A Report on “Robots and Jobs: Evidence from US Labor Markets” by Acemoglu and Restrepo (2020)

Reviewer 2

February 04, 2026

v1

Disclaimer

This report was generated by large language models, overseen by a human editor. It represents the honest opinion of The Catalogue of Errors Ltd, but its accuracy should be verified by a qualified expert. Comments can be made here. Any errors in the report will be corrected in future revisions.

I am wiser than this person; for it is likely that neither of us knows anything fine and good, but he thinks he knows something when he does not know it, whereas I, just as I do not know, do not think I know, either. I seem, then, to be wiser than him in this small way, at least: that what I do not know, I do not think I know, either.

Plato, *The Apology of Socrates*, 21d

To err is human. All human knowledge is fallible and therefore uncertain. It follows that we must distinguish sharply between truth and certainty. That to err is human means not only that we must constantly struggle against error, but also that, even when we have taken the greatest care, we cannot be completely certain that we have not made a mistake.

Karl Popper, 'Knowledge and the Shaping of Reality'

Overview

Citation: Acemoglu, D., and Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy*, Vol. 128, No. 6, pp. 2188–2244.

URL: https://doi.org/10.1086/705716

Abstract Summary: This study examines the effects of industrial robots on US labor markets using a theoretical model and empirical analysis across commuting zones. The findings show robust negative effects of robot exposure on both employment and wages, distinct from other forms of capital deepening.

Key Methodology: Theoretical model of automation and tasks; empirical analysis using a Bartik-style shift-share measure of exposure to robots, instrumented by European robot adoption trends (IV/2SLS) across US commuting zones (1990–2007).

Research Question: What are the equilibrium effects of industrial robots on employment and wages in local US labor markets?

Summary

Is It Credible?

Acemoglu and Restrepo investigate the equilibrium impact of industrial robots on US labor markets, offering a theoretical and empirical framework to distinguish automation from other forms of technological change. The article posits that unlike capital deepening or factor-augmenting technologies, which generally increase labor demand, robots and automation technologies create a “displacement effect” by directly replacing workers in specific tasks (p. 2195). Based on this framework, the authors claim to provide robust evidence that the adoption of industrial robots has led to negative effects on employment and wages. Specifically, the abstract states that “One more robot per thousand workers reduces the employment-to-population ratio by 0.2 percentage points and wages by 0.42%” (p. 2188). The article further asserts that these effects are “distinct from other capital and technologies” and that the negative impacts are concentrated in manufacturing industries and routine manual occupations (p. 2188).

The credibility of the article’s local estimates relies on a Bartik-style instrumental variable strategy, which uses robot adoption trends in European countries to isolate technological shocks in the US. While this approach is designed to filter out US-specific shocks, the identification strategy faces a significant challenge regarding its heavy reliance on a single sector. The authors acknowledge that “the share of employment in the automotive industry explains 67% of the cross-commuting zone variation in exposure to robots” (p. 2226, footnote 23). Consequently, the instrument may not strictly isolate exogenous technological progress if the global automotive industry experienced common unobserved shocks—such as changes in consumer preferences, trade dynamics, or non-robot manufacturing processes—that simultaneously affected European adoption and US employment. Although the authors perform robustness checks by separating the automotive industry from others, the overwhelming contribution of this single industry to the identifying variation suggests that the results might be less generalizable to the broader economy than the framing implies (p. 2227).

Furthermore, the headline aggregate claims presented in the abstract—that robots reduced the employment-to-population ratio by 0.2 percentage points and wages by 0.42%—are not direct empirical estimates but rather the output of a model simulation. The direct instrumental variable estimates for local labor markets are substantially larger, showing a 0.39 percentage point decline in employment and a 0.77% decline in wages for the 1990–2007 period (p. 2237). The authors derive the smaller aggregate figures by calibrating a structural model to account for trade spillovers and capital income gains (p. 2238). This calibration depends on several parameters, some of which are drawn from non-academic sources, including media reports and corporate white papers (p. 2239, footnote 29). Moreover, there appears to be a numerical inconsistency in the calibration exercise. The parameters and estimates reported in the article do not seem to satisfy the model's own structural equilibrium condition as defined in the appendix, where the left-hand side of the equation evaluates to approximately -0.20 while the right-hand side evaluates to approximately -0.08 (p. A-18). This discrepancy introduces uncertainty regarding the precision of the aggregate effects highlighted in the abstract.

The temporal stability of the findings also warrants scrutiny. The primary analysis focuses on the 1990–2007 period, finding a robust negative impact on employment. However, when the authors extend the analysis to 2014, the negative effect on the employment-to-population ratio diminishes by roughly 40%, falling from -0.388 to -0.250 (p. 2237). The authors suggest this "might reflect the fact that as wages have continued to adjust in the affected commuting zones, some of the initial employment response may have been reversed" (p. 2237, footnote 25). This attenuation of the effect over time is not prominent in the main discussion or conclusion, yet it suggests that the displacement effects of robots may be more transient or susceptible to long-run adjustment mechanisms than the headline results for the pre-recession period imply. Similarly, the magnitude of the baseline employment estimate is sensitive to specific controls; excluding the share of employment in light manufacturing reduces the estimated coefficient by approximately 30% (p. 2216, footnote 20).

Despite these issues, the article makes a significant contribution by empirically distinguishing the effects of automation from general capital accumulation. The analysis shows that controlling for exposure to IT capital, overall capital deepening, and value-added growth does not diminish the estimated negative impact of robots (p. 2228). In fact, these other forms of capital are often positively correlated with labor demand, supporting the authors' theoretical distinction between the displacement effects of automation and the productivity effects of other technologies. The findings regarding heterogeneous impacts are also robust; the negative employment effects are clearly concentrated in routine manual and blue-collar occupations within manufacturing, lending credibility to the mechanism of task displacement (p. 2233). While the precise magnitude of the aggregate effects remains subject to modeling assumptions and the identification relies heavily on the auto sector, the qualitative evidence for a distinct, negative local shock from industrial robots during the 1990s and 2000s appears credible.

The Bottom Line

Acemoglu and Restrepo provide compelling evidence that industrial robots exerted a negative pressure on employment and wages in local US labor markets between 1990 and 2007, an effect distinct from general capital investment. However, the specific aggregate figures highlighted in the abstract are model-based simulations rather than direct measurements and rely on calibration assumptions that appear numerically inconsistent. Furthermore, the estimated negative employment effects weaken significantly when the analysis is extended to 2014, suggesting that the long-term displacement impact may be less severe than the headline results imply.

Potential Issues

Instrumental variable validity: The article's identification strategy may be compromised by its heavy reliance on a single industry. The instrument, which uses robot adoption trends in other advanced economies to predict adoption in US commuting zones, is intended to isolate an exogenous technology shock. However, its variation is overwhelmingly driven by the automotive industry. The authors acknowledge that "the share of employment in the automotive industry explains 67% of the cross-commuting zone variation in exposure to robots" and that this industry contributes the vast majority of the identifying variation in the main estimates (p. 2226, footnote 23). They concede that this means the "estimates may be sensitive to other shocks affecting local labor markets specializing in the automotive industry during this period" (p. 2226, footnote 23). Because the global automotive sector is subject to common shocks—such as shifts in consumer demand, trade policy, or non-robot technological advances—that could simultaneously affect robot adoption in Europe and employment in US auto-centric regions, the instrument may violate the exclusion restriction by capturing these unobserved factors. While the authors conduct robustness checks by separating the automotive industry from others, the core identification remains tied to a single sector, which may limit the generalizability of the findings (p. 2227).

Nature of the headline aggregate findings: The article's most prominent quantitative claims are not direct econometric findings but are the output of a calibrated structural model. The abstract's claim that "One more robot per thousand workers reduces the employment-to-population ratio by 0.2 percentage points and wages by 0.42%" is a simulation result (p. 2188). The direct instrumental variable (IV) estimates are substantially larger. The authors are transparent about this, stating, "To explore these aggregate implications, we need to make further assumptions on cross-commuting zone spillovers (and this suggests greater caution in interpreting these aggregate estimates than the local effects...)" (p. 2238). However, the validity of this simulation rests on several parameters calibrated from sources that are not peer-reviewed academic literature, including popular media reports, a TV show's website, and a Boston Consulting Group report (p. 2239, footnote 29). Presenting these model-dependent calculations in the abstract may give a misleading impression about the certainty of the evidence for the article's aggregate claims.

Inconsistency in the aggregate model calibration: The article's quantification of aggregate effects appears to be mathematically inconsistent with the empirical estimates it is based on. The calibration exercise is designed to match key model parameters to the main IV estimates for local employment and wage effects. However, the reported calibrated values do not satisfy the model's own structural relationship as defined in equation (A48) of the appendix (p. A-18). Using the article's stated IV estimates ($\beta_L = -0.39$, $\beta_W = -0.77$) and its calibrated and assumed parameters ($\phi = 0.25$, $\psi = 0.02$, $\epsilon = 0.17$, and an implied labor share $\omega^L \approx 2/3$), the two sides of the equation do not balance (pp. 2237, 2239, A-18). The left-hand side evaluates to approximately -0.20, while the right-hand side evaluates to approximately -0.08. This numerical discrepancy, while potentially due to rounding, raises questions about the replicability and reliability of the calibration exercise that generates the article's headline aggregate results.

Temporal stability of the employment effect: The negative employment effect weakens substantially over a longer time period, a finding that is not prominently discussed in the main text. The primary analysis focuses on the 1990–2007 period, for which the IV coefficient for the effect on the employment-to-population ratio is -0.388. When the analysis is extended to 1990–2014, this effect shrinks by over a third to -0.250 (p. 2237). This result is discussed only in a footnote, where the authors suggest it "might reflect the fact that as wages have continued to adjust in the affected commuting zones, some of the initial employment response may have been reversed" (p. 2237, footnote 25). The fact that this result is not prominently featured in the abstract, introduction, and conclusion is a notable omission. The finding suggests a more complex dynamic where an initial employment shock may partially reverse over the longer run, which alters the article's primary message about the persistent displacement effects of robots.

Sensitivity of the main employment estimate: The magnitude of the main employment finding is sensitive to the inclusion of a specific control variable. The baseline estimate of the effect of robots on employment depends on controlling for the baseline share of employment in "light manufacturing" (textiles, paper, publishing, and printing). The authors acknowledge in a footnote that the employment estimates "are about 30% smaller in specifications that do not control for light manufacturing" (p. 2216, footnote 20). A comparison of the tables confirms this: the baseline employment coefficient of -0.448 falls to -0.295 when this control is excluded (pp. 2214, A-33). While the authors provide a valid justification for including the control—that these industries were declining for reasons unrelated to robots—the sensitivity of the point estimate to this choice suggests that the precise magnitude of the article's main finding is dependent on this specific modeling decision.

Scope of the theoretical framework: The article's theoretical model omits the creation of new tasks, a potentially important countervailing force to job displacement. The model is built around a framework where technology either automates and displaces labor from existing tasks or augments labor in remaining tasks (pp. 2193–2198). It does not explicitly incorporate a mechanism for the creation of entirely new tasks, job roles, or industries where labor has a comparative advantage. The "productivity effect" in the model is limited to increasing labor demand in non-automated tasks within existing industries. The authors acknowledge this limitation in the conclusion by citing their other work on the "reinstatement effect" of new tasks and noting that some general equilibrium effects "might emerge only slowly" (p. 2241). Nonetheless, the empirical analysis and its interpretation are guided by this displacement-centric model, which may not fully capture the long-run dynamics of technological change.

Interpretation of the local estimates: The article's primary empirical strategy does not account for inter-regional spillovers, meaning the local estimates should be interpreted as partial equilibrium effects. The research design estimates the effect of robot adoption within commuting zones, treating each as an independent unit. The authors acknowledge that this approach does not capture all equilibrium responses, particularly spillovers across regions through trade, supply chains, or migration (p. 2191, footnote 2). For example, a negative shock in a high-exposure commuting zone could reduce its demand for goods and services from other zones, potentially biasing the local estimates. While the article attempts to account for these spillovers in its separate, model-based calculation of aggregate effects, the core local IV estimates do not capture the full range of general equilibrium adjustments.

Generalization of the findings: The article's conclusions may over-generalize from the specific case of industrial robots to the broader categories of automation and artificial intelligence. The empirical analysis is based on data for "industrial robots" as defined by the International Federation of Robotics, with identifying variation coming predominantly from manufacturing sectors during the 1990s and 2000s (p. 2200). The article's conclusion, however, generalizes these findings back to the broader initial framing, discussing the implications for "artificial intelligence, and other automation technologies" (p. 2240). Extrapolating from the effects of industrial arms in factories to the potential societal impacts of modern AI and software automation may not be warranted without further evidence.

Minor methodological and presentation issues: Several smaller issues are present in the analysis and its presentation. First, the article does not discuss the heterogeneity implied by the difference between its weighted and unweighted results. The unweighted regressions show systematically larger negative effects for both employment and wages, suggesting that the impacts of robot exposure may be more severe in less populous commuting zones, a point that is not explored (p. 2214). Second, a calculation in footnote 26 for the total number of jobs lost appears to be internally inconsistent. The text states that an increase of 120,000 robots reduced employment by 756,000 jobs, implying a per-robot job loss of 6.3. However, the same footnote provides a formula that yields a per-robot job loss of 6.5, a minor discrepancy likely due to rounding (p. 2238).

Future Research

Diversified identification strategies: Future work could address the reliance on the automotive industry by exploiting variation in robot adoption within non-manufacturing sectors, such as logistics and warehousing, which have seen significant automation since 2010. Research could utilize firm-level data to construct instruments based on specific technological breakthroughs in sensor or gripping technology that affect a broader range of industries, thereby reducing the dependence on a single sector's global trends.

Long-run adjustment and reinstatement: Given the weakening of employment effects in the 2014 data, future research could investigate the long-run dynamics of automation shocks. Specifically, studies could empirically test for the "reinstatement effect"—the creation of new tasks—by tracking the emergence of new job titles or occupational categories in commuting zones with high historical robot exposure. This would help determine whether the attenuation of negative effects over time is due to wage adjustments or the creation of new types of labor demand.

Empirical measurement of spillovers: To validate the aggregate claims without relying on calibration, future research could attempt to estimate cross-regional spillovers directly. Researchers could estimate a spatial lag model or use data on inter-regional trade flows to quantify how robot adoption in one commuting zone affects prices and demand in connected zones. This would provide an econometric basis for aggregating local effects into a national total, offering a check on the model-based simulations.

isitcredible.com