The Path from Automation to Populist Political Behavior

Valentina González-Rostani

Department of Political Science, University of Pittsburgh

Abstract

I investigate the impact of automation exposure on political behavior in post-industrial societies, with a specific focus on the support for populism. I ask **how and to what extent automation affects political behavior**. I contribute by:

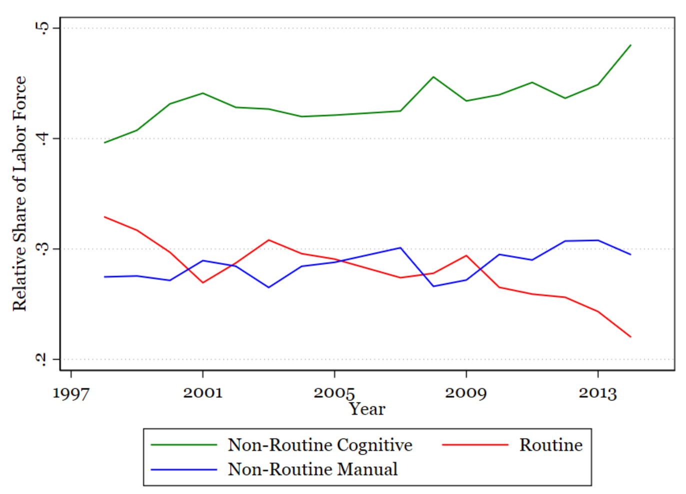
- Exploring the mechanisms through which the automation of jobs translates into party system change through its impact on cultural factors.
- Implementing a parallel encouragement design survey to unpack the causal mechanisms.
- Reflecting on ways to identify causal mediation effects.

Finally, I conclude with an extension using survey data across European countries to enhance external validity (not presented in the poster).

Motivation & Hypotheses

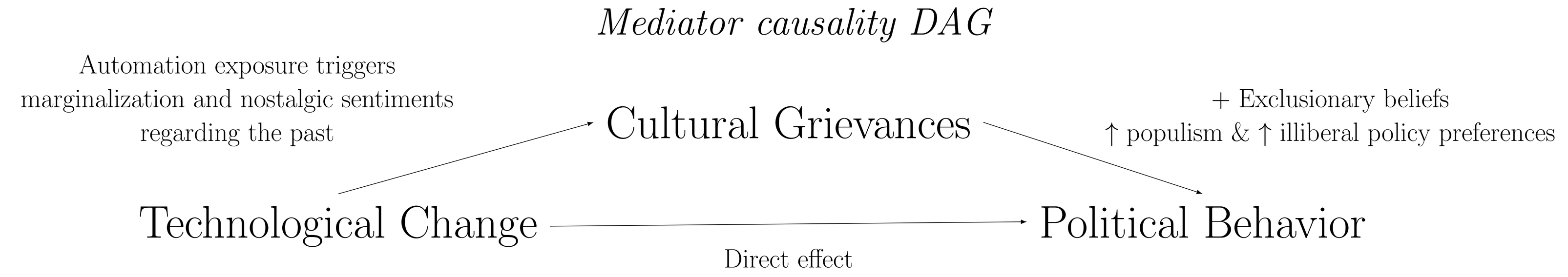
Technological changes reshape work, causing job displacement and altering the wage structure (Acemoglu and Restrepo, 2022). Automation is predicted to displace a significant number of jobs (47% in the US), particularly affecting middle-wage workers (Frey and Osborne, 2017). These transformations contribute to rising inequality, labor market polarization, and the shrinking of the middle class, which likely will have profound political consequences. As such, I argue:

Fig.1: Relative Share of Labor Force



- Technological change will affect political behavior (e.g., populist attitudes, illiberal policy preferences)
- It will do that partially by triggering feelings of marginalization and nostalgia (cultural grievances), which are coping mechanisms to alleviate the insecurity generated by automation of jobs.

Mediator causality DAG



Background

Social scientists have increasingly prioritized causal identification. However, many of these efforts primarily concentrate on establishing the existence of a relationship between a treatment and a political outcome, without delving into the underlying mechanisms, known as the "black box" of causality (Imai et al., 2011; Brady and Collier, 2010).

For example, we know from previous studies that exposure to automation is associated with political support for the radical right (e.g, Frey et al., 2017), but the underlying mechanisms have not been well-established (Gallego and Kurer, 2022). In this study, I aim to contribute by exploring these mechanisms.

Currently, researchers interested in opening the "black box" can rely on tools such as:

- Causal chain analysis (e.g., Spencer et al., 2005), which may be biased when there are heterogeneous treatment effects.
- Causal mediation analysis (Imai et al., 2013) addresses heterogeneous treatment groups but relies on sharp bounds (Balke and Pearl, 1997), which are not always informative about the direction of causation.

In this work, I employ a parallel encouragement design and discuss challenges.

Methodological Approach

I employ causal mediation analysis (Imai et al., 2011) to decompose the average treatment effect (ATE) into the average causal mediated effect (ACME) and the average direct effect (ADE), such that:

$$ATE = E[Y(1) - Y(0)]$$
$$ACME(t) = E[Y(t, M(1)) - Y(t, M(0))]$$
$$ADE(t) = E[Y(1, M(t)) - Y(0, M(t))]$$

with $Y(t)$ representing the expected outcome of interest under a given treatment status ($t \in [0, 1]$), and $M(t)$ representing the value of the mediator. In my analysis,

- The outcomes are indicators of political behavior and policy preferences. $Y \in$ [*populist attitudes, support for populist right, trade protectionism, anti-immigration*]
- The mediators are indicators of cultural grievances. $M \in$ [*nostalgia, marginalization*]

I implement this analysis in two different ways: **Study 1**: Parallel encouragement survey experiment design; and **Study 2**: Causal mediation analysis using the European Social Survey (ESS), 2002-2016.

Study 1 - Parallel Encouragement Survey Experiment Design

Each subject will be randomly assigned to one of two groups:

- Group 1: Randomization of treatment conditions only.
- Group 2: Randomization of treatment conditions and encouragement of the mediator.

To date, only a handful of studies have utilized survey experiments to explore the effects of automation of jobs on individuals' attitudes (e.g., Wu, 2022; Mutz, 2021). My **design is novel** in several ways accounting for some of the limitations of previous studies:

- The survey experiment proposes a **realistic task-oriented experience** about news, in which subjects are asked to evaluate two news stories under the guise of helping a news website (Lelkes and Westwood, 2017). My goal is to provide an ecologically valid task-oriented experience.
- Instead of comparing different risks, I compare exposure to **job automation risk** with a **neutral technology** control condition.
- The design allows me to better explore possible mediated pathways.

Manipulation of the treatment: in all the cases, participants were exposed to 2 news articles.

- $t = 1$ News discussing the automation of jobs, focusing on either manufacturing jobs (robots) or white-collar workers dealing with artificial intelligence (AI).
- $t = 0$ News related to technological advancement was presented in a neutral manner without discussing its impact on jobs.

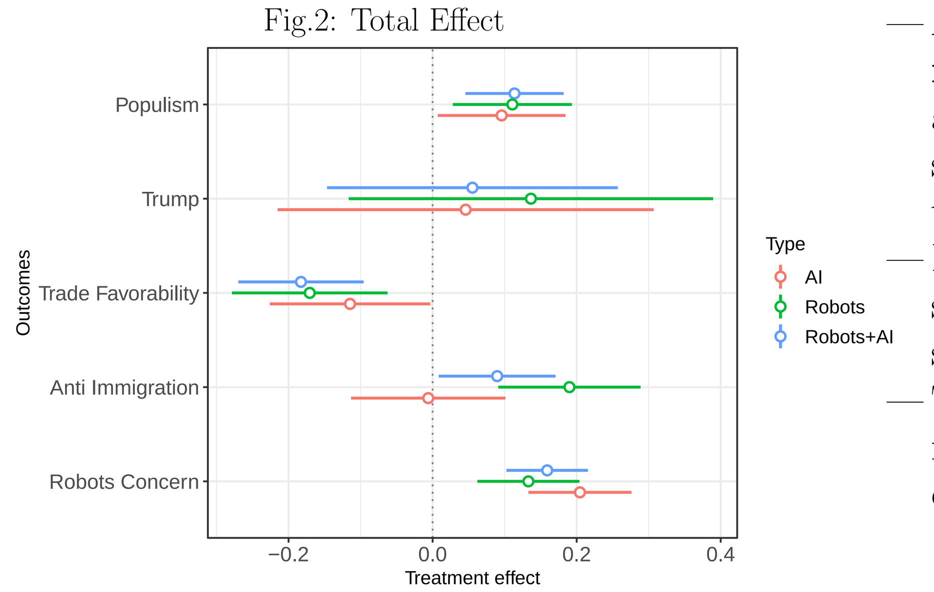
Encouragement of mediators: Subjects completed a 90-second writing exercise (autobiographical emotional memory task). The frame was that a news organization was deciding whether to add a new section called "letters to the editor." Participants were given a prompt to think about a time in their life that made them feel a particular emotion (nostalgia or marginalization).

Implementation

- The survey got **IRB approval**, and was **pre-registered** at `https://osf.io/3kdpq`, with power calculations based on the pilot using Declare Design framework (Blair et al., 2019).
- Fielded using CloudResearch-Mturk toolkit between May 23, and May 29, 2023, collecting 3133 responses from adults-US citizens, part of the workforce (currently working or looking for a job).
- I implemented several measures to ensure data quality, including CAPTCHA, location screening, attention checks, manipulation checks, and respondents with a greater 90% of rate of survey approval HITS on CloudResearch, among others.

Results: Total Effect of Automation on Political Behavior $[T \to Y]$

Fig.2: Total Effect

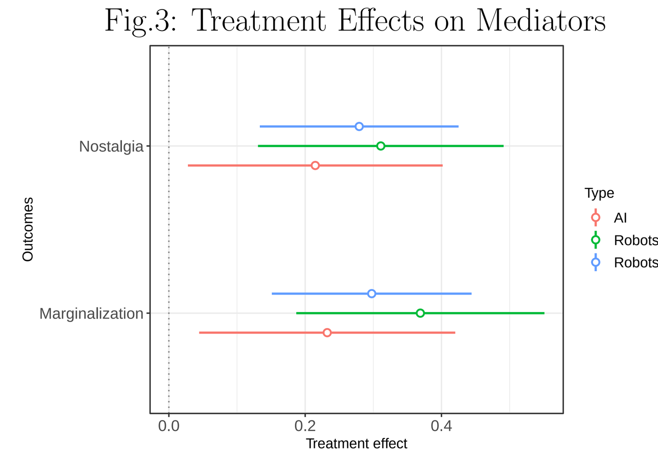


- As expected, **exposure to automation risks** increases populist attitudes, specifically anti-elitism, and fosters support for illiberal policies such as reducing immigration and opposition to trade.
- Both the **robots and AI** treatments exhibit similar average treatment effects (ATE), with AI showing a slightly lower effect.
- The treatment effectively triggers concerns about robots, as evidenced by the **manipulation check**.

Results: Treatment Effect on Mediator $[T \to M]$

Consistent with my theoretical expectations, subjects exposed to the job automation treatment condition reported an increase in feelings of nostalgia towards the past and a heightened sense of marginalization.

Fig.3: Treatment Effects on Mediators



Analyzing Mediator Effects on Outcomes $[M \to Y]$

I implement three strategies:

1. **Matching** on covariates, examining the mediator effect based on assignment to the mediator's encouragement.
2. **Intention to Treat**, instrumenting the mediator through assignment to the encouragement, and considering compliers with high values on the mediators for those who completed the task (wrote more words than the median).
3. **Sharp Bounds** following Balke and Pearl (1997), after implementing linear programming optimization techniques. Sharp bounds are defined as follows:

$$= max \begin{Bmatrix} p_{00.0} + p_{11.1} - 1, \\ p_{00.1} + p_{11.1} - 1, \\ p_{11.0} + p_{00.1} - 1, \\ p_{00.0} + p_{11.0} - 1, \\ 2p_{00.0} + p_{11.0} + p_{10.1} + p_{11.1} - 2, \\ p_{00.0} + 2p_{11.0} + p_{00.1} + p_{01.1} - 2, \\ p_{10.0} + p_{11.1} + 2p_{00.1} + p_{11.1} - 2, \\ p_{00.0} + p_{01.0} + p_{00.1} + 2p_{11.1} - 2 \end{Bmatrix}$$

$$= min \begin{Bmatrix} 1 - p_{10.0} - p_{01.1}, \\ 1 - p_{01.0} - p_{10.1}, \\ 1 - p_{01.0} - p_{10.1}, \\ 1 - p_{01.1} - p_{10.1}, \\ 2 - 2p_{01.0} - p_{10.0} - p_{10.1} - p_{11.1}, \\ 2 - p_{01.0} - 2p_{10.0} - p_{00.1} - p_{01.1}, \\ 2 - p_{10.0} - p_{11.0} - 2p_{01.1} - p_{10.1}, \\ 2 - p_{00.0} - p_{01.0} - p_{01.0} - p_{01.1} - 2p_{10.1} \end{Bmatrix}$$

Results: Mediator Effects on Outcome $[M \to Y]$

Fig.4: Mediator Effects on Outcomes (ITT & Bounds)

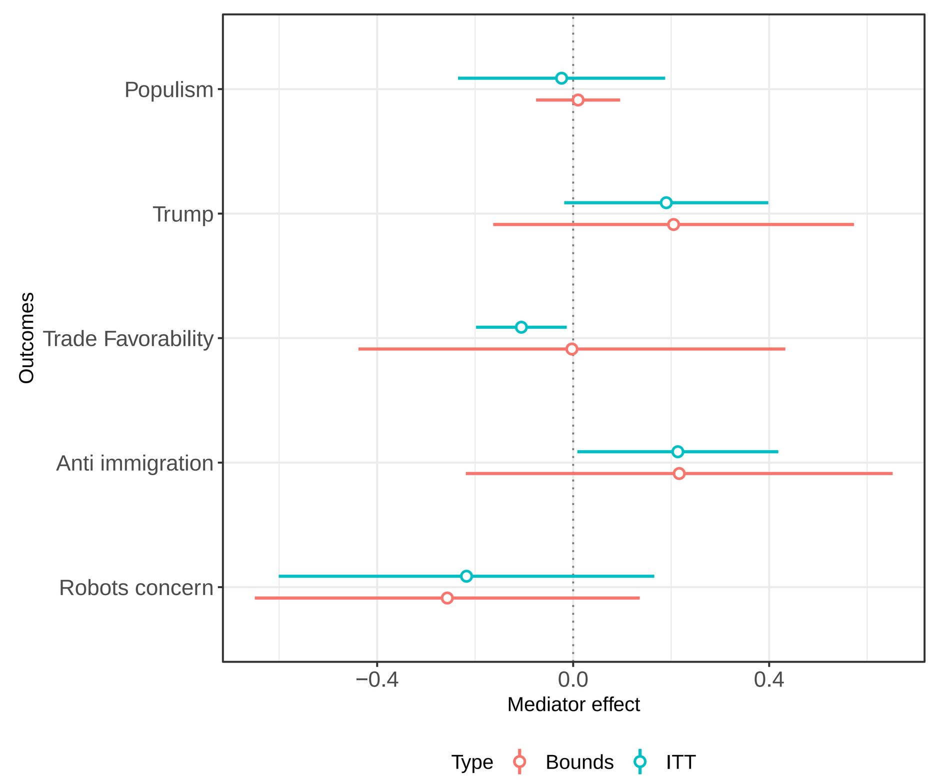
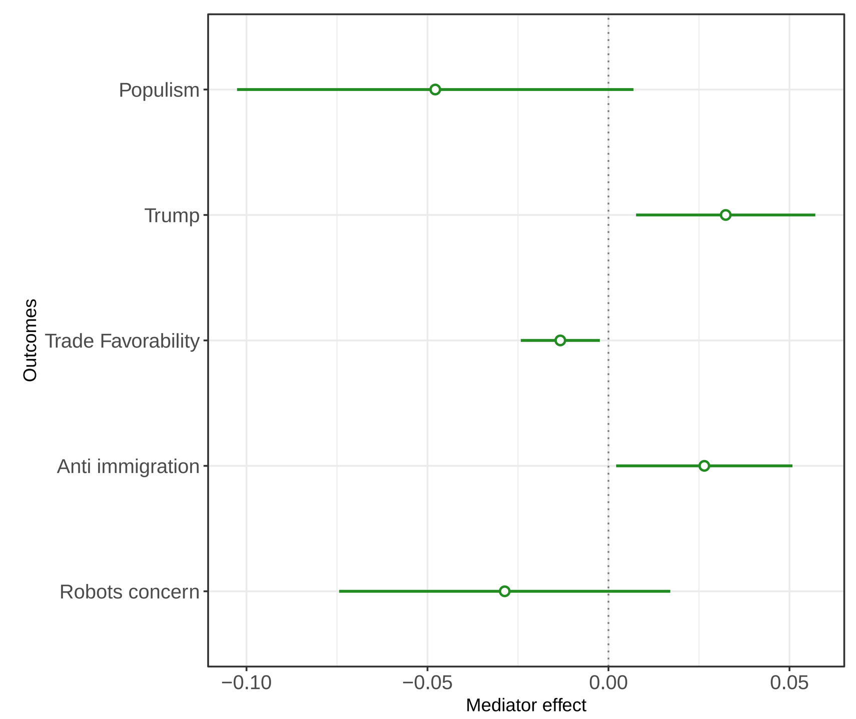


Fig. 5: Encouragement Effects on Outcomes (Matching)



- As expected, the encouragement of cultural grievances increases support for Trump, opposition to trade, and xenophobia.
- Using Matching and ITT, I obtain similar significance tests results.
- Bounds are not very informative due to significant uncertainty.
- As a **manipulation check**, the encouragement effectively triggered the mediators, resulting in an approximate 15% increase.

Robustness checks:

- I estimated the **ACME based on observed values of nostalgia** (model-based inference), providing support for the hypothesis that nostalgia and marginalization play a role in explaining changes in political behavior.
- To enhance **external validity**, I conducted **Study 2** consisting of causal mediation analysis using the European Social Survey, and the results are consistent.

Conclusions

The question of how to explain the causal relationship between political-economic phenomena and political behavior is complex. The evidence presented here suggests three broad conclusions:

1. **The exposure to automation risks triggers populist attitudes, illiberal policy preferences, nostalgia, and feelings of marginalization.**
2. **There is some evidence of a causal pathway from automation to political behavior through cultural grievances, driven by nostalgia and marginalization.**
3. **There are challenges for the identification of causal mechanisms when the researcher only manipulates the encouragement (e.g., weak effects).**

Moving forward, I will focus on the following areas:

- Deriving standard errors for the causal chain ACME.
- Exploiting subjects' open-ended responses to re-assess compliance.
- Developing techniques to identify type heterogeneity to address treatment heterogeneity and calculate type-specific ACMEs.

References

Acemoglu, D. and P. Restrepo (2022). Tasks, Automation, and the Rise in US Wage Inequality. *Econometrica*, 75.

Balke, A. and J. Pearl (1997). Bounds on Treatment Effects from Studies with Imperfect Compliance. *Journal of the American Statistical Association 92*(439), 1171–1176.

Blair, G., J. Cooper, A. Coppock, and M. Humphreys (2019). Declaring and Diagnosing Research Designs. *American Political Science Review 113*(3), 838–859. Publisher: Cambridge University Press.

Brady, H. E. and D. Collier (2010). *Rethinking Social Inquiry: Diverse Tools, Shared Standards.* Rowman & Littlefield Publishers.

Frey, C. B., T. Berger, and C. Chen (2017). Political machinery: Automation anxiety and the 2016 US presidential election. *University of Oxford*.

Frey, C. B. and M. A. Osborne (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change 114*, 254–280.

Gallego, A. and T. Kurer (2022). Workplace automation and digitalization: Implications for political behavior. *Annual Review of Political Science*.

Imai, K., L. Keele, D. Tingley, and T. Yamamoto (2011, November). Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies. *American Political Science Review 105*(4), 765–789. Publisher: Cambridge University Press.

Imai, K., D. Tingley, and T. Yamamoto (2013, January). Experimental designs for identifying causal mechanisms: *Experimental Designs for Identifying Causal Mechanisms. Journal of the Royal Statistical Society: Series A (Statistics in Society) 176*(1), 5–51.

Lelkes, Y. and S. J. Westwood (2017, April). The Limits of Partisan Prejudice. *The Journal of Politics 79*(2), 485–501.

Mutz, D. C. (2021). (Mis)Attributing the Causes of American Job Loss: The Consequences of Getting It Wrong. *Public Opinion Quarterly 85*(1), 101–122.

Spencer, S. J., M. P. Zanna, and G. T. Fong (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology 89*(6), 845–851.

Wu, N. (2022). "Restrict foreigners, not robots": Partisan responses to automation threat. *Economics & Politics*.

PolMeth XL, Stanford University, July 9th 2023

Email: mag384[at]pitt.edu

Web: gonzalez-rostani.com