)	(-1.21)
Always Treated # Pessimistic # HV	-0.149	0.102	0.00231	0.0637	0.0598	-0.0892	0.255	-0.640
	(-0.90)	(0.67)	(0.43)	(0.42)	(0.42)	(-0.57)	(0.65)	(-1.88)
Treated # Optimistic # HV	-0.124	0.140	-0.00361	-0.203	-0.115	-0.0883	0.902	0.819
	(-0.59)	(0.72)	(-0.52)	(-1.02)	(-0.63)	(-0.45)	(1.85)	(1.92)
Constant	3.618***	3.505***	4.723***	3.627***	3.708***	4.213***	6.502***	8.451***
	(11.91)	(12.25)	(22.04)	(14.17)	(16.46)	(18.51)	(10.86)	(13.02)
N	5275	5983	5983	5275	5983	5275	5369	5983

t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table S18: Treatment Effect Heterogeneity for New Policies

Outcome: ⟶	(1) Automation Tax	(2) Automation Tax / Redistribute	(3) Gen AI Regulation	(4) Codetermination	(5) Employment Protection
Ideological Self-Placement	0.470	0.894	-0.603	-0.279	0.158
	(1.01)	(0.84)	(-1.69)	(-0.56)	(0.13)
Treated	0.102	-0.0528	0.0518	-0.0353	-0.0250
	(1.15)	(-0.26)	(0.67)	(-0.45)	(-0.12)
Treated # Ideological Self-Placement	-0.151	-0.0939	-0.0844	-0.00196	-0.00172
	(-1.33)	(-0.37)	(-0.87)	(-0.02)	(-0.01)
Pessimistic	-0.235	-0.326	0.134	-0.388**	-0.495
	(-1.86)	(-1.06)	(1.21)	(-3.29)	(-1.64)
Always Pessimistic	0.215*	-0.293	0.520***	-0.243**	0.0806
	(2.26)	(-1.31)	(6.34)	(-2.88)	(0.37)
Optimistic	0.558***	1.020***	0.719***	0.257**	0.875***
	(5.36)	(4.38)	(8.04)	(2.76)	(3.71)
Pessimistic # Ideological Self-Placement	0.136	-0.370	0.0160	0.276*	0.353
	(0.89)	(-1.00)	(0.12)	(1.98)	(0.98)
Always Pessimistic # Ideological Self-Placement	-0.0145	-0.0638	-0.0273	0.129	0.0298
	(-0.12)	(-0.23)	(-0.27)	(1.24)	(0.11)
Optimistic # Ideological Self-Placement	-0.234	-0.912**	-0.242*	-0.169	-0.359
	(-1.75)	(-3.03)	(-2.04)	(-1.43)	(-1.17)
Treated # Pessimistic	0.509**	0.857*	0.395**	0.438**	1.178**
	(3.08)	(2.14)	(2.62)	(2.76)	(2.96)
Treated # Always Pessimistic	-0.0630	0.0860	0.0164	0.209	-0.133
	(-0.49)	(0.28)	(0.15)	(1.83)	(-0.44)
Treated # Optimistic	-0.852***	-1.359***	-0.593***	-0.267	-1.066**
	(-4.99)	(-3.48)	(-4.12)	(-1.79)	(-2.77)
Treated # Pessimistic # Ideological Self-Placement	0.0797	0.113	-0.0145	-0.153	-0.595
	(0.39)	(0.23)	(-0.08)	(-0.81)	(-1.23)
Treated # Always Pessimistic # Ideological Self-Placement	0.124	0.113	0.151	-0.0210	0.350
	(0.77)	(0.30)	(1.09)	(-0.15)	(0.93)
Treated # Optimistic # Ideological Self-Placement	0.473*	1.058*	0.358*	-0.00248	0.508
	(2.26)	(2.22)	(1.98)	(-0.01)	(1.07)
Constant	0.131	5.746***	0.381*	0.624***	5.211***
	(0.66)	(10.16)	(2.00)	(3.30)	(9.03)
N	5369	5369	5369	5369	5369

t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table S19: Treatment Effect Heterogeneity by Gender for New Policies

	(1)	(2)	(3)	(4)	(5)
Outcome: $\longrightarrow$	Automation Tax	Automation Tax / Redistribute	Gen AI Regulation	Codetermination	Employment Protection
Gender	-0.282	-1.030	-0.0980	-0.429	-0.465
	(-1.08)	(-1.41)	(-0.36)	(-1.87)	(-0.66)
Treated	-0.0143	-0.151	-0.0194	-0.0672	-0.142
	(-0.21)	(-0.96)	(-0.32)	(-1.10)	(-0.91)
Treated # Gender	0.0414	0.0874	0.00715	0.110	0.314
	(0.38)	(0.37)	(0.08)	(1.22)	(1.32)
Pessimistic	-0.192*	-0.592**	0.219*	-0.205*	-0.292
	(-2.02)	(-2.58)	(2.51)	(-2.39)	(-1.33)
Always Pessimistic	0.241**	-0.345	0.553***	-0.206**	0.0343
	(3.20)	(-1.94)	(8.12)	(-3.05)	(0.20)
Optimistic	0.460***	0.512**	0.639***	0.197**	0.932***
	(5.48)	(2.69)	(8.33)	(2.67)	(4.99)
Pessimistic # Gender	0.212	0.199	-0.141	0.0861	0.257
	(1.59)	(0.64)	(-1.19)	(0.74)	(0.84)
Always Pessimistic # Gender	-0.0495	0.0696	-0.180	0.0653	0.177
	(-0.45)	(0.29)	(-1.91)	(0.71)	(0.72)
Optimistic # Gender	-0.0850	-0.0394	-0.204	-0.0801	-0.575*
	(-0.68)	(-0.14)	(-1.79)	(-0.75)	(-1.98)
Treated # Pessimistic	0.615***	1.092***	0.328**	0.288*	0.979***
	(4.82)	(3.62)	(2.74)	(2.46)	(3.34)
Always Treated # Pessimistic	-0.0587	0.149	0.0842	0.215*	0.105
	(-0.56)	(0.61)	(0.91)	(2.29)	(0.43)
Treated # Optimistic	-0.775***	-1.159***	-0.578***	-0.486***	-1.304***
	(-5.87)	(-3.78)	(-5.00)	(-4.12)	(-4.41)
Treated # Pessimistic # Gender	-0.213	-0.410	0.0831	-0.0595	-0.661
	(-1.16)	(-0.97)	(0.50)	(-0.37)	(-1.57)
Always Treated # Pessimistic # Gender	0.0915	-0.105	0.114	-0.0572	-0.204
	(0.61)	(-0.31)	(0.88)	(-0.45)	(-0.59)
Treated # Optimistic # Gender	0.441*	0.798	0.402*	0.316*	0.987*
	(2.35)	(1.89)	(2.45)	(1.97)	(2.29)
Constant	0.762***	7.374***	0.360*	0.977***	6.119***
	(3.76)	(14.07)	(2.07)	(5.80)	(11.35)
N	5983	5983	5983	5983	5983

t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table S20: Treatment Effect Heterogeneity for Economic Nationalism

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome: ⟶	Limiting Imports	Limiting Imports	Barriers to Immigration	Barriers to Immigration	Barriers to Immigration	Barriers to Immigration	Barriers to Immigration
Heterogeneity by variable HV: ⟶	*ideological self-placement*	*conservative vote*	*ideological self-placement*	*conservative vote*	*radical right vote*	*right-wing authoritarianism*	*social dominance orientation*
HV	-1.412***	1.088**	-1.343***	0.674*	0.804*	0.105*	0.130**
	(-4.44)	(3.22)	(-4.27)	(2.01)	(2.40)	(2.28)	(2.92)
Treated	0.00658	-0.0121	-0.0226	0.102	0.0649	0.0587	-0.0561
	(0.08)	(-0.21)	(-0.27)	(1.79)	(1.17)	(1.04)	(-0.87)
Treated # HV	-0.0927	-0.0345	0.0837	-0.264*	-0.190	-0.132*	-0.0955
	(-0.90)	(-0.33)	(0.79)	(-2.45)	(-1.59)	(-2.09)	(-1.60)
Pessimistic	-0.00703	0.00306	0.185	0.185**	0.169*	0.175**	0.201*
	(-0.06)	(0.04)	(1.50)	(2.59)	(2.45)	(2.62)	(2.38)
Always Pessimistic	0.225**	0.0392	0.231*	0.325***	0.286***	0.306***	0.244***
	(2.70)	(0.67)	(2.48)	(5.37)	(4.86)	(5.41)	(3.51)
Optimistic	0.520***	0.134	-0.243*	0.178*	0.0975	0.0652	-0.102
	(5.27)	(1.79)	(-2.05)	(2.36)	(1.33)	(0.92)	(-1.19)
Pessimistic # HV	-0.118	-0.0804	0.0343	0.0799	0.109	0.0458	0.0676
	(-0.82)	(-0.58)	(0.23)	(0.57)	(0.64)	(0.55)	(0.91)
Always Pessimistic # HV	-0.294**	0.150	0.0641	-0.0827	-0.0207	-0.0558	-0.0371
	(-2.74)	(1.37)	(0.55)	(-0.71)	(-0.16)	(-0.80)	(-0.60)
Optimistic # HV	-0.592***	0.359**	0.452**	-0.470**	-0.307	-0.184*	-0.162*
	(-4.62)	(2.63)	(3.10)	(-2.98)	(-1.70)	(-2.21)	(-2.11)
Treated # Pessimistic	0.427**	0.158	-0.0920	-0.0617	-0.0892	-0.0639	-0.0860
	(2.76)	(1.52)	(-0.54)	(-0.60)	(-0.90)	(-0.66)	(-0.73)
Treated # Always Pessimistic	0.00832	0.0826	0.145	-0.0492	-0.00930	0.00284	0.130
	(0.07)	(1.04)	(1.16)	(-0.60)	(-0.12)	(0.04)	(1.34)
Treated # Optimistic	-0.674***	-0.277**	0.544**	-0.0192	0.0686	0.124	0.393**
	(-4.27)	(-2.65)	(3.22)	(-0.19)	(0.69)	(1.28)	(3.14)
Treated # Pessimistic # HV	-0.205	0.348	0.0979	-0.0134	0.109	0.0104	-0.0854
	(-1.05)	(1.79)	(0.47)	(-0.07)	(0.46)	(0.08)	(-0.81)
Treated # Always Pessimistic # HV	0.114	-0.0338	-0.130	0.321*	0.263	0.156	0.104
	(0.78)	(-0.23)	(-0.83)	(2.06)	(1.52)	(1.52)	(1.21)
Treated # Optimistic # HV	0.683***	-0.252	-0.581**	0.758***	0.610*	0.428***	0.236*
	(3.50)	(-1.25)	(-2.83)	(3.60)	(2.53)	(3.57)	(2.14)
Constant	0.491**	-0.400	0.609**	-0.0754	-0.219	0.129	0.255
	(2.78)	(-1.43)	(2.80)	(-0.29)	(-0.84)	(0.42)	(0.83)
N	5369	5983	5369	5983	5983	5983	5983

t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

S11 Political preferences

Table S21: Warmth towards Backlash and Non-backlash Parties by Past Vote

	(1) Warmth Towards Backlash Parties	(2) Warmth Towards Non-Backlash Parties
Treated	1.862**	1.132
	(0.874)	(1.112)
$\hat{\mu}$	8.741***	6.318***
	(1.203)	(1.583)
Treated # $\hat{\mu}$	-2.817**	-0.528
	(1.392)	(1.815)
Treated # Voted for Backlash Party	3.797**	
	(1.805)	
Treated # Voted for Backlash Party # $\hat{\mu}$	-6.503***	
	(2.457)	
Treated # Voted for Non-Backlash Party		1.084
		(1.598)
Treated # Voted for Non-Backlash Party # $\hat{\mu}$		-6.802***
		(2.259)
Observations	5,983	5,983
R-squared	0.370	0.387
Mean Dep. Var.	34.81	42.71
S.D. Dep. Var.	26.63	27.20

Note: In line with Figure 3 in the main text, the estimated treatment effect is evaluated at $\hat{\mu} = -1$ for individuals pushed towards pessimism and at $\hat{\mu} = 1$ for individuals pushed towards optimism. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

GREEN
Centre for Geography, Resources, Environment, Energy and Networks
via Röntgen, 1
20136 Milano - Italia

www.green.unibocconi.eu