Draft: Please Do Not Circulate

Robots Replacing Trade Unions: Novel Data and Evidence from Western Europe[*]

PAOLO AGNOLIN[†] MASSIMO ANELLI[‡] ITALO COLANTONE[§]
PIERO STANIG[¶]

November 2024

Abstract

Historically, labor unions have played a crucial role in liberal democracies by hindering the increasing wage inequality, by channeling political demands and discontent into an organized voice, and by linking blue-collar constituencies to mainstream left parties. However, the importance and effectiveness of unions in the democratic process have progressively diminished in the last decades, combined with an atomization of political demands. We suggest that technological change, and robotization in particular, have directly contributed to weakening the role of unions. We employ novel granular data, at the subnational and sector level, on union density in Western Europe over two decades, to estimate the impact of industrial robot adoption on unionization rates. We find that regions more exposed to automation experience a decrease in union density. The decline in unionization occurs not so much through an erosion of union strength within sectors directly affected by automation, but rather through a broader labor market shift away from traditionally unionized industries toward less unionized sectors. This evidence contributes to explain why technologically-driven economic grievances tend to express a decidedly right-wing character and do not favor pro-redistribution left-wing parties.

[*]Research supported by a Fondazione CARIPLO Science and Technology Grant.
Research Assistance Acknowledgments: Andrea Cancellieri, Simone Foresti, Juan David Garcia Gonzalez, Maria Elena Lasiu, Mariadolores Schiavone and Jakob Wall

[†]Bocconi University, Dondena Centre for Research on Social Dynamics, and Duke University. Contact: paolo.agnolin@unibocconi.it.

[‡]Bocconi University, CESifo, IZA, Dondena Centre for Research on Social Dynamics. Contact: massimo.anelli@unibocconi.it.

[§]Bocconi University, Baffi-Carefin Research Centre, CESifo and FEEM. Contact: italo.colantone@unibocconi.it.

[¶]Bocconi University, Dondena Centre for Research on Social Dynamics, Yale-NUS and NUS. Contact: piero.stanig@unibocconi.it.

1 Introduction

The wave of automation in recent decades has brought productivity and welfare gains, but it has also created substantial distributional effects. Workers in regions historically specialized in industries that adopted more robots, and those whose skills were more easily substituted by new technologies, have been particularly penalized. Traditional, relatively well-paid and stable jobs, prevalently in the manufacturing sector, have largely disappeared, while new jobs have been created in service sectors or in new sectors (i.e. the Gig economy sector). These new jobs, however, often come with lower earning potential, temporary contracts, and minimal employment protection. Overall, automation has exacerbated earning inequality and triggered economic discontent of the "losers" of this process, which in turn triggered important political shifts. Research by Anelli *et al.* (2021) demonstrates that exposure to robot adoption increases support for nationalist and radical-right parties at both regional and individual levels.

Throughout the 20th century, labor unions have played a crucial role in liberal democracies by hindering the increasing wage inequality and by channeling political demands and discontent into an organized voice. With technological change and automation polarizing wages, unions could have played a crucial role in cushioning the unequal distribution of welfare gains. However, the advent of the new century has seen a diminished importance and effectiveness of unions in the democratic process, combined with an atomization of political demands.

In this paper, we suggest that technological change, and robotization in particular, have directly contributed to weakening the role of unions. We employ novel granular data at the subnational and sector levels on union density across Western Europe over two decades to estimate the impact of industrial robot adoption on unionization rates. In this way, we shed new light on the mechanism by which regions and individuals more exposed to automation tilt towards nationalist, isolationist, and radical right parties.

On the economics side, we theorize that the technological process undermines unions by drastically reducing employment in the large highly-unionized sectors, such as manufacturing, while boosting job creation in new sectors unions' traditional reach. This shift decreases the overall proportion of unionized workers. If unions are weaker, in terms of bargaining power, in sectors more exposed to automation, then they are less effective in translating productivity gains in workers' salaries.

While studying the direct impact of technological change on worker representative bodies appears crucial to understand the structural societal changes taking place in western democracies and societies, the academic literature has so far produced limited empirical analysis to study this phenomenon. Moreover, there is slender empirical evidence about the evolution of labor unions in liberal democracies, especially in western Europe. This is most likely due to the absence of adequate data.

New, fine-grained data on union density at the subnational level is crucial to explain economic and social phenomena that are politically consequential. Automation, by shifting bargaining power toward capital owners (Kristal, 2013), likely contributed to weakening the inclusive "social market economy" model, particularly in Western Europe. The crisis of embedded liberalism underpins the realignment currently taking place in many Western advanced democracies, contributing to the rise of radical-right parties that advocate economic nationalism, isolationism, and skepticism towards redistribution and the welfare state (Colantone and Stanig, 2018).

From a voting behavior perspective, labor unions have historically served as a critical conduit between leftist political parties and blue-collar constituencies (e.g. Przeworski and Sprague, 1986). This link operated in two directions: on the one hand by inducing mainstream left parties to focus on the economic distributional issues relevant to workers (Pontusson and Rueda, 2010), and on the other mobilizing union members in support of mainstream left parties. The reduced importance of unions in the workplace led to a decreased importance of their role as brokers for mainstream social-democratic and labor parties (Piazza, 2001). The weakening of unions also made it more appealing for social democratic parties to move to the center on redistribution issues, in order to capture middle-class constituencies attracted to their cosmopolitan stances on non-economic issues (Kitschelt, 2012). This further weakened the connection between blue-collar constituencies and mainstream left parties. Recent studies, such as Balcazar (2022), show that robot adoption reduces the likelihood of political representatives voting in alignment with union interests in the United States. Moreover, research by Becher and Stegmueller (2024, 2020b) illustrates that economic shocks from global markets weaken labor unions, diminishing their influence on political representation and legislative support for compensatory policies, which exacerbates inequality and heightens dissatisfaction with democratic processes.

The weakening of unions might explain why those affected by globalization and automation are turning towards nationalist and radical-right forces (Kitschelt, 2012; Betz, 1993). This shift suggests that support for radical-right parties is lower, while support for redistributionist left parties is higher, in regions with stronger unions.

Thus far, research on trade unions have been hindered by data limitations. In this paper, we collected data on union membership at the region-by-sector level across 15 western European countries. To generate usable measures of unionization at the necessary level of detail, we employed multilevel regression with post-stratification (MrP) (Park *et al.*, 2004). MrP makes it possible to exploit information from nationally representative samples to estimate more disaggregated summaries at the sub-national level. At this stage, we use this approach to combine information from the European Social Survey (ESS) and census information from 15 Western countries over two decades (2002-2018). We further validate our estimates by comparing them with administrative data from countries with reliable subnational estimates.

Then, we employ our novel granular data on subnational and sector level union density to estimate the impact of industrial robot adoption on unionization rates. Our findings reveal that regions more exposed to technological change exhibits decreasing union density.

We estimate that an increase of one standard deviation of robot adoption leads to more than 14% decrease of union density, after accounting for country-year and region fixed effects. Through a systematic analysis of automation's impacts at regional, sectoral, and individual levels, we show that the decline in unionization is not significantly driven by the weakening of unions within sectors directly affected by automation, but rather by a broader labor market shift away from traditionally unionized industries toward less unionized sectors.

The decline in union membership due to automation carries profound political implications. Our data shows that support for mainstream left parties is generally higher in electoral district with higher union density. We argue that, as unions weaken, their ability to mobilize working-class voters and influence electoral outcomes is jeopardized. This shift helps explain why regions and voters historically aligned with social-democratic parties are increasingly turning to nationalist and radical-right alternatives. Our findings suggest that the erosion of union strength is a key factor in the broader political transformations observed in many advanced democracies, highlighting the essential role of unions in sustaining support for social-democratic parties.

2 Background

Economics, political science, and sociology have started investigating the consequences of the latest spurts of technological change, in particular the ICT revolution and the robotization of manufacturing and services.

The starting point is that shifts in technology have distributional consequences: technological innovation produces economic winners and losers, at least in relative terms. New opportunities open for workers endowed with skills that are complementary to new technologies, while workers more substitutable by machines lose out. The identity of winners and losers varies depending on the nature of technological changes. On top of direct economic winners and losers, technological change also affects the distribution of power within society, and therefore political competition in democracy.

2.1 Technological change and trade unions

The wave of automation in recent decades has led to productivity and welfare gains, but it has substantial distributional effects. Economic research has shown how automation penalizes workers in regions that were historically specialized in industries adopting more robots, and individuals whose skills were more substituted than complemented by the new technologies (e.g. Acemoglu and Restrepo, 2020). Automation has also shaped the allocation of labor market opportunities. Traditional, relatively well-paid and stable jobs, especially in the manufacturing sector, have disappeared, while new jobs have been created in service sectors or in new sectors (i.e. the Gig economy sector). However, these newly created jobs are characterized by lower earning trajectories, temporary contracts, discontinuous careers and little or no employment protection (Kaine and Josserand, 2019). As argued in Anelli *et al.* (2021), the structural process by which automation leads to increase in earning inequality and to economic discontent of "losers" of this process, is politically consequential. In fact, higher exposure to robot adoption increases support for nationalist and radical-right parties, both at the regional and at the individual level.

Throughout the 20th century, labor unions have played a crucial role in liberal democracies by hindering the increasing wage inequality and by channeling political demands and discontent into an organized voice (e.g. Ahlquist, 2017). With technological change and automation polarizing wages, unions could have played a crucial role in cushioning the unequal distribution of welfare gains. However, the advent of the new century has seen a diminished importance and effectiveness of unions in the democratic process, combined with an atomization of political demands.

Importantly, the adoption of robots and ICT has consequences for labor organizations, hence it is possible for technological change to affect the distribution of income in two ways: a direct one through labor markets, and an indirect one which is mediated by labor organizations. In fact, over the past decades, the role of labor unions has been weakened by processes related to technological change (as well as by other processes like global trade). In a study of wage inequality in the United States, Kristal and Cohen (2015) (see also Kristal, 2013) find that the indirect effects of computerization on inequality–channelled mainly through weakening unions and to a lesser extent by enhancing the rise of non-standard employment relations– were even greater than the direct effects.

Meyer (2019) suggests that routine task-biased technological change might cause union decline through three mechanisms. First of all, routine task occupations were among the most unionized. Second, routine- biased technological change increases between-worker skill heterogeneity, which reduces workers incentives for collective action. Finally, formerly routine task workers compete for lower-skill jobs, and the increased competition for these low-skill jobs gives employers greater leverage to resist attempts at unionization. Labor unions played an important role in terms of wage setting. Extant evidence (e.g., Western and Rosenfeld, 2011 for the US) shows that the decline of unionization is key to explain increases in inequality. It is therefore important to understand to which extent the now well-documented effect of automation on labor market polarization, and on increased inequality, is mediated by weakened labor organization and wage bargaining position rather than straightforward labor market effects. On the economics side, we theorize that the technological process has directly contributed to weakening the role of unions, by drastically reducing employment in the large highly-unionized sectors (i.e. manufacturing) while boosting the creation of new jobs and new sectors which are outside the traditional scope of unions. We propose three hypotheses that we plan to test empirically:

1. Technological change, in the form of adoption of robots, decreases the proportion of workforce that is unionized at the regional level. We call this the extensive margin of technology on unions.

2. Technological change, in the form of adoption of robots, decreases the proportion of workforce that is unionized within industrial sectors. We call this the intensive margin of technology on unions.

3. Higher exposure to robot adoption at the individual level decreases the probability to be union member.

While studying the direct impact of technological change on worker representative bodies appears crucial to understand the structural societal changes taking place in western democracies and societies, the academic literature has so far produced limited empirical analysis to study this phenomenon. Moreover, there is slender empirical evidence about the evolution of labor unions in liberal democracies, especially in western Europe. This is most likely due to the absence of adequate data. We have collected region-by-sector data on union membership in 15 European countries to test our theory empirically.

In labor economics research, technological change has been isolated as a main driver of the increase in wage inequality and educational premia, which have fostered social cleavages in Western democracies (Acemoglu and Autor, 2011). Since routine jobs –both manual and cognitive– were mostly middle-income and middle- skill jobs, a polarization of the labor market has been documented both in the US and in Europe (Autor and Dorn, 2013; Goos *et al.*, 2014). Polarization involves an increase in employment at the two tails of the wage and skill distribution, along with a shrinkage of the traditional middle class. For instance, computers have destroyed many decently paid clerical jobs, while the computer-based automation of production processes has reduced job opportunities for relatively skilled blue-collar workers. Workers (both actual and prospective) substituted by computer-based technology have been largely absorbed by the service sector in non-routine jobs, typically at lower wages and with less favorable contractual conditions (e.g., drivers and fast-food workers). The main computerization winners have been the high-skill (college-educated) workers in cognitive occupations: their incomes have been diverging from those of the impoverished middle class, which has been falling in the group of losers together with lowskill workers. The latter, even if employed in non-routine tasks, have been complemented by the new technologies much less than the high skilled, and their wage dynamics have been compressed by the additional supply of displaced middle-skill workers competing for the same jobs (Autor, 2015).

A growing literature documents the economic effects of the most recent automation wave. In particular, a strand of the literature uses data on the adoption of industrial robots at the industry level made available for many countries by the International Federation of Robotics (IFR). According to these data, the stock of operational robots in advanced economies has increased substantially between 1993 and 2016, a phenomenon commonly referred to as the "robot shock". Focusing on the US, Acemoglu and Restrepo (2020) find that, at the level of commuting zones, a stronger exposure to the robot shock has a negative effect on local employment rates and wages. To illustrate, the adoption of one extra robot in a commuting zone reduces employment by around 6 workers. The negative effect of robots on employment is stronger in the manufacturing sector, and especially in industries that are most exposed to robots. Moreover, it is more pronounced for workers with less than college education, for blue collars employed in routine manual tasks and assembling, for machinists and transport workers, and for men in general. The negative effect of robots on wages is concentrated in the bottom half of the wage distribution, contributing to the increase in wage inequality. Graetz and Michaels (2018), using industry-level data on a larger sample of countries, find that robot adoption has a positive effect on productivity, but a negative impact on the share of hours worked by low- skill workers. Chiacchio *et al.* (2018) focus on six European countries and find a negative effect of robot adoption on employment at the level of local labor markets. Dauth *et al.* (2018) investigate the impact of industrial robots using matched employer-employee data for Germany and find that the adoption of robots leads to job losses in manufacturing, which are compensated by employment gains elsewhere, mostly in the business service sector. Importantly, fewer manufacturing jobs become available for new entrants in the labor market. Using individual data, they find that affected workers mostly stay with the same employer, but change their occupation and incur wage losses. Overall, automation increases wage inequality: it benefits managers and high-skill workers performing abstract tasks, while low- and medium-skill workers see their earnings decrease, leading to a general decline in the labor share of income.

2.2 Technological change and political behavior: the mediating role of trade unions

There is limited evidence, thus far, on the consequences of the most recent spurts of technological change on political preferences and behavior. Gallego *et al.* (2018), using data from the UK, show that the winners of computerization - educated workers in IT-heavy sectors - become more likely to vote Conservative and less likely to vote Labour, while losers are more likely to support the radical-right option, namely the UKIP. Yet, due to limited data to answer this type of question, they refrain from making more general claims about the radical-right turn of the losers in the British setting. Studying the 2016 US presidential election, Frey *et al.* (2018) show how voters in regions more affected by robotization in manufacturing were more supportive of the Republican candidate, Donald Trump, who was running on a nationalist platform, both in economic and in identitarian terms. Im *et al.* (2019), using data on eleven countries from the European Social Survey, show that workers in occupations at higher risk of automation are more prone to vote for radical right parties. Finally, Dal Bó *et al.* (2019) study patterns of support for the Sweden Democrats in local elections. They show that the share of automation-vulnerable workers in a municipality is robustly correlated with support for the radical-right option. Anelli *et al.* (2021) provide evidence on the effect of automation vulnerability on voting behavior using sub-national and individual-level data on automation exposure for fourteen Western European countries. Regions and individuals that are more exposed to automation tilt towards nationalist, isolationist, and radical right parties.

When it comes to politics, labor unions played an important role in sustaining the embedded liberalism model that obtained in advanced democracies after WWII (Ruggie, 1982). In a nutshell, such a model entailed trade openness – and, more in general, an approximation to economic efficiency– accompanied by policies that compensated the possible losers from structural changes and more in general distributed the gains from economic growth according to criteria of equality or inclusiveness.

Embedded liberalism started entering a crisis towards the end of the XX century, in part due to increased capital mobility –that reduced revenues to finance the welfare state– and increased import competition with emerging economies –that increased the demand for compensation itself. (Rodrik, 1998; Colantone and Stanig, 2018). Plausibly, automation, by tilting the bargaining power in favor of the owners of capital (Kristal, 2013), contributed to weakening the more inclusive "social market economy" model adopted especially by Western European countries. The crisis of embedded liberalism underpins the realignment currently taking place in many Western advanced democracies, and the success of (so-called "populist") radical right parties proposing platforms of economic nationalism, that are isolationist and protectionist in matters of international relations, but unsupportive or skeptical of redistribution and the welfare state (Colantone and Stanig, 2018).

From the perspective of voting behavior, labor unions have historically provided an important link between left parties and blue-collar constituencies (e.g. Przeworski and Sprague, 1986). This link operated in two directions: on the one hand by inducing mainstream left parties to focus on the economic distributional issues relevant to workers (Pontusson and Rueda, 2010), and on the other mobilizing union members in support of mainstream left parties. Leighley and Nagler (2007) demonstrated that unions positively influence voter turnout. Also, trade unions contributed to frame the political discourse and to make workers think in terms of class conflict rather than cultural and ethnic conflict. The reduced importance of unions in the workplace led to a decreased importance of their role as brokers for mainstream social-democratic and labor parties (Piazza, 2001). The weakening of unions also made it more appealing for social democratic parties to move to the center on redistribution issues, in order to capture middle-class constituencies attracted to their cosmopolitan stances on non-economic issues (Häusermann and Kitschelt, 2023; Kitschelt, 2012; Kriesi, 1998). This further weakened the connection between blue-collar constituencies and mainstream left parties.

As suggested by Kitschelt (2012), this decoupling of working-class constituencies from social-democratic parties may partially explain why individuals adversely affected by structural changes such as globalization and automation have increasingly gravitated towards nationalist and radical-right parties (Betz, 1993, 1994; Betz and Meret, 2012). Rennwald and Mosimann (2023) have argued that union members are less likely to realign their vote based on cultural preferences and more inclined to support parties that align with their pro-redistribution stance, compared to non-members. In the context of the United States, Balcazar (2022) demonstrates that automation has political implications by reducing public policy to unions interests. Consequently, we anticipate that radical-right parties, along with nationalist political platforms, will garner lower support in regions with strong unions, while support for the economic left's redistribution platforms will be higher in such areas. Levi (2017) states that "the decline of labor unions has also facilitated the rise of populism by eliminating a source for a framework for understanding the situation of workers."

2.3 The absence of fine-grained data for Europe

Research on the role of unions in recent developments within advanced democracies has been hindered by data limitations. As Ahlquist (2017) aptly notes, "we have too many explanations chasing too few data points that are themselves interdependent in both time and space," and therefore recommends "research designs explicitly taking advantage of heterogeneity in context and population." To pursue such paths, better disaggregated data is required. Additionally, as Pontusson and Rueda (2010) remark, the distribution of union members across income categories varies considerably across countries. Thus, to better understand the role unions play in the political economy, further data collection is imperative.

At present, the most important source of macro-level data on unionization in Europe, collected and compiled by third parties, is the Visser (2019) dataset. This dataset incorporates and updates information about union membership, density, and concentration from the Golden et al. (2009) and Ebbinghaus and Visser (2000) datasets, which have underpinned seminal political science research on the role of unions. However, these cross-national data sources have one main limitation: a level of aggregation that is too coarse for the purposes we aim to address. As detailed in the previous section, country-level data do not significantly advance our knowledge due to the limited number of data points. A sub-national and sectoral analysis is required to achieve clean identification of causal effects.

In the US, researchers have leveraged more fine-grained data available via the Current Population Survey, which includes an item on union membership and large samples that allow for disaggregation (see, for instance, Ahlquist and Downey, 2023). Recently, a burgeoning stream of research on trade unions in the United States has utilized the fine-grained data from Becher *et al.* (2018), with Balcazar (2022) and Becher and Stegmueller (2020b,a, 2024) being notable examples.

3 New data on union density

Sufficiently fine-grained data on union membership is crucial for studying how structural changes, and in particularly automation, have influenced trade unions and political orientations. Unfortunately, as previously noted, such data for Western Europe is challenging to obtain. Official national statistics only rarely collect or provide data on union density at any level, and administrative sources allow to retrieve such information only in a few contexts. While trade unions generally maintain records of their membership numbers, these records are often based on varying collection procedures, definitions of union membership, criteria for including students and retirees, and differing update frequencies. Moreover, statistics derived from the main trade unions' records must also account for the presence of union members in smaller trade unions.

Given these limitations, it is crucial to produce estimates of union density using harmonized methods across different countries and regions. Surveys provide a valuable means to achieve this goal.[1] By adopting a survey-based approach, we develop a new dataset that offers a more comprehensive and harmonized view of union membership across Western Europe.

The strategy that we adopt to derive usable measures of unionization at the necessary level of granularity is multilevel regression with post-stratification (MrP). This approach was originally proposed by Gelman and Little (1997) and Park *et al.* (2004), and it was initially popularized in applied political science research by Lax and Phillips (2009) (see also Gelman and Hill (2007) and Leemann and Wasserfallen (2020) for textbook treatments).

MrP allows researchers to exploit information from nationally representative samples to estimate more disaggregated summaries at the sub-national level. In our study, we use this approach to combine data from the European Social Survey (ESS) and census information from 15 Western European countries. The ESS includes individual-level data on union membership. Importantly, along with basic demographic information, the ESS also reports the occupation (ISCO), industry (NACE) and region of residence (NUTS) of each individual. We leverage census data on the distribution of these variables in the population to post-stratify multi-level probit models for unionization. Complete information on the data collection and on the harmonization of the variables in different sources can be found in Sections A and B of the Online Appendix, respectively.

MrP involves two main steps: prediction by type and aggregation, each critical for deriving accurate sub-national estimates of union membership.

The initial step involves predicting union membership at the individual level using a probit model based on socio-demographic types, including effects for sub-national areas. This model leverages data from the European Social Survey (ESS) to estimate the probability of union membership. We can then weight these predicted probabilities by the prevalence of a cell ("type") in the population, based on census data, and aggregate them at the

[1] The use of surveys to estimate key metrics has proven effective, particularly in the absence of comprehensive administrative records. For example, Thomas Piketty's influential work on inequality extensively utilizes survey data to derive comprehensive estimates of income and wealth distribution.

desired level, for instance region or industry. Finally, we assess the performance of the approach we propose. This refers both to the performance of the multi-level regression models themselves, and the performance of the post-stratification exercise relative to the true target values that we are trying to compute.

3.1 First step: models

For each country, we estimate models of the form

$$\begin{aligned} Pr(Union_i = 1) = \quad & Probit^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} \\ & + \beta \cdot round + \boldsymbol{\beta} \cdot \boldsymbol{X_c}) \end{aligned}$$

where $Union_i$ is an indicator variable for whether the respondent is member of a union, and the α terms are random effects for gender, age category, education level, NUTS2 region of residence, NACE (rev 1.1) code of the industry in which the respondent is employed, and 2-digit ISCO for the occupation of the respondent. F() is the probit link. $\boldsymbol{X}$ is a vector of regional, context-level, pre-sample variables that includes the share of low and middle skilled workers, the share of low and medium tech, the share of finance and business, the share of primary and service sector, and the share of foreign-born workers. Throughout, we estimate the probit models separately for each country.

From the estimates, we can compute predicted probabilities for cells ("types") defined by gender, age group, education level, region, industry and occupation. For instance, the model estimate the probability of union membership for a woman in her 30s, with tertiary education, working as a doctor in the healthcare industry in the Paris region. The model predicts probabilities for each possible socio-demographic type (λ). Let θ_λ represent the predicted probability of unionization for type λ.

In principle, different model specification can be employed to predict union membership. In total, we estimate two dozen different models. Our specifications mostly follow a dynamic MrP approach (Gelman *et al.*, 2019) to model time variation. Our preferred baseline specification allows for different time trends for industrial sector and is specified as follows:

$$
\begin{aligned}
Pr(Union_i = 1) = \quad & Probit^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} \\
& + \beta \cdot round + \beta^{ind} \cdot round + \boldsymbol{\beta} \cdot \boldsymbol{X_c})
\end{aligned}
$$

To enhance the accuracy of our estimates, we assess the relative performance of the different multi-level mixed effects models. The full list of models can be found in Sections C of the Online Appendix. We implement a cross-validation procedure and compute the root mean square error (RMSE) for each model specification to identify the most effective models. While cross-validation is a standard technique in statistical learning and machine learning, it is less commonly used in economics and political science. We thus provide an explanation of its methodology and relevance.

3.2 Cross-validation

An initial evaluation of the procedure focuses on the predictive performance of the probit models for unionization. Given that our primary objective is to predict unionization rates out of sample, we are not primarily concerned with the in-sample performance of the probit regressions themselves. A potential drawback of models exhibiting very high in-sample predictive power is the risk of overfitting: the model might be very accurate in picking up what are ultimately idiosyncratic features of the specific sample one is dealing with. When these idiosyncratic features are then reproduced in out-of-sample prediction, they lead to lower predictive performance. Conversely, more parsimonious models might omit some idiosyncratic characteristics of the current sample but tend to perform better out of sample.

For this reason, we use cross-validated predictions to assess model performance. In practice, we randomly split the sample into K = 10 folds. We then iteratively pick one fold k (which becomes the “test set”) and use the sample without fold k as the training set. We estimate the model in 1 on the training set, and form predictions for fold k based on the estimates of the parameters from the training set and the values of the observable features in fold k. We repeat the operation for all folds, to arrive at a vector of predictions for all observations in the sample. We can then run diagnostics on these cross-validation predictions. As a performance metric, we can calculate the root mean square error as the square root of the average squared difference between the midpoint of the bin, and the sample proportion within that bin. Notice that this is expressed in the same units as the variable of interest. In this case, the variable is a probability, so the rmse can be interpreted (loosely) as a measure of how far "on average" the predicted probability and the actual proportion are from each other.

We develop a specific RMSE measure, inspired by the fact that we aim at group-wise unionization rates. For group g defined by observable characteristic (e.g., NUTS region) define Ω_g set of observations in group g, with cardinality N_g. We thus compute the average of the predicted probabilities for group g:

$$\hat{P}_g = \frac{\sum_{i \in \Omega_g} \hat{P}_i}{N_g}$$

Analogously, let us define the empirical frequency:

$$F_g = \frac{\sum_{i \in \Omega_g} (Union_i = 1)}{N_g}$$

Finally, we define the group-wise calibration RMSE based on groups G as

$$\mathrm{RMSE_G} = \left(\frac{\sum_g (\hat{P}_g - F_g)^2}{G} \right)^{\frac{1}{2}}$$

In a nutshell, this measure compares the (cross-validated) predicted probabilities for a given region or industry with the empirical frequency (i.e. the naive estimate of the unionization rate) observed among survey respondents from that region or industry. In loose terms, it captures the average divergence between the predicted probabilities and the actual unionization rates in the survey. Although this group-based approach may appear more demanding, it aligns with the principles of the standard calibration method, which involves binning predicted probabilities.

This metric enables us to rank estimation models according to their relative performance. To predict union density, we thus select the model that ranks on average best across the countries we study, according to the RMSE. The model has the following form:

$$\begin{aligned} Pr(Union_i = 1) = \quad & Probit^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} \\ & + \beta \cdot round + \alpha^{occ,edu} + \boldsymbol{\beta} \cdot \boldsymbol{X_c}) \end{aligned}$$

Note that this specification, which overall performs better based on our RMSE metric, includes a random effect for the combination of occupation and educational attainment. It goes without saying that the best model according to the RMSE is not necessarily the same in all countries considered. Figure 5 in Section C of the Online Appendix displays the RMSE of the different models in the 15 countries in our sample.

While the cross-validation allows to assess the predictive performance of different models in estimating the predicted probabilities of unionization of every type, to retrieve aggregate measures of union density the MrP approach requires that such predicted probabilities are post-stratified using weights from census data.

3.3 Post-stratification

The second step fo MrP involves using census data to weight the predicted probabilities of union membership for each socio-demographic type θ_λ at the chosen level of aggregation. This allows for the estimation of union density at various sub-national levels.

To illustrate, we can aggregate the unionization probabilities at the regional level using the following formula:

$$U_r = \sum_{\lambda \in \Lambda_r} \left(\frac{N_\lambda}{N_r} \right) \theta_\lambda$$

Here:

- Λ_r denotes the set of types in region r.
- N_λ is the number of λ-type individuals in region r.
- N_r is the total population in the region.

This approach is flexible and can be adjusted to aggregate data at different levels depending on the research objectives. For example, if the goal is to provide unionization data at the region r / industry i level, the aggregation formula becomes:

$$U_{ri} = \sum_{\lambda \in \Lambda_{ri}} \left(\frac{N_\lambda}{N_{ri}} \right) \theta_\lambda$$

In this case:

- Λ_{ri} represents the set of types for which the region is r and the industry is i.
- N_λ is the number of λ-type individuals in region r and industry i.
- N_{ri} is the total population in the region-industry.

For instance, we weight the probability of unionization for our female ideal type by the proportion of women with the same socio-demographic characteristics among the total workforce in the healthcare industry in Paris. This process is repeated for each ideal type within the region-industry and summed over all types.

By utilizing the Multilevel Regression with Poststratification (MrP) approach, we can generate more precise and regionally detailed estimates of union density. This methodology also provides the flexibility to produce union density estimates at varying levels of aggregation, such as by region and industrial sector.

3.4 External validation

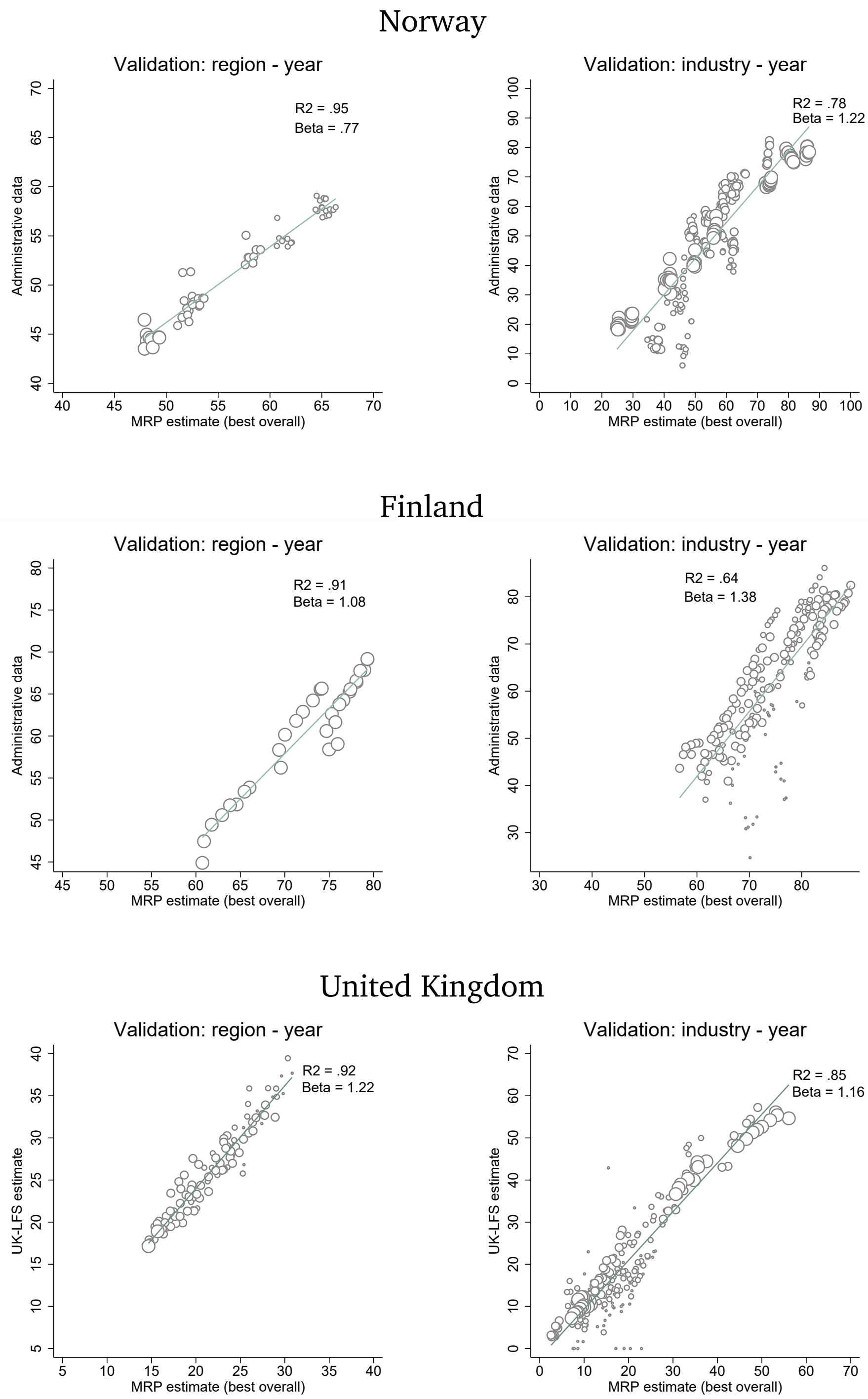
The primary advantage of employing a MrP approach to measure union density at the subnational level lies in its ability to generate estimates by combining a nationally representative survey with highly informative census data on subnational population composition. Nevertheless, it is crucial to evaluate the reliability of the estimates produced by this method. Where available, administrative data and surveys representative at the subnational level enable us to assess these estimates through a process that we define as "external validation." However, such robust sources are available only in a limited number of European countries. In this study, we obtained administrative data from Norway and Finland to produce alternative estimates of unionization at the regional and industry levels. These administrative data provide a direct comparison with our MrP estimates.

In Figure 1, we compare our estimates of union density, aggregated at the regional (NUTS-2) and industry (NACE subsections) levels, with the corresponding values from administrative data in Finland and Norway, as shown in the first two rows. The last row presents a different context for external validation: in the United Kingdom, the Labour Force Survey includes a measure of union membership, allowing for the estimation of subnational union density that is representative at the regional and industry levels, particularly for larger units.

Each data point in the charts represents a region (left panels) or industry (right panels) for a given year. The results indicate a high correlation between unionization estimates derived from administrative or survey data and those from the MrP method. The R-squared values from an OLS regression, where data from external sources are regressed on MrP estimates, are 95% in Norway, 91% in Finland, and 92% in the UK for regional variation. On average, in countries with available administrative data (Norway and Finland), the MrP approach predicts higher union density, but the variation across regions and over time is consistent across the two sources. For industry-level variation, the R-squared values tend to decrease, yet the overall convergence between the methods is still evident.

The strong convergence between the predictions of the two approaches reinforces the credibility of the estimated values. In conclusion, the results of our external validation provide robust evidence supporting the reliability of our approach.

Figure 1: External validation with administrative data or labour force survey



Note: Each observation in the chart represents a region in a given year in the first row or an industry (NACE subsection) in a given year in the second raw. The size of circles is proportional to population. External validation data are sourced from administrative data provided by Statistics Norway and Statistics Finland and from the United Kingdom Labour Force Survey.

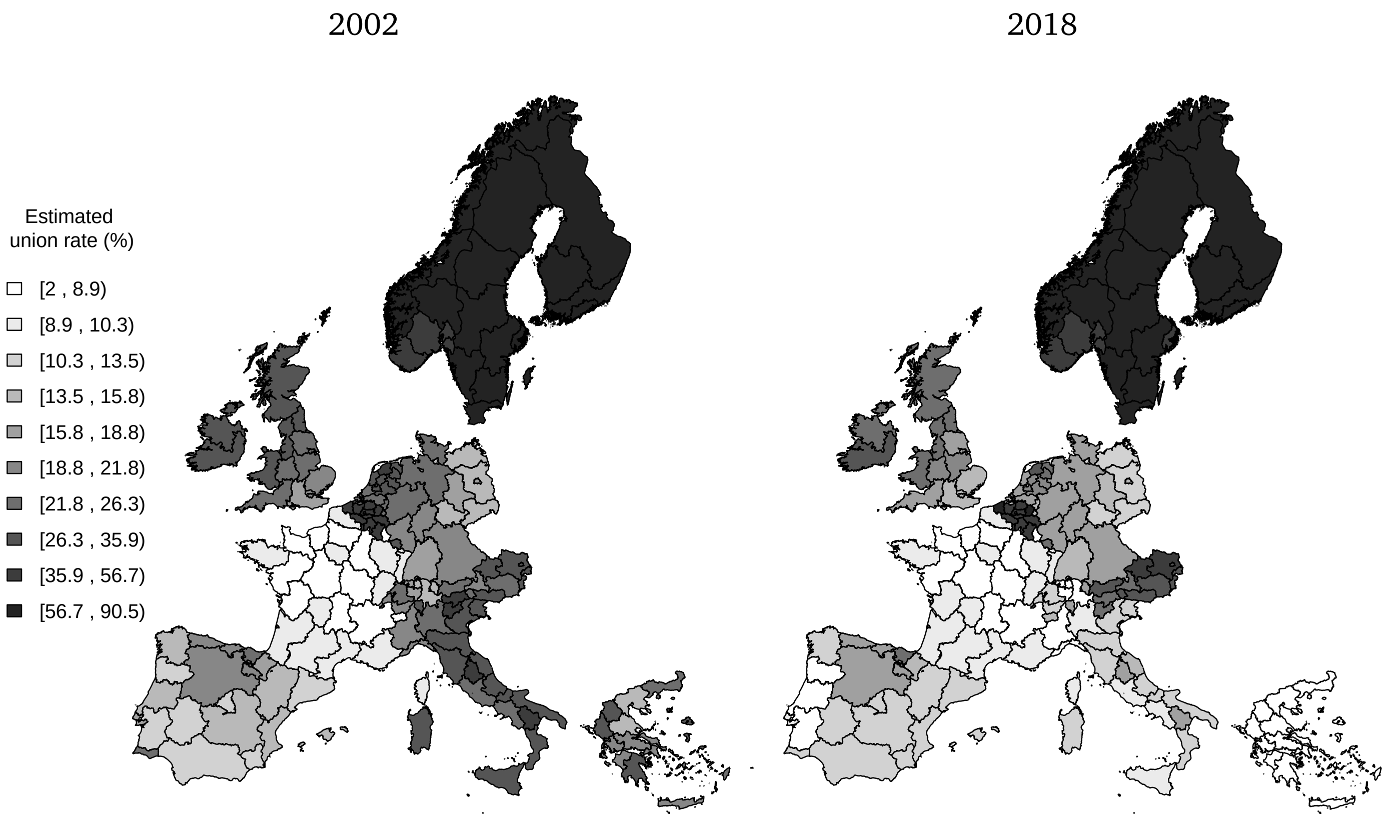
3.5 Union density in Western Europe

Our novel dataset on union density spans over around two decades (2002-2018) and includes 15 western European countries. These are: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the UK. Union density is measured at the region-by-industry level. Regions are defined at the NUTS-2 level for all countries except Germany and the UK, where the NUTS-1 level is used. Industrial sectors are classified according to the NACE revision 1.1 at the 2-digit level. Our measure of union density the share of individuals in the labor force that are union members, excluding self-employed workers.

Figure 2 illustrates union density across Western European regions in 2002 and 2018, representing the first and last years currently covered by our data. The results reveal a rich heterogeneity in unionization levels, with countries such as Finland, Sweden, and Norway exhibiting union densities exceeding 50%, while countries like France, Portugal, and Greece display significantly lower levels of unionization. Beyond the well-documented variation between countries—rooted in deep-seated historical and institutional factors—our data also uncovers substantial intra-country variation. For example, union density is notably higher in northern Spain compared to the south, and in the industrial regions of West Germany compared to other parts of the country. In France, union density remains uniformly low, ranging from 8% to 10% in most regions.

Figure 3 unveils a pronounced trend of deunionization in several European regions, especially in Greece, Ireland and Italy. Importantly, the implications of a higher or lower union density levels are shaped by complex, context-specific factors that can vary significantly across nations. As Ahlquist (2017) insightfully notes, trade unions are inherently diverse organizations, differing across multiple dimensions. Therefore, while cross-national comparisons of unionization trends should be approached with nuanced understanding, subnational variation provides a valuable analytical lens. To rigorously investigate trade

Figure 2: Estimated union density by region



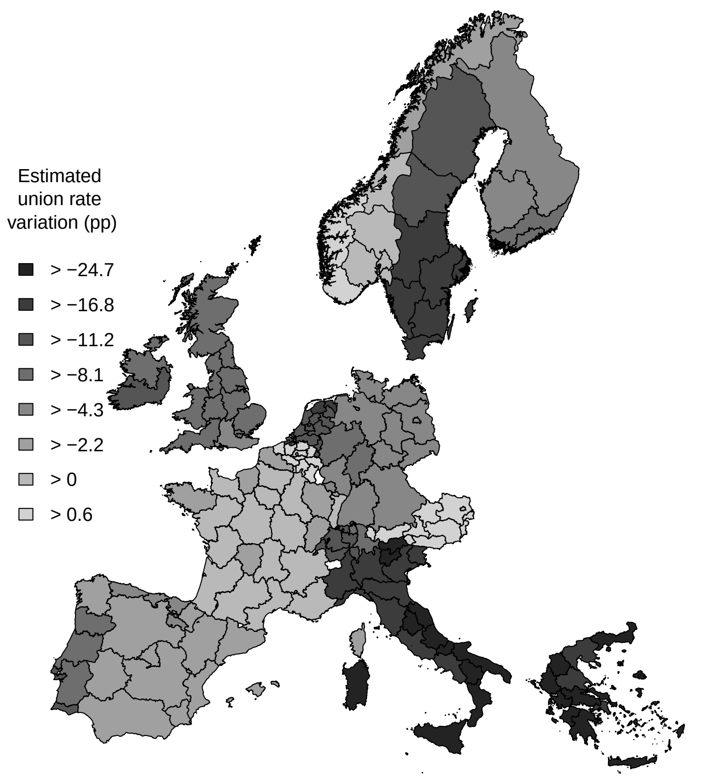
unions, it is crucial to employ research designs that capitalize on this subnational diversity.

Our novel granular data on on subnational and sector level union density are crucial to allow researchers to shed new light on a broad set of question about the role of trade unions in economic, social and political processes. In the next section, we present the results of our analysis aimed at investigating the impact of automation on unionization in Western Europe.

4 Empirical analysis

In this study, we leverage our novel dataset on union density to estimate the impact of industrial robot adoption on unionization rates. To comprehensively analyze the effect of robotization on unions, we employ a multi-faceted empirical approach. First, we examine the impact of regional exposure to robotization on union density at the regional level, which

Figure 3: Estimated union density variation (2002-2018)



allows us to capture the ’extensive margin’ of robotization’s effects. In fact, we expect industrial robot adoption to decrease the proportion of the labor force that is unionized. Next, we assess the impact of industrial robot adoption within specific sectors on union density within those sectors. This analysis provides insights into whether trade unions are weaker in industries more exposed to automation, thus less effective at bargaining wages (“intensive margin”). Finally, we estimate the effect of individual exposure to automation on the likelihood of union membership, as well as on various attitudes toward trade unions and individual working conditions.

4.1 Cross-region analysis

We measure regional exposure to robotization based on the ex-ante industry specialization of each region. Following Acemoglu and Restrepo (2020), we measure the time-varying exposure to automation at the regional level as:

$$\text{Regional Exposure}_{crt} = \sum_{j} \frac{L_{crj}^{\text{pre-sample}}}{L_{cr}^{\text{pre-sample}}} * \frac{R_{cj}^{t-1} - R_{cj}^{t-n}}{L_{cj}^{\text{pre-sample}}}, \tag{1}$$

where: c indexes countries, r NUTS-2 regions, j manufacturing industries, and t years, $R_{cj}^{t-1} - R_{cj}^{t-n}$ is the change in the operational stock of industrial robots over the past n years, in country c and industry j. This change is normalized by the pre-sample number of workers employed in the same country and industry, $L_{cj}^{presample}$. This ratio provides a measure of the intensity of robot adoption at the industry level. To retrieve the regional-level exposure, we take a weighted summation of the industry-level changes, where the weights capture the relative importance of each industry in each region. Specifically, each weight is the ratio between the number of workers employed in a given region and industry ($L_{crj}^{presample}$), and the total number of workers employed in the same region ($L_{cr}^{presample}$). Importantly, weights are based on pre-sample figures, dating before the surge in the adoption of industrial robots observed from the mid-1990s onwards. Intuitively, regions that were initially specialized in industries in which the adoption of robots has later been faster are assigned stronger exposure to automation. Data on robot adoption by industry is sorced from the International Federation of Robotics. Detailed information on the data sources used to construct our measure of robot adoption can be found in Section D of the Online Appendix.

In order to address endogeneity concerns, we instrument robot adoption with a set of measures of technological progress. Specifically, we utilize the following measures: the producer price index of computers sourced from FRED, single-thread performance, and the number of transistors per microprocessor, derived from Rupp's 50 Years of Microprocessor Trend Data. These instruments are designed to capture global technological shifts in robotics and computing, which are assumed to be exogenous to local economic dynamics.

The instrumental variables for regional exposure to robotization are computed as follows:

$$\text{IV Regional Exposure}_{crt} = \sum_{j} \frac{L_{crj}^{\text{pre-sample}}}{L_{cr}^{\text{pre-sample}}} * \text{Rep}_{\text{j}} * \Delta\text{Index}_{\text{t-1,t-n}} \tag{2}$$

Here, Rep_j, representing labor replaceability for industry j, is derived from the labor replaceability index developed by Graetz and Michaels (2018). The index, based on the 1980 U.S. Census data, reflects the proportion of hours worked within industry j in occupations susceptible to automation by robots. This captures how automatable is labor in sectors in which the region was historically specialized. The term $\Delta\ Index_{t-1,t-n}$ denotes the change in the producer price index of computers, single-thread performance, or the number of transistors per microprocessor over time. Consequently, the instruments are constructed as the interaction between these measures of technological progress (which vary over time but are common across countries) and regional automatability (which is time-invariant).

In our empirical analysis, we thus estimate the impact of robot adoption on union density in western European regions. The baseline specification we estimate has the general form:

$$Union\_rate_{crt} = \alpha + \beta \text{ Regional Exposure}_{crt} + \lambda_{ct} + \lambda_r + \epsilon_{rt}$$

where r indexes NUTS-2 regions, c countries, and t years. $Union\_rate_{crt}$ represents union density in NUTS-2 region r and year t, $Regional\ Exposure_{crt}$ is robot adoption at the regional level over the past three years in the baseline estimates. The terms λ_{ct} and λ_r are country-year and NUTS-2 region fixed effects, respectively. Robust standard errors are utilized in the estimation.

The dependent variable, union density at the regional level for a given year, is derived from our novel dataset. We consider three alternative measures of union density based on distinct specifications used in generating our estimates. The first measure is our baseline MrP specification, the second is the specification that overall performed best according to our RMSE metric, and the third includes the optimal specification for each country.[2]. These measures are highly correlated, with a correlation coefficient close to 1.

We employ two different models. First, we regress union density on automation shock in an OLS regression with country-year and region fixed effects. Next, we conduct a 2SLS regression where robot adoption at the country level is instrumented with our measures of technological progress. The results are presented in Table 1.

Table 1 reveals a negative and statistically significant relationship between robot adoption and unionization rates across all models. The effect of automation on unionization is stronger in the models with instrumental variables. The impact of automation on unionization is more pronounced in the models using instrumental variables, particularly when employing the union density measure based on the best MrP specifications (columns (5)

[2]The specification that overall performed is the specification that, on average, rankEd better (i.e. had the lowest RMSE) across all countries. This means that the estimates of union density refers to the same specification for every country. Instead, the optimal specification by country includes a different specification for every country.

Table 1: Cross-region analysis

VARIABLES	(1) Union (baseline)	(2) Union (best)	(3) Union (rel. best)	(4) Union (baseline)	(5) Union (best)	(6) Union (rel. best)
Robots Regional	-0.176***	-0.177***	-0.184***	-0.347**	-0.430**	-0.466***
Exposure	[0.048]	[0.051]	[0.052]	[0.161]	[0.171]	[0.177]
Estimator	OLS	OLS	OLS	2SLS	2SLS	2SLS
Observations	1,413	1,413	1,413	1,413	1,413	1,413
Country-year FE	X	X	X	X	X	X
Region FE	X	X	X	X	X	X
Std dev. Y	0.521	0.526	0.533	0.521	0.526	0.533
Std dev. X	0.421	0.421	0.421	0.421	0.421	0.421
Magnitude	-0.143	-0.141	-0.145	-0.281	-0.344	-0.368
First stage F-stat				15.45	15.45	15.45

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

and (6)). In both the OLS and IV specifications, the p-value of the robots' regional exposure coefficient is less than 0.05.

In the models shown in columns (1), (2), and (3), a one standard deviation increase in robot density (equivalent to 17 robots per 100,000 workers) leads to a 14% reduction in a standard deviation of unionization in European regions, after accounting for fixed effects. This finding provides evidence for the 'extensive margin' of technology on unions.. The coefficient is statistically significant (p-value < 0.01). In the IV models presented in columns (4) through (6), the coefficients remain negative, are larger in absolute magnitude, and are statistically significant at the 95% confidence level. Specifically, a one standard deviation increase in exposure to robotization results in a decrease in unionization ranging from 28% to 37% of a standard deviation, depending on the specification.

In summary, regions with higher levels of industrial robot adoption tend to experience a decline in the proportion of the workforce that is unionized. These findings are robust across various specifications and estimation strategies. Table 2 presents a series of robustness checks. The first row reports the estimate from the IV specification using union density from the best overall model (i.e., column 5 of Table 1). In rows 2 through 6, we enhance the specification by controlling for interactions between year dummies and a set of pre-sample, context-level variables. These variables include the shares of low-, middle-, and high-skilled workers, as well as the share and stock of foreign-born workers. The purpose of these robustness checks is to ensure that the observed decrease in unionization in regions more exposed to robotization is not driven by other regional trends related to the characteristics of the labor force.

To further verify that the effect is not driven by larger regions, we conduct additional robustness checks in rows 7 through 10 by excluding the top 10% and 25% of regions by size within each country, as well as the largest regions overall. Importantly, the coefficient on individual exposure to automation remains consistently negative and statistically significant. Another potential concern is that the observed effect might be influenced by regions with a higher concentration of certain sectors that are particularly susceptible to robotization. To test this, we conduct regressions where robot adoption is instrumented using instrumental variables with employment weights recalculated after excluding specific sectors. We exclude the automotive, electronics, and metallurgical and mineral sectors[3], with results presented in rows 11 to 13. Once again, the coefficients remain negative and statistically significant, reinforcing the robustness of our findings.

Finally, we examine whether the effect is sensitive to the choice of different lags for changes in the stock of robots. Our baseline specification considers robot adoption over the past three years. We re-estimate the model using alternative time periods, ranging from two to six years. Results, displayed in Figure 6 in Section E of the Online Appendix, consistently show that the coefficient on regional exposure to robotization remains negative and statistically significant, with the magnitude of the effect increasing as the lag period

[3]In particular, we exclude the NACE sectors DM (manufacture of transport equipment), DL (manufacture of electrical and optical equipment) and DI-DJ(manufacture of other non-metallic mineral products and manufacture of basic metals and fabricated metal)

Table 2: Robustness checks

VARIABLE	Union density (best)
1) Baseline	-0.430**
	[0.171]
2) Year dummy * Initial share low-skill workers	-0.427**
	[0.176]
3) Year dummy * Initial share med-skill workers	-0.406**
	[0.174]
4) Year dummy * Initial share high-skill workers	-0.401**
	[0.174]
5) Year dummy * Initial share foreign workers	-0.431**
	[0.178]
6) Year dummy * Initial stock foreign workers	-0.417**
	[0.178]
7) Excluding top 10% regions by area	-0.424**
	[0.193]
8) Excluding top 25% regions by area	-0.414*
	[0.232]
9) Excluding top 10% overall regions by area	-0.521***
	[0.178]
10) Excluding top 25% overall regions by area	-0.708***
	[0.210]
11) Excluding automotive sector	-0.419*
	[0.228]
12) Excluding electronics sector	-0.397**
	[0.177]
13) Excluding metals and minerals sectors	-0.434***
	[0.163]

Robust standard errors in brackets
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

lengthens.

We provide robust evidence that technology impacts the 'extensive margin' of unionization, showing a decline in unionization rates in regions more exposed to robotization. This finding aligns with our hypothesis that, in regions where robot adoption accelerates, there is a shift in employment from historically highly unionized sectors to less unionized ones. An alternative explanation could be that robotization not only facilitates the reallocation of employment across sectors but also reduces the share of unionized workers within automated sectors. To test this hypothesis, we examine the variation in union density across different sectors in the next section.

4.2 Cross-industry analysis

We measure robot adoption at the industry level in a given country based on change in the stock of Industrial robots according to the IFR. In particular, we measure the time-varying automation at the industry level as:

$$\text{Industry Exposure}_{cjt} = \frac{R_{cj}^{t-1} - R_{cj}^{t-n}}{L_{cj}^{\text{pre-sample}}} \tag{3}$$

where: c indexes countries, j manufacturing industries, and t years, $R_{cj}^{t-1} - R_{cj}^{t-n}$ is the change in the operational stock of industrial robots over the past n years, in country c and industry j. This change is normalized by the pre-sample number of workers employed in the same country and industry, $L_{cj}^{presample}$.

Similarly to the cross-region analysis, we instrument robot adoption with our set of measures of technological progress. The instrumental variables for regional exposure to robotization are computed as follows:

$$\text{IV Industry Exposure}_{cjt} = \mathrm{Rep_j} * \Delta\mathrm{Index_{t-1,t-n}} \tag{4}$$

In our empirical analysis, we proceed to estimate how robot adoption affects union density across industries in Western Europe. The baseline specification is represented by the following equation:

$$Union\_rate_{cjt} = \alpha + \beta \text{ Industry Exposure}_{cjt} + \lambda_{ct} + \lambda_j + \epsilon_{cjt}$$

where j indexes NACE industries[4], c denotes countries, and t represents years. Here, $Union\_rate_{cjt}$ refers to union density within NACE industries j in country c and year t, while $Industry\ Exposure_{cjt}$ measures robot adoption at the industry level over the past three years in the baseline estimates. Industries without robot adoption, as recorded by the IFR, are assigned an exposure value of 0. The terms λ_{ct} and λ_j are country-year and NACE industry fixed effects, respectively. Robust standard errors are applied in the estimation.

Consistent with our regional analysis, we incorporate three measures of union density: the baseline MrP specification, the specification that performed overall best based on the RMSE, and the country-specific optimal specification. We then perform a 2SLS regression, instrumenting country-level robot adoption with our technological progress measures. The results are reported in Table 3.

Table 3 indicates that the relationship between robot adoption and unionization rates is negative across all models, but not statistically different from 0. After accounting for fixed effects, the estimated impact is minimal, ranging from -0.7% to -1.7%, depending on the specification, and lacks precision. Consequently, the data do not support the conclusion that automation significantly alters unionization rates within industrial sectors.

When these findings are considered alongside the results from the regional analysis, the evidence suggests that industrial automation may have induced a compositional shift in the labor force. This shift appears to move employment from highly unionized sectors to those with lower unionization, thereby reducing overall union presence without substantially

[4]Industries are distinguished according to the NACE 2-digit classification, revision 1.1, with manufacturing sectors in section D grouped into the following subsections: DA, DB-DC, DD, DE, DF-DG-DH, DI-DJ, DK, DL, DM, and DN. A separate category is included for unemployed workers.

Table 3: Cross-industry analysis

VARIABLES	(1) Union (baseline)	(2) Union (best)	(3) Union (rel. best)
Robots Industry Exposure	-0.130	-0.057	-0.103
	[0.654]	[0.642]	[0.648]
Estimator	2SLS	2SLS	2SLS
Observations	6,061	6,061	6,061
Country-year FE	X	X	X
Industry FE	X	X	X
Std dev. Y	6.630	6.630	6.604
Std dev. X	0.865	0.865	0.865
Magnitude	-0.0169	-0.00742	-0.0135
First stage F-stat	5.451	5.451	5.451

Robust standard errors in brackets
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

affecting the proportion of unionized individuals within those sectors.

4.3 Individual level analysis

To further substantiate our findings and provide additional insight into the regional and sectoral analyses, we now conduct an individual-level analysis to examine the effect of personal exposure to automation on the likelihood of being a union member and on a set of associated individual outcomes.

In conducting individual-level analyses, it is crucial to use measures that extend beyond the automation exposure associated with a person's current occupation. This is because an individual's current occupation may already be influenced by prior automation-driven changes, either through direct or indirect displacement. For example, some workers may have transitioned to low-automation service occupations after being displaced by robots from previous manufacturing jobs. Similarly, new labor market entrants may find themselves in less secure and lower-paying occupations due to the shrinking availability of "good jobs" in manufacturing, a consequence of automation. These individuals, based on their current occupations, might appear to have low automation exposure, yet they represent classic cases of automation-induced displacement. Moreover, unemployed individuals, who may have been displaced due to automation, lack current occupational data altogether, further complicating the assessment.

To address these issues, we employ a counterfactual measure of individual vulnerability to automation that accounts for potential displacement. This approach is particularly relevant for new labor market entrants who may face unemployment or find themselves in suboptimal, low-automation jobs. We adopt the individual automatability measure developed by Anelli *et al.* (2021).

This methodology utilizes historical data on labor market conditions from the early 1990s, prior to the most recent wave of automation. Instead of relying on observed current occupations, we estimate predicted probabilities of employment in various occupations based on factors such as education, age, gender, and region of residence. These estimates are derived from the early 1990s European Labor Force Survey (EU-LFS) data, calculated separately for each country. We then apply these estimates to respondents from the European Social Survey (ESS), who were interviewed between 2000 and 2018—a period marked by rapid robotization. By combining these predicted probabilities with the automatability measure for each occupation, we generate a metric of "individual vulnerability to automation." This metric is then multiplied by the pace of robot adoption in each country and year, yielding a measure of "individual exposure to automation."

Our approach effectively assigns higher automation exposure to individuals who, based on pre-automation labor market conditions, would have been more likely to be employed in automatable occupations. This method captures the reality that, due to automation, certain job opportunities have diminished, potentially resulting in unemployment or employment in less desirable occupations.

Following Anelli *et al.* (2021), we define individual exposure to automation as:

$$\text{Individual Exposure}_{icrt} = \Delta R_{ct} * \underbrace{\sum_j \widehat{Pr}(o_i = j|\text{age}, \text{gender}, \text{edu}, r) * \theta_j}_{\text{Individual Vulnerability}}$$

where: $\Delta R_{crt} = \sum_j \frac{L_{crj}^{\text{pre-sample}}}{L_{cr}^{\text{pre-sample}}} * \frac{R_{cj}^{t-1} - R_{cj}^{t-n}}{L_{cj}^{\text{pre-sample}}}$ represents region-level exposure to robot adoption over the past n years. The term $\widehat{Pr}(o_i = j|\text{age}, \text{gender}, \text{edu}, r)$ denotes individual i's predicted probability of working in occupation j, estimated using pre-shock data from the European Labor Force Survey (EU-LFS). The parameter θ_j captures the automation threat associated to occupation j, based on the scores provided by Frey and Osborne (2017). We calculate these individual exposure scores for respondents of the European Social Survey (ESS) interviewed between 2000 and 2018 across the 15 Western European countries included in our study.

We now proceed to estimate the effect of individual exposure to automation on likelihood of being a union member. Specifically, we estimate the following equation:

$$\text{Union}_{it} = \alpha + \beta_1 \text{Individual Exposure}_{icrt} + \beta_2 X_{it} + \lambda_{ct} + \lambda_r + \varepsilon_{icrt}$$

where i indexes individuals, c countries, r regions, and t years. Individual Exposure$_{icrt}$ represents our measure of individual exposure to automation. The percentage change in total operational robots is between year t – 1 and t – 3, in the country c where individual i resides. X_{it} is a vector of individual control variables, including age, gender, and years of education. Standard errors are clustered by region-year. As in our previous regression analysis, the change in the stock of robots is instrumented using our three indexes of technological progress.

Table 4 presents the results of the regression analysis. The coefficient on individual

Table 4: Individual level analysis

VARIABLES	(1) Union (best)	(2) Union (best)
Individual Exposure	-0.011***	-0.074***
	[0.004]	[0.013]
Estimator	OLS	2SLS
Instrument		Technological
Observations	224,922	224,922
Age, gender, edu	X	X
Country-year FE	X	X
Region FE	X	X
Std dev. Y	0.375	0.375
Std dev. X	0.421	0.421
Magnitude	-0.0121	-0.0831
R-squared	0.166	
Kleibergen-Paap F-stat		54.04

Standard errors clustered by region-year
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

exposure to automation is both negative and statistically significant in the OLS and 2SLS estimations. In the IV model, a one standard deviation increase in individual exposure to automation is associated with a 7.4 percentage point reduction in the probability of being a union member. After accounting for fixed effects and control variables, this reduction corresponds to approximately 8 percentage points. These results suggest that robotization significantly contributes to lower unionization rates among individuals more vulnerable to automation.

It is important to emphasize that our measure of exposure is counterfactual, indicating a decreased propensity to unionize among those who would have been more likely to work in automatable occupations. This finding reinforces the expectation that, due to automation and the resulting structural changes in the labor market, many workers have been displaced into employment contexts with lower union density.

To gain a deeper understanding of our findings, we re-estimate the model using a different set of dependent variables designed to capture economic conditions and perceptions. These variables are constructed based on specific items from the European Social Survey (ESS) and include whether a trade union is present at the respondent's workplace, whether the respondent has a permanent contract, their perception of being fairly compensated relative to their efforts and achievements, their sense of job security, their likelihood of becoming unemployed within the next 12 months, and their agreement with the statement that strong trade unions are necessary to protect working conditions and wages.

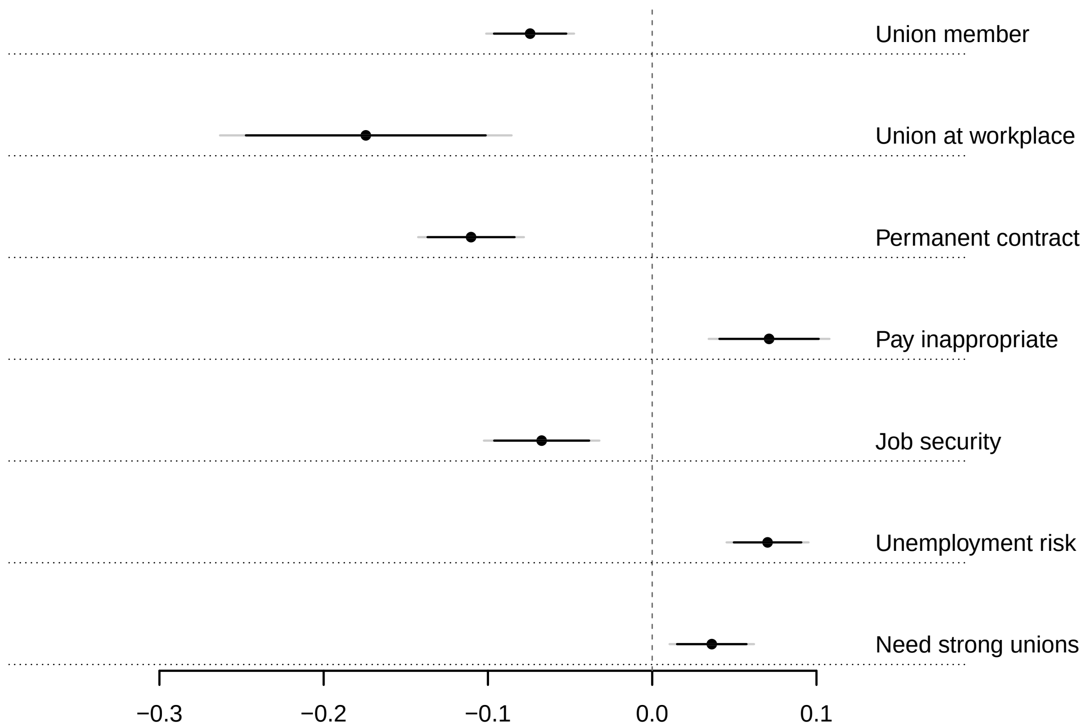
The results, presented in Figure 4, reveal several key findings. In addition to the previously discussed negative effect on individual union membership, the analysis suggests that individuals with higher exposure to automation are less likely to work in a unionized workplace, less likely to hold a permanent contract, more likely to perceive their pay as inadequate, and more likely to be in insecure jobs with a higher risk of unemployment. Additionally, these individuals are more inclined to believe in the importance of strong unions to safeguard working conditions and wages.

In summary, our analysis shows that industrial automation has led to a decline in unionization, largely by shifting employment from highly unionized sectors to those with lower union presence. The individual-level analysis further supports this, indicating that workers more vulnerable to automation are less likely to be union members and experience less secure working conditions. These findings suggest that robotization is reshaping labor markets, contributing to a reduced overall union presence.

4.4 Unions and electoral results

In this study, we have focused on the impact of automation on trade unions and the resulting decline in union membership. This decline carries significant political implications, as we hypothesize that the reduction in unionization rates driven by automation is politically consequential. Specifically, the weakening of a key intermediary institution like trade unions may help explain why many voters and regions have shifted their support away from traditional socio-democratic parties.

Figure 4: Effect of individual exposure to robots on economic conditions and perceptions



Historically, close ties between labor unions and social democratic parties were crucial for mobilizing working-class voters and securing electoral victories (e.g. Przeworski and Sprague, 1986). However, empirical evidence for this relationship has been limited due to data constraints. Our new data can be leveraged to test the hypothesis that support for social democratic parties is indeed linked to higher union density. By connecting unionization rates with district-level voting behavior across fifteen Western European countries, we provide empirical insights into this long-assumed relationship.

The outcome variables are derived from an updated version of the data used in Colantone and Stanig (2018) and Anelli *et al.* (2021). Official election results are sourced from the Constituency-Level Election Archive (CLEA, Kollman *et al.*, 2019), the Global Election Database (GED, Brancati, 2016), and various national sources. For each district in each election, we have information on vote shares at the party level.

In particular, we define p_{ldt} as the vote share for party l, in district d, at time (election) t. We then estimate the relationship between regional unionization rates and district-level electoral outcomes using the following equation:

$$Electoral\_outcome_{cdt} = \alpha + \beta \text{ Union\_rate}_{cd(r)t} + \lambda_{ct} + \epsilon_{rt}$$

Here, d indexes electoral districts, $d(r)$ maps the electoral districts to the corresponding NUTS-2 regions, c indicates countries, and t denotes years. $Union\_rate_{crt}$ represents union density in NUTS-2 region r and year t, while $Electoral\ outcome_{cdt}$ is one of the dependent variable used to characterize electoral outcomes in district d at time t. The model includes country-year fixed effects λ_{ct} and we cluster standard errors at the NUTS-2 regional and year level.

To provide a comprehensive analysis of the relationship between union density and voting patterns, we estimate our baseline regression using four alternative outcome variables: the vote shares for Radical Left, Mainstream Left, Mainstream Right, and Radical Right party families.

Table 5 presents the estimates from four models that regress district-level voting behavior on the unionization rate (as predicted by the best MrP specification overall) in the corresponding NUTS-2 region. The data encompass elections held between 2002 and 2019 across fifteen western European countries.

Our findings indicate that in regions with higher union density, mainstream left parties tend to secure larger vote shares compared to regions with lower union density. Conversely, union density is negatively associated with support for the other party families. While these results should not be interpreted causally, they provide empirical support for the expectation that social democratic parties are stronger in contexts where trade unions remain robust.

5 Conclusions

This study contributes to the ongoing debate on the political consequences of economic

Table 5: Voting and unionization

VARIABLES	(1) Radical Left	(2) Mainstr. Left	(3) Mainstr. Right	(4) Radical Right
Union rate	-0.230***	0.850***	-1.031***	-0.102**
(best)	[0.055]	[0.176]	[0.188]	[0.049]
Estimator	OLS	OLS	OLS	OLS
Observations	7,194	7,194	7,194	7,194
R-squared	0.531	0.351	0.363	0.728
Country-year FE	X	X	X	X
Std dev. Y	0.0493	0.121	0.135	0.0408
Std dev. X	0.0962	0.0962	0.0962	0.0962
Magnitude	-0.449	0.679	-0.734	-0.241

Robust standard errors in brackets
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

transformations, particularly those driven by technological change. Our analysis has shown that automation plays a significant role in undermining the strength of trade unions in Western European countries, thereby weakening a key institutional pillar that traditionally supported social democratic parties.

Our findings provide the first comprehensive evidence that industrial automation, especially through the adoption of robots, has significantly contributed to the decline in union density across Western European regions. By systematically analyzing the regional, sectoral, and individual-level impacts of automation, we have shown that robotization reshapes labor markets in ways that weaken the role of trade unions. This occurs not so much through an erosion of union strength within sectors directly affected by automation, but rather through a broader labor market shift away from traditionally unionized industries toward less unionized sectors.

The reallocation of labor driven by technological change—from unionized to less unionized sectors—emerges as a key mechanism behind the decline in union density. The implications of this shift extend beyond the economic sphere, as lower union density correlates with a weakening of mainstream left political support and a rise in support for radical-right parties. This political dimension highlights the broader societal consequences of declining union strength, as it may contribute to the growing appeal of populism and nationalism in response to economic dislocation.

In line with Kitschelt (2012), we argue that the erosion of the link between working-class constituencies and social democratic parties is pivotal in understanding the political realignment seen in many Western democracies. As automation displaces workers and diminishes the influence of trade unions, the economic grievances of these "losers" of structural changes are increasingly channeled into support for nationalist and radical-right parties. Our evidence suggests that the direct impact of automation on trade unions exacerbates these grievances, further fueling the shift away from traditional left-wing parties.

The innovative dataset developed in this study, which combines census and survey data to measure union density at the region-by-industry level, represents a significant contribution to the field. This data opens up new avenues for research into the role of trade unions in contemporary labor markets. By making this dataset publicly available, we aim to facilitate further empirical investigations into the economic, social, and political ramifications of declining union influence.

The weakening of trade unions not only reflects broader trends in labor market transformations but also provides a crucial explanatory factor for the rise of populist and radical-right movements. As technological advancements continue to reshape the labor market, understanding these dynamics will be essential for both scholars and policymakers aiming to address the root causes of political discontent and polarization.

References

ACEMOGLU, D. and AUTOR, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, vol. 4, Elsevier, pp. 1043–1171.

— and RESTREPO, P. (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy*, **128** (6), 2188–2244.

AHLQUIST, J. S. (2017). Labor unions, political representation, and economic inequality. *Annual Review of Political Science*, **20**, 409–432.

— and DOWNEY, M. (2023). The effects of import competition on unionization. *American Economic Journal: Economic Policy*, **15** (4), 359–89.

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).

AUTOR, D. H. (2015). Why are there still so many jobs? the history and future of workplace automation. *Journal of Economic Perspectives*, **29** (3), 3–30.

— and DORN, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American economic review*, **103** (5), 1553–1597.

BALCAZAR, C. F. (2022). Unions and robots: Automation and the power of labor. *Available at SSRN: https://ssrn.com/abstract=4360315*.

BECHER, M. and STEGMUELLER, D. (2020a). Global economic shocks, local economic institutions, and legislative responses. *Unpublished manuscript*.

— and — (2020b). Reducing unequal representation: The impact of labor unions on legislative responsiveness in the u.s. congress. *Perspectives on Politics*, **19** (1), 92–109.

— and — (2024). Global competition, local unions, and political representation: Disentangling mechanisms. *American Journal of Political Science*.

—, — and KÄPPNER, K. (2018). Local union organization and law making in the us congress. *The Journal of Politics*, **80** (2), 539–554.

BETZ, H.-G. (1993). The new politics of resentment: radical right-wing populist parties in western europe. *Comparative politics*, pp. 413–427.

— (1994). *Radical Right-Wing Populism in Western Europe*. Palgrave Macmillan.

— and MERET, S. (2012). Right-wing populist parties and the working-class vote: What have you done for us lately? In *Class Politics and the Radical Right*, Routledge, pp. 107–121.

BRANCATI, D. (2016). Global elections database. *Global Elections Database, New York*, **15**.

CHIACCHIO, F., PETROPOULOS, G. and PICHLER, D. (2018). The impact of industrial robots on eu employment and wages: A local labour market approach. *Bruegel Working Paper*, **2018** (04).

COLANTONE, I. 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.

DAL BÓ, E., FINAN, F., FOLKE, O., PERSSON, T. and RICKNE, J. (2019). Economic losers and political winners: Sweden's radical right. *Working Paper*.

DAUTH, W., FINDEISEN, S. and SUEDEKUM, J. (2018). Adjusting to robots: Worker-level evidence. *Opportunity and Inclusive Growth Institute Working Papers*, **13**.

EBBINGHAUS, B. and VISSER, J. (2000). *Trade Unions in Western Europe since 1945*. Societies of Europe, Palgrave Macmillan.

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.

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

GALLEGO, A., KURER, T. and SCHÖLL, N. (2018). Not so disruptive after all: How workplace digitalization affects political preferences. *Economics Working Papers*, **1623**.

GELMAN, A. and HILL, J. (2007). Data analysis using regression and multilevel/hierarchical models. *Cambridge University Press*.

—, LAX, J., PHILLIPS, J., GABRY, J. and TRANGUCCI, R. (2019). Using multilevel regression and poststratification to estimate dynamic public opinion. *Unpublished manuscript, Columbia University*, **2**.

— and LITTLE, T. C. (1997). Poststratification into many categories using hierarchical logistic regression. *Survey Methodology*, **23** (2), 127–135.

GOOS, M., MANNING, A. and SALOMONS, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American economic review*, **104** (8), 2509–2526.

GRAETZ, G. and MICHAELS, G. (2018). Robots at work. *The Review of Economics and Statistics*, **100** (5), 753–768.

HÄUSERMANN, S. and KITSCHELT, H. (eds.) (2023). *Beyond Social Democracy: The Transformation of the Left in Emerging Knowledge Societies*. Cambridge, United Kingdom: Cambridge University Press Assessment.

IM, Z. J., MAYER, N., PALIER, B. and ROVNY, J. (2019). The losers of automation: A reservoir of votes for the radical right? *Research & Politics*, **6** (1), 2053168018822395.

KAINE, S. and JOSSERAND, E. (2019). The organisation and experience of work in the gig economy. *Journal of Industrial Relations*, **61** (4), 479–501.

KITSCHELT, H. (2012). Social class and the radical right: Conceptualizing political preference formation and partisan choice. *J. Rydgren (ed.), Class Politics and the Radical Right*, pp. 224–251.

KOLLMAN, K., HICKEN, A., CARAMANI, D., BACKER, D. and LUBLIN, D. (2019). Constituency-level elections archive (clea)(data file and codebook). *Ann Arbor, MI: Center for Political Studies, University of Michigan (Producer and Distributor). Available at: http://www. electiondataarchive. org*.

KRIESI, H. (1998). The transformation of cleavage politics the 1997 stein rokkan lecture. *European Journal of Political Research*, **33** (2), 165–185.

KRISTAL, T. (2013). The capitalist machine: Computerization, workers' power, and the decline in labor's share within u.s. industries. *American Sociological Review*, **78** (3), 361–389.

— and COHEN, Y. (2015). What do computers really do? computerization, fading pay-setting institutions and rising wage inequality. *Research in Social Stratification and Mobility*, **42**, 33–47.

LAX, J. R. and PHILLIPS, J. H. (2009). Gay rights in the states: Public opinion and policy responsiveness. *American Political Science Review*, **103** (3), 367–386.

LEEMANN, L. and WASSERFALLEN, F. (2020). Measuring attitudes–multilevel modeling with post-stratification (mrp). In *The Sage Handbook of Research Methods in Political Science and International Relations*, SAGE Publications Ltd, pp. 371–384.

LEIGHLEY, J. E. and NAGLER, J. (2007). Unions, voter turnout, and class bias in the u.s. electorate, 1964–2004. *The Journal of Politics*, **69** (2), 430–441.

LEVI, M. (2017). Labor unions as critical intermediate associations. *FSI Memo*, **2017** (3).

MEYER, B. (2019). Financialization, technological change, and trade union decline. *Socio-Economic Review*, **17** (3), 477–502.

PARK, D. K., GELMAN, A. and BAFUMI, J. (2004). Bayesian multilevel estimation with poststratification: State-level estimates from national polls. *Political Analysis*, **12** (4), 375–385.

PIAZZA, J. (2001). De-linking labor: Labor unions and social democratic parties under globalization. *Party Politics*, **7** (4), 413–435.

PONTUSSON, J. and RUEDA, D. (2010). The politics of inequality: Voter mobilization and left parties in advanced industrial states. *Comparative Political Studies*, **43** (6), 675–705.

PRZEWORSKI, A. and SPRAGUE, J. (1986). Paper stones: A history of electoral socialism.

RENNWALD, L. and MOSIMANN, N. (2023). *Union Membership and Electoral Demand for Redistribution among Left-authoritarians*. Working Paper 37, FORS, Lausanne, Switzerland, Line.Rennwald@fors.unil.ch.

RODRIK, D. (1998). Has globalization gone too far? *Challenge*, **41** (2), 81–94.

RUGGIE, J. G. (1982). International regimes, transactions, and change: Embedded liberalism in the postwar economic order. *International Organization*, **36** (2), 379–415.

VISSER, J. (2019). Ictwss: database on institutional characteristics of trade unions, wage setting, state intervention and social pacts.

WESTERN, B. and ROSENFELD, J. (2011). Unions, norms, and the rise in us wage inequality. *American Sociological Review*, **76** (4), 513–537.

Robots Replacing Trade Unions: Novel Data and Evidence from Western Europe

Online Appendix

Paolo Agnolin, Massimo Anelli, Italo Colantone, and Piero Stanig

Contents

A Data sources

Country	1st stage data	2nd stage data	Validation data available
Austria	ESS	Austrian National Statistics Offices (National LFS)	-
Belgium	ESS	Statbel (National LFS)	-
Finland	ESS	Statistics Finland	Statistics Finland
France	ESS	Census (IPUMS)	-
Germany	ESS	Mikrozensus (National LFS)	-
Greece	ESS	Census (IPUMS)	-
Ireland	ESS	Census (IPUMS)	-
Italy	ESS	Rilevazione sulla Forza Lavoro (National LFS)	-
The Netherlands	ESS	Survey Workforce New Series EEBnw (National LFS)	-
Norway	ESS	Statistics Norway	Statistics Norway
Spain	ESS	Census (IPUMS)	-
Sweden	ESS	Statistics Sweden	-
Switzerland	ESS	National LFS + Admnistrative data	-
Portugal	ESS	Census (IPUMS)	-
United Kingdom	ESS	Office for national statistics (National LFS)	National LFS

B Data harmonization

We harmonize six key variables using data from the European Social Survey (ESS) and national census data for 15 European countries. The objective is to create consistent variables for age, education, region, industry, occupation, and gender.

For the regional level, we use the NUTS-2 classification for all countries except the UK and Germany, where data are at the NUTS-1 level. Our data excludes Northern Ireland, Ceuta and Melilla, and the French overseas territories.

The education variable is categorized into three levels: “no secondary education”, “secondary education”, and “tertiary education”. For the occupation variable, we adopt the ISCO-88 classification at the 2-digit level. When original data are reported in different classifications (e.g. ISCO-08, SOC, or national classifications), we apply crosswalks to convert these to ISCO-88.

Similarly, the industry variable is classified according to NACE Rev. 1.1 at the 2-digit level. We also implement crosswalks to translate data from NACE Rev. 2 to NACE Rev. 1.1. This process ensures that our variables are standardised and comparable across different data sources, countries, and years.

Table B.1: Summary of harmonized variables and categories

Variable	**Categories**
Gender	Male, Female
Education	No secondary education, Secondary education, Tertiary education
Age	Below 25, 25-34, 35-44, 45-54, 55-64, Over 65
Region	NUTS-2 (NUTS-1 UK, DE)
Occupation	ISCO-88 2d
Industry	NACE Rev. 1.1. 2d

We provide a detailed account of the data sources, harmonization, and cleaning processes for the second stage data of MrP for each country.

Austria

Source of data: The data were obtained from the Austrian National Statistics Office. Microcensus data are available free of charge via AUSSDA. The microcensus provides information on housing and internationally comparable data on employment, unemployment, and education. The data are available quarterly (March, June, September, December). Documentation is accessible online at: https://webapps.ilo.org/surveyLib/index.php/catalog/6535/related-materials.

Harmonization and cleaning:

- **Education, age, industry, gender, region, self-employment, labor force:** These variables were treated consistently with the procedures applied to other countries.
- **Year:** Data were available only for the years 2004, 2006, 2008, 2010, 2012, 2014, 2016, and 2018. Consequently, the 2004 data were duplicated and imputed for the year 2002.

- **Occupation:** A crosswalk was employed to convert ISCO-08 3-digit codes to ISCO-88 2-digit codes.

Belgium

Source of data: The data were sourced from the EU Labour Force Survey (EU LFS) microdata through Statbel, the Belgian statistical office. An official Microdata request was required to obtain the data. The dataset covers 23 years of the survey, from 1999 to 2021.

Harmonization and cleaning:

- **Occupation, education, age, industry, gender, labor force:** These variables were treated consistently with the procedures applied to other countries.
- **Occupation:** ISCO-88 2-digit classification was used. Data classified according to ISCO-08 from 2011 onwards were converted to ISCO-88 2-digit.
- **Labour force:** Only those in the labor force were retained.
- **Self-employment:** Self-employed individuals (with or without employees) and unpaid family workers were filtered and excluded using the variable `stapro`. The remaining category comprised employees and unemployed individuals in `wstator`. Others were dropped.
- **Region:** NUTS2 regions were used. When possible, the region of residence was substituted with the region of work.
- **Weights and context variables:** Relative weights over absolute frequencies were used. Context variables were merged at the NUTS1 level.

Finland

Source of data: The data were provided by Statistics Finland, consisting of two tables: one based on census data and the other on union density (for external validation). These tables were constructed from administrative data, with basic population information derived from the Finnish Population Information System. Employment data were based on various registers, including the national income register.

Harmonization and cleaning:

- **Education, age, industry, gender, region, self-employment, labor force:** These variables were treated consistently with the procedures applied to other countries.
- **Occupation and industry:**
  - 2001-2006: Census data were classified using ISCO-88 and NACE Rev.1.1.
  - Note: For the years 2001-2003, occupation data were unavailable, so the distribtuion of occupation was imputed from 2004.
  - 2007-2009: Census data contained ISCO-88 and NACE Rev.2 (harmonized to NACE Rev.1.1).
  - 2010-2020: Census data contained ISCO-08 and NACE Rev.2 and were harmonized to ISCO-88 and NACE Rev.1.1.

France

Source of data: The data were sourced from IPUMS International, which collects and distributes census microdata globally. The relevant datasets include the 2006 Census (cycle 2004-2008, rolling census) and the 2011 Population Census (cycle 2009-2013), both provided by INSEE (Institut National de la Statistique et des Études Économiques).

Population universe: The dataset includes residents of France of any nationality, excluding French citizens living abroad, foreign tourists, and transient individuals.

Harmonization and cleaning:

- **Education, age, industry, gender, region, self-employment, labor force:** These variables were treated consistently with the procedures applied to other countries.

- **Year:** Data were available for the years 2006 and 2011.
- **Weight interpolation (2007-2010):** Linear interpolation was used to generate weights for 2007-2010 based on weights from 2006 and 2011:

$$\text{perwt\_year} = \text{perwt\_2006} + \frac{1}{5} \times (\text{perwt\_2011} - \text{perwt\_2006}) \times (\text{year} - 2006)$$

- **Weight extrapolation (2002-2005)**
  - We sum weights within each stratum industry-region for 2006 and ensures all records within each stratum (industry-region-2006) have the same summed weight (total sum of weights for each stratum industry-region-2006).
  - we generate a share variable by dividing the weights by the total weight within each stratum. (share variable sums to one within each stratum)
  - we repeat the summing process for weights from year 2007.
  - we generate weights for 2002-2005 using extrapolation based on weights from 2006 and 2007
  - the extrapolated weights are adjusted using the proportional share within each stratum.
- **Weight extrapolation (2012-2018)**
  - We sum weights within each stratum industry-region for 2011 and ensures all records within each stratum (industry-region-2011) have the same summed weight (total sum of weights for each stratum industry-region-2011).
  - we generate a share variable by dividing the weights by the total weight within each stratum. (share variable sums to one within each stratum)
  - we repeat the summing process for weights from year 2010.

  - we generate weights for 2012-2018 using extrapolation based on weights from 2010 and 2011
  - the extrapolated weights are adjusted using the proportional share within each stratum.

- **Occupation:** Detailed conversion procedures were applied to harmonize occupational classifications to ISCO-88. For the 2006 Census, crosswalks were used to convert PCS 1982 values to ISCO-88. Missing categories were manually recoded using additional crosswalks. For the 2011 Census, PCS 2003 values were converted using similar procedures.

Germany

Source of data: The data were obtained from the Forschungsdatenzentrum, Statistical Offices of the Federation and the Federal States, through the German Microcensus. This is the largest annual household survey in Germany, conducted since 1957, covering approximately one percent of the population (about 810,000 individuals).
Harmonization and cleaning:

- **Age, education, industry, gender, self-employment, labor force:** These variables were treated consistently with the procedures applied to other countries.
- **Region:** NUTS-1 level.
- **Year:** The available years were 2002, 2004, 2006, 2008, 2010, 2012, 2014, 2016, and 2018.

Greece

Source of data: The data were sourced from IPUMS International, which collects and distributes census microdata globally. The relevant datasets include the National Population

Housing Census 2001 by the National Statistical Service of Greece and the 2011 Population and Housing Census by the Hellenic Statistical Authority (El.STAT).

Population universe: The dataset includes all individuals (present or temporarily absent) regardless of age, gender, and citizenship, residing usually within the census tract boundaries.

Harmonization and cleaning:

- **Occupation, education, age, industry, gender, region, self-employment, labor force:** These variables were treated consistently with the procedures applied to other countries.
- **Year:** Data were available only for 2001 and 2011.
- **Weight interpolation (2002-2010):**

$$\text{perwt\_year} = \text{perwt\_2001} + \frac{1}{10} \times (\text{perwt\_2011} - \text{perwt\_2001}) \times (\text{year} - 2001)$$

- **Weight extrapolation (2012-2018):**
  - we sum weights within each stratum industry-region for 2011 and ensures all records within each stratum (industry-region-2011) have the same summed weight (total sum of weights for each stratum industry-region-2011).
  - we generate a share variable by dividing the weights by the total weight within each stratum. (share variable sums to one within each stratum)
  - we repeat the summing process for weights from year 2010.
  - we generate weights for 2012-2018 using extrapolation based on weights from 2010 and 2011.
  - the extrapolated weights are adjusted using the proportional share within each stratum.

Ireland

Source of Data: IPUMS International provides census microdata from various countries. For Ireland, we used data from the Central Statistics Office: Census of Population for the years 2002, 2006, 2011, and 2016.

Population Universe: All persons present in Ireland on the census night, including visitors and usual residents who were temporarily absent but answered a subset of questions.

Variables: Education, age, gender, self-employment, and labor force variables are treated similarly to those in other countries.

Years: Data are available for 2002, 2006, 2011, and 2016. Missing years were addressed as follows:

- **Interpolation of Weights:**
  - For 2003-2005:
  $$\text{perwt\_year} = \text{perwt\_2002} + \frac{1}{4} \times (\text{perwt\_2006} - \text{perwt\_2002}) \times (\text{year} - 2002)$$
  - For 2007-2010:
  $$\text{perwt\_year} = \text{perwt\_2006} + \frac{1}{5} \times (\text{perwt\_2011} - \text{perwt\_2006}) \times (\text{year} - 2006)$$
  - For 2012-2015:
  $$\text{perwt\_year} = \text{perwt\_2011} + \frac{1}{5} \times (\text{perwt\_2016} - \text{perwt\_2011}) \times (\text{year} - 2011)$$
- **Extrapolation of Weights for 2017 and 2018:** Data from 2016 are duplicated for these years.

Industry: For 2002 and 2006, NACE rev. 1.1 was used. For 2011, NACE rev. 2 and ISIC rev. 4 (3-digit) were employed, so we converted the data to rev 1.1. In 2016, NACE rev. 2 (sections) was initially utilized. Given that the NACE rev.2 sections, in fact, cannot be uniquely assigned to NACE rev.1.1 categories, we assigned weighted values from NACE rev.1.1 to the opbservations of the 2016 dataset. The weights for the distribution of 2-digits NACE rev.1.1 industry classification codes among industry subsections in 2016 are derived from the Labour Force Survey (LFS) of the United Kingdom for the same year. This choice of using UK data is made because a similar survey is not available for Ireland in 2016, which would have made it of course unnecessary to rely on IPUMS international data for Ireland. The UKLFS dataset for 2016 is therefore used to determine the distribution weights of NACE rev.1.1 2-digit codes within each section of the NACE rev.2 standard. To increase the precision of such conversion, the weighting is performed for ideal types of workers defined by the variables gender, age, education level, and self-employment status.

Occupation: Harmonization involved mapping occupation codes (SOC 2000 and SOC 2010 3 digits) to ISCO-88 1-digit codes. For the 2016 data, SOC 2010 1-digit codes were employed, so we assigned ISCO-88 1-digit codes across all years for consistency.

Italy

Source of Data: Weights for Italy are derived from Eurostat data (ITFS) obtained from ISTAT. The "Rilevazione sulla Forze di Lavoro" (RFL) survey is conducted quarterly and involves a representative sample of families across Italy.

Population Universe: Families are randomly selected and interviewed multiple times over a 15-month period.

Variables: Occupation, education, age, industry, gender, region, self-employment, and labor force variables are harmonized similarly to other countries.

Years: Data are available for 2004, 2005, 2006, 2012, 2016, and 2018. Missing years were handled as follows:

- Interpolate weights for 2007-2011 using 2006 and 2012 data.

- Interpolate weights for 2013-2015 using 2012 and 2016 data.
- Interpolate the weight for 2017 using 2016 and 2018 data.
- Extrapolate weights for 2002-2003 using data from 2004 and 2005, then adjusts by an earlier share.

Netherlands

Source of Data: The EBBnw (Survey Workforce new series) from the Centraal Bureau voor de Statistiek (CBS) is an ongoing survey with annual data from 2003 to 2022. The survey includes all household members aged 15 or older.
Variables: Occupation, education, age, industry, gender, region, self-employment, and labor force variables are harmonized similarly to other countries.
Notes:

- **Region:** NUTS2. Observations with missing regions were dropped, totaling 9,268 cases. Their weights are negligible relative to the total population.
- **Year:** Data are available from 2003 onwards. Observations from 2003 were duplicated and imputed for 2002.
- **Imputation for censored observations:** Cells with less than 10 observations are dropped in the data extracted from the Dutch platform. We imputed 33,012 weights for censored observations (combination of occupation-industry-gender-age-nuts-educ) using the average weight based on the marginal distribution of year, gender, region, self-employment status, labor force status.

Norway

Data Source: Statistics Norway (SSB). The initial objective was to acquire a table for union data (for external validation) and a census dataset. The census dataset as well as a table of data for external validation was obtained from Statistics Norway. The dataset spans 22 years, from 2000 to 2021.

Harmonization and Cleaning:

- **Variables:** Education, age, gender, and industry are treated consistently with other countries.

- **Labour Force:** All information pertains to the labour force, which includes categories such as "Wage earner", "Self-employed", and "Unemployed".

- **Occupation:** Data on occupation is available only from 2003 onwards. Data from 2000 to 2002 is excluded. The classification follows the ILO's International Standard for Classification of Occupations 2008 (ISCO-08). The Norwegian adaptation of this classification is known as STYRK-08, replacing the previous STYRK classification based on ISCO-88. In the dataset, only STYRK codes based on ISCO-88 are available, although some codes may not fully conform to ISCO-88.

- **Industry:** The NACE Rev 1.1 classification is used until 2007. From 2008 onwards, NACE Rev 2 was used and thus we harmonize it to NACE Rev 1.1.

Portugal

Data Source: IPUMS International, which collects and distributes census microdata globally. The project aims to collect, preserve, harmonize, and disseminate data free of charge. The relevant datasets are Censos 2001 (XIV Recenseamento Geral da População e IV Recenseamento Geral da Habitação) and Censos 2011 (XV Recenseamento Geral da População; V Recenseamento Geral da Habitação), Instituto Nacional de Estatística (INE).

Year: Data are available for 2001 and 2011 only.

Interpolation and Extrapolation:

- **Interpolation (2002-2010):** Linear interpolation is used to generate weights for

2002-2010 based on weights from 2001 and 2011. The formula for computing weights is:

$$\text{perwt\_year} = \text{perwt\_2001} + \frac{1}{10} \times (\text{perwt\_2011} - \text{perwt\_2001}) \times (\text{year} - 2001)$$

Weight extrapolation (2012-2018)

- We sum weights within each stratum industry-region for 2011 and ensures all records within each stratum (industry-region-2011) have the same summed weight (total sum of weights for each stratum industry-region-2011).
- we generate a share variable by dividing the weights by the total weight within each stratum. (share variable sums to one within each stratum)
- we repeat the summing process for weights from year 2010.
- we generate weights for 2012-2018 using extrapolation based on weights from 2010 and 2011
- the extrapolated weights are adjusted using the proportional share within each stratum.

Spain

Data Source: IPUMS International. The datasets used are Census of Population and Housing 2001 and Censos de Población y Viviendas 2011, Instituto Nacional de Estadística (INE).

Population Universe: Residents of Spain at the time of the census, excluding those temporarily abroad.

Harmonization and Cleaning:

- **Occupation:** Data are crosswalked from the Spanish classification CNO94 to ISCO-88.

- **Interpolation (2002-2010):** Linear interpolation is used similarly to the Portuguese data.

- **Extrapolation (2012-2018):** The method is similar to that described for Portugal.

Sweden

Data Source: Statistics Sweden (SCB). Data were purchased directly from SCB and cover 21 years between 2001 and 2021.

Harmonization and Cleaning:

- **Variables:** Education, age, gender, region are treated consistently with other countries.

- **Age:** Collected for individuals aged 16 and older between 2001 and 2009; from 2010, data are collected for individuals aged 15 and older.

- **Occupation and Industry:** The classification employed varies over time but it's always eventually converrted to ISCO-88 2 -digits:

  - **2001-2006:** SNI92 (NACE Rev 1.1) and SSYK96 (ISCO-88).

  - **2007-2009:** SSYK96 (ISCO-88) and SNI2007 (NACE Rev 2).

  - **2010-2011:** SSYK96 (ISCO-88) and SNI2007 (NACE Rev 2).

  - **2012-2013:** SNI2007 (NACE Rev 2) and SSYK96_4digits.

  - **2014-2021:** SNI2007 (NACE Rev 2) with additional crosswalks from SSYK2012 to ISCO-08 and then to ISCO-88.

- **Labour Force:** Excludes military personnel and only includes employed or unemployed individuals. A new category for unemployed was constructed using data from reliable sources (from 2005-2021 swedish institute of statistics and from 2001-2004 eurostat).

Switzerland

Source of Data: The EU Labour Force Survey (EU-LFS) is the largest sample survey of European households or individuals. Its main statistical objective is to classify people into three mutually exclusive groups that cover the entire target population: employed, unemployed, and those outside the labour force (neither employed nor unemployed).

- EU-LFS data provide information only at by education, region, occupation (ISCO-88 2 digits), gender, age, year and NACE 1-digit level. To obtain necessary weights at the NACE 2-digit level:
  - We downloaded data from Eurostat on the number of employed workers at the region-NACE 2-digit-year level. However, Eurostat only provides this information for the years 2001, 2005, 2008, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, and 2019. Therefore, data for the years 2002, 2003, 2004, 2006, 2007, 2009, and 2010 are missing.
    - **Interpolation of Weights:**
      - We used linear interpolation to generate weights for the missing years based on available weights from 2001 to 2019. We performed this interpolation using the Stata command `ipolate`.
      - For years after 2009, the data are classified according to the NACE Rev. 2 classification, and we applied our crosswalk to convert to NACE Rev. 1.1 classification.
  - We computed the relative weight for each cell (region-NACE 2-digit-year) compared to the total of the region-NACE 1-digit-year cell. This involves calculating the proportion of employees in each NACE 2-digit category relative to the total number of employees in the corresponding NACE 1-digit category within the same region and year.

- We then applied these weights to the actual EU-LFS data to determine the number of individuals in each age-education-gender-NACE 2-digit-ISCO 2-digit-year cell.
- When needed, we applied the crosswalk to convert from the ISCO-08 occupation classification to ISCO-88 (after 2011) and from NACE Rev. 2 2-digit to NACE Rev. 1.1 2-digit classification (after 2009).

United Kingdom

Source of Data: Weights are constructed using Labour Force Survey data (UKLFS) obtained from the UK Data Service. The Labour Force Survey (LFS) is the largest household study in the UK and provides official measures of employment and unemployment.

Details:

- **Years:** We construct weights for the following years: 2002, 2004, 2006, 2008, 2010, 2012, 2014, 2016, and 2018.
- **Industry:**
  - Up to 2008, the industry of the worker is coded using SIC92. We applied a crosswalk to convert these codes to NACE Rev. 1.1.
  - After 2010, the industry is reported using SIC07, which is identical to the NACE Rev. 2 classification. We applied our usual crosswalk to convert these codes to NACE Rev. 1.1.
- **Occupation:**
  - Before 2010, occupation is reported using the SOC 2000 classification. We applied a crosswalk to convert these codes to ISCO-88.
  - After 2010, the classification employed is SOC 2010, which is identical to SOC 2000. Therefore, we applied the same crosswalk to convert to ISCO-88.

- **Region:** Data are available only at the NUTS-1 level.

- **Education:** Individuals are systematically categorized into three educational levels: "No secondary education," "Secondary education," and "Tertiary education." This categorization is based on a series of variables reflecting educational qualifications, recorded annually.

  - Initially, the variable `HIQUALD`, representing high-quality educational data, was used to assess the distribution of educational qualifications. A new variable, `edu_level`, was created to store the educational level, initially set to missing. Based on `HIQUALD`, `edu_level` was assigned as follows:

    * Value 6 indicated "No secondary education," assigned as 1.
    * Values 4, 3, and 2 indicated "Secondary education," assigned as 2.
    * Value 1 indicated "Tertiary education," assigned as 3.

  - If `HIQUALD` was missing, secondary sources of educational data (TYPHST, QUALS01, QUALCH1, etc.) were used.

  - For individuals categorized under `HIQUALD` value 5 (Other qualification), the age at completion of education (YERQAL2) was used to more accurately assign them to either secondary or tertiary levels based on prevalent age-related trends.

- **Age, Gender, Self-employed, and Labour Force Variables:** These are treated similarly to the other countries.

C List of models

1. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{\text{gndr}} + \alpha^{\text{age}} + \alpha^{\text{edu}} + \alpha^{\text{ind}} + \alpha^{\text{occ}} + \alpha^{\text{region}} + \beta \cdot \text{round} + \beta^{\text{ind\_sec}} \cdot \text{round})$

2. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{\text{gndr}} + \alpha^{\text{age}} + \alpha^{\text{edu}} + \alpha^{\text{ind}} + \alpha^{\text{occ}} + \alpha^{\text{region}} + \beta \cdot \text{round})$

3. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{\text{gndr}} + \alpha^{\text{age}} + \alpha^{\text{edu}} + \alpha^{\text{ind}} + \alpha^{\text{occ}} + \alpha^{\text{region}} + \beta \cdot \text{round} + \beta^{\text{occ\_sec}} \cdot \text{round})$

4. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{gndr} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} + \beta \cdot round + \beta^{age} \cdot round)$

5. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} + \beta \cdot round + \beta^{edu} \cdot round)$

6. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{edu} + \alpha^{age} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} + \beta \cdot round + \beta^{gndr} \cdot round)$

7. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} + \beta \cdot round + \alpha^{occ\_1d,ind\_sec})$

8. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} + \beta \cdot round + \alpha^{occ\_1d,reg})$

9. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} + \beta \cdot round + \alpha^{occ\_1d,age})$

10. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} + \beta \cdot round + \alpha^{occ\_1d,gndr})$

11. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} + \beta \cdot round + \alpha^{occ\_1d,edu})$

12. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} + \beta \cdot round + \alpha^{ind\_sec,reg})$

13. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} + \beta \cdot round + \alpha^{ind\_sec,age})$

14. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} + \beta \cdot round + \alpha^{ind\_sec,gndr})$

15. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{gndr} + \alpha^{age} + \alpha^{edu} + \alpha^{ind} + \alpha^{occ} + \alpha^{region} + \beta \cdot round + \alpha^{ind\_sec,edu})$

16. $Pr(\text{Union}_i = 1) = \text{Probit}^{-1}(\alpha^{\text{gndr}} + \alpha^{\text{age}} + \alpha^{\text{edu}} + \alpha^{\text{ind}} + \alpha^{\text{occ}} + \alpha^{\text{region}} + \alpha^{\text{round}} + \beta \cdot \text{round})$

Figure 5: First-stage models by RMSE

D Measure of exposure to robot adoption: data sources

To create our measure of exposure to robot adoption, we retrieve information on industrial robots for the 15 European countries of interest between 1993 and 2021, except for Greece and Ireland, where the information is available only from 1999 and 2002, respectively. In these latter cases, the robot shock exposure information starts in 2006, considering that before all the robots were classified as "unspecified" and there was no variation across industries. Our focus is on eleven industries encompassing the entire manufacturing sector, primarily corresponding to the subsections of Section D "Manufacturing" of the NACE Rev. 1.1. The subsections are: 1) subsection DA: Food products, beverages and tobacco, 2) DB-DC: Textiles, textile products, leather and leather products, 3) DD: Wood and wood products, 4) DE: Pulp, paper and paper products, 5) DF-DG-DH: Coke, refined petroleum products, nuclear fuel, chemicals, chemical products, man-made fibres, rubber and plastic products, 6) DI: Other non-metallic mineral products, 7) DJ: Basic metals and fabricated metal products, 8) DK: Machinery and equipment, 9) DL: Electrical and optimal equipment, 10) DM: Transport equipment, and 11) DN: Other manufacturing products.

Yearly data on the stock of operational industrial robots by country and industry come from the International Federation of Robotics (IRF). IRF consolidates data reported by robot suppliers also with support from different national robotics associations. The operational stock of robots measures the number of robots deployed in a country each year, assuming an average service life of 12 years with an immediate withdrawal from service afterward (World Robotics 2022 – Industrial Robots). IRF relies on the ISO 8373:2012 definition of an industrial robot as an "automatically controlled, reprogrammable, multipurpose manipulator programmable in three or more axes, which can be either fixed in place or mobile for use in industrial automation applications" (World Robotics 2022 – Industrial Robots).

Sourced from national databases and Eurostat, we retrieve annual employment information for the 15 European countries of interest with the starting year varying between 1988 and

1995. The data is consistently at the NUTS-2 regional level, except for Germany and the United Kingdom, where the data is at the NUTS1 level. In the following, we describe the employment data source by country:

- Austria: Data on regional employment by industry come from Eurostat Regional Employment Statistics, while Austria Statistics provides total regional employment data.
- Belgium: information sourced from the National Bank of Belgium.
- Finland: data provided by Statistics Finland.
- France: data sourced from the National Institute of Statistics and Economic Studies (INSEE).
- Germany: data come from the Federal Employment Agency. These data are available for a fee.
- Greece: data sourced from the Hellenic Statistical Authority - Enterprise Census of 1988.
- Ireland: Data on regional employment by industry from Eurostat Regional Employment Statistics, while the Central Statistics Office provides total regional employment data.
- Italy: data sourced from the National Institute for Statistics - ISTAT.
- Netherlands: data provided by CBS Statistics Netherlands.
- Norway: data provided by Statistics Norway.
- Portugal: data sourced from Statistics Portugal - INE.
- Spain: data come from the National Statistics Institute - INE.

- Sweden: data sourced from Statistics Sweden - SCB. These data are available for a fee.

- Switzerland: data provided by Swiss Statistics - SFSO.

- United Kingdom: data sourced from the Office for National Statistics - ONS.

We normalize the change in the operational stock of robots by the pre-sample thousands of workers employed in the same country and industry. To retrieve the regional-level exposure to robots, we take a weighted summation of the industry-level changes, where the weights capture the relative importance of each industry in each region. Specifically, each weight is the ratio between the pre-sample number of workers employed in a given region and industry, and the pre-sample total number of workers employed in the same region.

In the case of technological instruments, we use annual data from 1993 until 2021 for three variables designed to capture technological shifts in computing: a producer index of computer prices and two indexes of advances in computing. A first instrument is the Producer Price Index by Industry "Electronic Computer Manufacturing", sourced from the Federal Reserve Bank of St. Louis's FRED Economic Data. The electronic computer manufacturing category includes both primary and secondary products. The former encompasses both single-user computers and other types of computers, including host and multiuser computers. This index is annual, not seasonally adjusted, and set to 100 in 2004. A second instrument is the single-thread performance (SpecINT x 10^3), which measures the speed at which a single thread (a sequence of instructions within a program) can be executed by a processor core. A third instrument is the number of transistors in thousands per microprocessor, which influences the complexity and performance of calculations. The information for the last two instruments of advances in computing is sourced from karlrupp.net. The three technological variables multiply an industry-level replaceability index (i.e., the share of hours worked within a given industry in occupations replaceable by robots), as computed by Graetz and Michaels (2018) based on US Census data of 1980.

E Additional results

Figure 6: Robustness: different periods

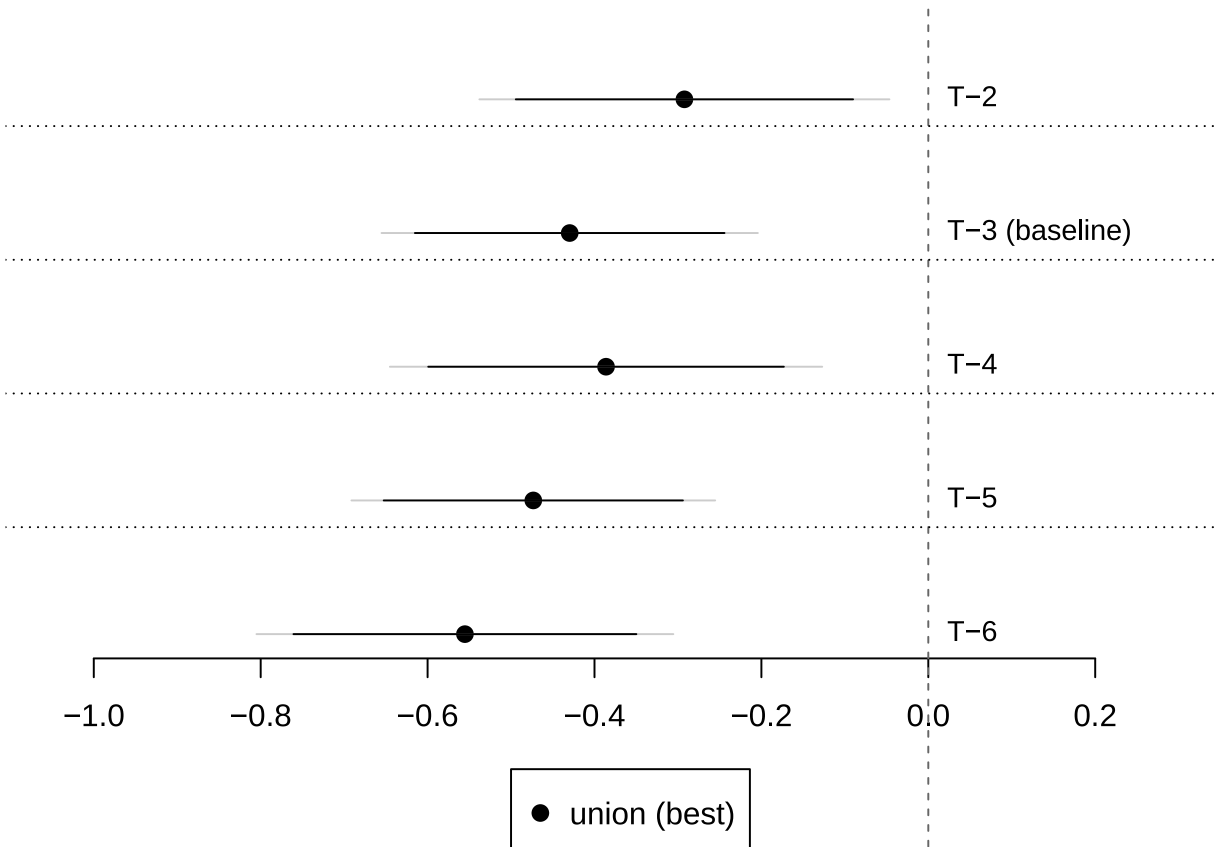


Table E.1: Voting and unionization (baseline)

VARIABLES	(1) Radical Left	(2) Mainstr. Left	(3) Mainstr. Right	(4) Radical Right
Union rate (baseline)	-0.232*** [0.055]	0.847*** [0.176]	-1.037*** [0.188]	-0.104** [0.050]
Estimator	OLS	OLS	OLS	OLS
Observations	7,194	7,194	7,194	7,194
R-squared	0.531	0.351	0.363	0.728
Country-year FE	X	X	X	X
Std dev. Y	0.0493	0.121	0.135	0.0408
Std dev. X	0.0963	0.0963	0.0963	0.0963
Magnitude	-0.453	0.677	-0.739	-0.245

Robust standard errors in brackets
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table E.2: Voting and unionization (relative best)

VARIABLES	(1) Radical Left	(2) Mainstr. Left	(3) Mainstr. Right	(4) Radical Right
Union rate (rel. best)	-0.235*** [0.056]	0.855*** [0.177]	-1.038*** [0.189]	-0.103** [0.050]
Estimator	OLS	OLS	OLS	OLS
Observations	7,194	7,194	7,194	7,194
R-squared	0.531	0.351	0.363	0.728
Country-year FE	X	X	X	X
Std dev. Y	0.0493	0.121	0.135	0.0408
Std dev. X	0.0963	0.0963	0.0963	0.0963
Magnitude	-0.458	0.683	-0.740	-0.243

Robust standard errors in brackets
*** $p<0.01$, ** $p<0.05$, * $p<0.1$