Automation Is Not a Sufficient Statistic:

Robots, AI, and Inequality

Carlos Chávez [1]

July 2026

Abstract

Empirical work on automation and inequality often collapses technological change into a single exposure index. This paper studies whether that scalar is a sufficient statistic for the direction of inequality change. Across 722 U.S. commuting zones between 2005 and 2019, robot exposure tracks rising wages for both education groups and a compressing skill premium, while AI exposure tracks rising college wages and a widening one. Because the two carry opposite signs, a pooled index of the two explains 4 percent of the cross-zone variation against 31 percent for the separated frontiers, and under controls it changes sign. The sign of the aggregate coefficient therefore depends on the mix of the two frontiers as much as on the strength of either. A two-frontier task model, in which physical and cognitive automation displace and create different tasks, locates the switch at a robot share of roughly two-thirds.

JEL Codes: J23, J24, J31, O33

Keywords: Automation, robots, artificial intelligence, task model, skill premium, inequality

1 Introduction

Does automation widen inequality or compress it? The empirical literature answers both ways with equal confidence, and this paper argues the disagreement itself is the finding. I examine seven widely used measures of automation exposure, spanning patent-based, industry-adoption, occupational-exposure, and routine-task approaches, across 722 U.S. commuting zones. Their coefficients on the college wage premium range from strongly negative to strongly positive. Occupation-level robot exposure compresses the skill premium while industry-level robot deployment widens it. AI patent exposure widens it. Language-model

[1]Center for the Economics of Human Development, University of Chicago. Email: cchavezpadilla@uchicago.edu.

exposure, built as an AI measure, behaves like the robot deployment measures instead. A composite index, under controls, yields a small and statistically insignificant coefficient. Every possible conclusion, that automation raises inequality, lowers it, or has no effect, finds support in at least one well-identified study.

The disagreement is a predictable consequence of aggregating two technological frontiers that push inequality in opposite directions.

A pooled automation index explains 4.1 percent of the cross-commuting-zone variation in skill premium changes. Entering robot and AI exposure as separate regressors explains 30.8 percent of the variation, more than seven times as much, with individually significant, opposite-signed coefficients. Once standard controls are added the pooled coefficient even flips sign and turns insignificant, while the two-index model stays significant. A researcher using only the composite index would conclude, incorrectly, that automation has no distributional effect.

The critique has a precise target. Studies that analyze robot, AI, and software exposure separately, as distinct regressors with separate coefficients, are not subject to the aggregation problem formalized below. Fossen et al. (2022) do this for U.S. individual wages and find that the three exposures carry different signs, which is the pattern a multi-frontier environment produces. Neither are studies that measure a single frontier, such as Graetz and Michaels (2018) on industrial robot adoption, since a scalar index of one frontier aggregates nothing. The critique applies to specifications that summarize technological change with a scalar: aggregate routine-task intensity, composite technology exposure indices, or reduced-form regressions that estimate a single automation coefficient without distinguishing which frontier generates the variation. The distinction is methodological, and both approaches coexist in the work of the same research groups.

I formalize this argument in a two-frontier task model extending Acemoglu and Restrepo (2018b, 2019), in which automation advances along separate physical and cognitive task continua, each with its own displacement rate and reinstatement dynamics. The main the-

oretical result is a masking theorem: because the two frontiers have opposite-signed effects on the skill premium, any scalar index that aggregates them can yield a precisely estimated null even when each frontier has an economically large effect. The model derives a closed-form threshold for the automation mix at which the skill premium effect switches sign, the boundary of the region where the two effects cancel under pooling. Conditional on regularity and at least two opposite-signed frontier effects, the result holds locally without requiring the paper's specific calibration.

The baseline model already generates sign opposition. Asymmetric reinstatement, sorting, and spatial pass-through are what reverse the effects into the directions observed in the data. Robot adoption displaces manual tasks and at the same time creates new roles, in maintenance, supervision, quality control and human-machine coordination, that non-college workers disproportionately fill. AI adoption also creates new cognitive tasks, but at a substantially lower rate, so cognitive displacement is only partly offset. That displacement margin, routine-cognitive exposure, is what the AI-labeled occupational measures capture, and it lowers both wages, non-college workers more, and widens the premium.

A calibrated version of the model reproduces this asymmetry. The physical-frontier effective reinstatement rate that matches the wage moments is 1.23 at the reference measurement mapping, inside the range implied by Humlum (2022)'s firm-level estimates from Danish manufacturing, which come from a different country, unit of analysis and identification strategy. The cognitive rate is finite but substantially weaker, and it is implied by the external task-rate ratio instead of being read off the cognitive moments. The evidence for masking is the reduced-form sign matrix, and the model's role is to show that the mechanism is quantitatively coherent and to locate the automation mix at which pooling reverses sign.

The baseline displacement model correctly predicts sign opposition between frontiers but gets the direction wrong. It predicts that robots raise the skill premium and AI lowers it, but the data show the reverse. The directional failure is informative, because each extension moves a distinct wage margin in the reported cumulative decomposition. Asymmetric rein-

statement is what rationalizes the finding that robot exposure raises both wages. Without task creation in the physical sector, robot shocks cannot raise wages at all. Within-sector task sorting near the AI frontier is what rationalizes the finding that AI exposure depresses non-college wages more than college wages. Without sorting, displacement in the cognitive sector would affect wages more uniformly across skills. Differential spatial pass-through is what rationalizes the finding that robot exposure raises non-college wages more than college wages at the commuting-zone level. Without lower geographic mobility for non-college workers, marginal product gains would not translate into larger local wage increases for the less mobile group. Each extension addresses an independently documented empirical regularity, namely physical-frontier reinstatement (Humlum, 2022), task sorting (Autor et al., 2003), and differential mobility (Notowidigdo, 2020). Along the reported cumulative path, no intermediate specification matches both empirical signs. The reported optimum matches the four independent wage targets to within 0.22 standard errors. A cumulative decomposition shows the progressive reversal. Pure displacement produces opposite signs in the wrong direction, asymmetric reinstatement blunts the robot effect and moves the cognitive one only slightly, from -0.83 to -0.76. Common frontier concentration shifts levels without touching either sign. Differential sorting flips the cognitive premium, and differential mobility, applied after a common pass-through, flips the physical one back.

Five empirical results test the framework.

First, the sign matrix across seven measures aligns with the frontier-loading decomposition. Robot patent exposure produces the reinstatement signature, in which both wage levels rise, non-college workers gain more, and the premium compresses. Industry robot deployment and the two AI-labeled occupational indices produce the mirror image, with both wages falling and the premium widening. A wage-decomposition diagnostic classifies the five measures that trigger a signature and marks the remaining two, routine-task intensity and computerization, as uninformative. The empirical frontier grouping of the five feasibility measures in the classification principal-component analysis (PCA) is established ex ante,

with no outcome information. Among these five (Webb Robot, Webb AI, Language Model Occupational Exposure (LMOE), AI Occupational Exposure (AIOE), and routine-task intensity), principal components find exactly two effective dimensions, one tracking robot patent exposure and one tracking AI patent exposure. Language-model and AI occupational exposure, both built as AI measures, load almost entirely on the physical frontier, correlating near minus 0.95 with robot exposure and near zero with AI patent exposure. The reclassification rests on the data, with no modeling choice. The deployment-based Acemoglu–Restrepo (A&R) Robot measure is held out of the five-measure classification PCA, though it enters the supplementary seven-measure factor analysis, and its empirical grouping is assigned from its construction as a robot-deployment measure. The wage diagnostic separately assigns its displacement signature.

Second, an instrumental variables strategy, adapting the European adoption instrument of Acemoglu and Restrepo (2020), produces a theory-predicted sign reversal. Occupation-level robot vulnerability captures the net effect of displacement and reinstatement, and the ordinary least squares (OLS) coefficient on the skill premium is negative. Industry-level robot deployment, instrumented by European adoption shocks, recovers a deployment-weighted short-run margin whose sign is the one the displacement model predicts, and the IV coefficient is positive, with a first-stage F-statistic of 39.4, though exposure-robust inference does not place it away from zero. The displacement-only model matches the IV signs, and the full model matches the OLS signs. A formal three-channel decomposition, covering geographic leakage, temporal separation, and the wedge between feasibility and deployment measures, explains why applying the same instrument to different endogenous variables generates different sign diagnostics within the same two-frontier structure. An AI instrument based on European AI-adoption shifts produces the cognitive wage-sign pattern in a complementary AI-IV exercise, with a weak first stage and baseline imbalance, and the point estimate is positive on college wages and undetectable on non-college wages.

Third, the masking curve traces the estimated coefficient continuously as the index weight

on robot exposure varies from zero to one. The coefficient declines monotonically from positive (pure AI) through zero to negative (pure robot), crossing the null at a robot weight of 0.67, close to the threshold the structural model implies, 0.66, without coinciding with it. The band of weights over which the pooled coefficient is smaller than its own standard error covers the equal-weight index and robot-tilted indices out to roughly 0.84. The pattern varies across physical-frontier measures. When both frontiers push inequality in the same direction (as with computer adoption), aggregation amplifies instead of masking, so the pathology is specific to mixed-sign environments.

Fourth, within every tercile of baseline college share, robot exposure compresses the skill premium and AI exposure widens it. All nine tercile-specific robot coefficients carry the predicted negative sign. Formal interaction tests fail to reject equal slopes across community types. The geography exercise uses Webb Robot, the routine occupation share measured in 1990, and Babina computer adoption, which correlate with college share in opposite directions (strongly negative, moderately positive, and strongly positive respectively). The Webb Robot and computer-adoption signs are stable across terciles, and the routine-share estimates are mixed. This diagnostic design provides evidence against the possibility that the results reflect divergent trajectories of high- versus low-education communities rather than automation per se. Across the flexible-control battery the robot coefficient is stable. Most specifications move it further from zero, by at most about seven percent without the baseline skill premium and sixteen percent with it, and the kitchen-sink specification moves it by 5.0 and 12.5 percent. The one exception is the tercile-fixed-effects specification with baseline SP, which attenuates it modestly, from -0.0144 to -0.0136.

Fifth, the physical channel changes sign at 2019 in the Webb AI horse race, and not in the specification that uses LMOE as the cognitive control, where both windows are indistinguishable from zero. In the Webb AI pairing, robot exposure compresses the premium through the diffusion window and widens it afterward, and stacking the windows with period-specific controls puts the difference between the two coefficients at $+0.025$ (SE 0.005), so equality

is rejected. The reversal runs through non-college wages, which the same measure raises before 2019 and lowers after, and it survives adjustment for observed worker characteristics. Inside the diffusion window the structure is stable, and neither channel differs detectably between 2005–2010 and 2010–2019 once the controls vary by period. The cognitive channel also weakens after 2019, by −0.009 (SE 0.004), though it remains positive over the full 2005–2024 window.

A calibrated structural model, presented after the evidence, quantifies the mechanism. Three parameters are calibrated by minimum distance to four independent reduced-form wage moments, a calibration conditional on the reference mapping from exposure measures to frontiers. The relative reinstatement rate and the two migration elasticities are disciplined outside the wage data, and the single overidentifying restriction is not rejected. The model reproduces the full sign pattern, locates the masking threshold at a robot weight of 0.66, and puts the pooled-index prediction error at 3.1 log points with sign reversal. Leave-one-moment-out analysis, the external disciplines on mobility and on the reinstatement ratio, and the remaining estimation machinery are in the online appendix.

The contribution extends beyond the robot–AI application. The masking result holds whenever a scalar exposure index aggregates treatments from channels with opposite-signed effects on the outcome. The same three pathologies, sign dependence on weights, existence of a masking zone, and predictable sign error under calibration to aggregates, can arise in trade liberalization, immigration, and climate shocks whenever the relevant treatments have heterogeneous skill bias. The implication is that automation exposure must be treated as a vector-valued object. Empirical measures load on different technological frontiers, so they estimate different economic objects, and disagreement across studies using different measures is a predicted feature of the multi-frontier environment.

Formalizing this observation requires extending the task-based approach to automation. Zeira (1998) is an early canonical statement of automation as machines replacing workers in tasks instead of as factor augmentation, and Acemoglu and Autor (2011) set out the task-

assignment framework this literature builds on. Acemoglu and Restrepo (2018b, 2019) formalized the displacement–reinstatement tradeoff within a single-frontier, single-labor-type framework. Acemoglu and Restrepo (2018a) extended the model to two skill groups and showed that low-skill automation raises inequality while high-skill automation reduces it, but analyzed each frontier in isolation, as separate comparative statics that never interact. That two simultaneous technological changes can have opposing skill biases, and that an aggregate technology measure can conceal them, is documented by Combemale et al. (2021) for within-occupation skill demand in semiconductor production. Their result is a demonstration in one production process. The question left open is when a scalar index is sufficient for the direction of the effect, and how much is lost when it is not. The present paper moves to the joint setting, with two task continua, each with its own automation frontier, analyzed simultaneously. The masking result, cross-frontier spillovers, and the frontier-loading decomposition arise because both frontiers move at the same time. None of these results obtain when each frontier is studied in isolation, whether in the task framework of Acemoglu and Restrepo (2018a) and Acemoglu et al. (2025) or in the nested constant elasticity of substitution (CES) approach of Bloom et al. (2025), which separates industrial robots from AI and which Ribeiro and Prettner (2025) extend across countries, and Lankisch et al. (2019), who study automation and wage inequality without that separation. Both lack the task-level microfoundation needed to generate reinstatement dynamics or aggregation bias. The online appendix provides detailed positioning relative to the broader literature.

The paper proceeds as follows. Section 2 develops the two-frontier framework and derives the masking theorem. Section 3 presents the empirical evidence: the sign matrix, the pooled-versus-decomposed comparison, the IV sign reversal, and the masking curve. Section 4 develops the extensions the data demand, calibrates the model, and quantifies the masking. Section 5 concludes.

2 A Two-Frontier Task Model

2.1 Automation Exposure as a Vector-Valued Object

Before developing the task-based framework, this subsection establishes the core measurement result in reduced form. The argument requires no production function, functional forms, or assumptions about factor markets.

Suppose technological change can be decomposed into $N \geq 2$ components $T = (T_1, \ldots, T_N)$, each representing the advance of a distinct production frontier. Let $\theta_j \equiv \partial \mathcal{P} / \partial T_j$ denote the marginal effect of frontier j on the log skill premium $\mathcal{P} \equiv \ln W_H - \ln W_S$, the gap between college (H) and non-college (S) log wages.

Condition 1 (Heterogeneous skill bias)**.** There exist frontiers j, k such that $\theta_j > 0 > \theta_k$: at least two components of technological change have opposite-signed effects on the skill premium.

Under a local linear approximation that treats the frontier effects θ_j as constant over the relevant variation, the population regression coefficient of $\mathcal{P}$ on a scalar exposure index $M = \boldsymbol{\lambda}' T$ satisfies:

$$\beta_M \equiv \frac{\mathrm{Cov}(M, \mathcal{P})}{\mathrm{Var}(M)} \approx \sum_{j=1}^{N} \tilde{\lambda}_j \, \theta_j \,, \tag{1}$$

where $\tilde{\lambda}_j$ are variance-adjusted weights.[2]

Condition 1 makes three consequences possible: (i) **Sign indeterminacy**, different measures can produce opposite-signed coefficients, (ii) **Masking**, weights exist for which the scalar index returns a null even when every frontier has a nonzero effect, and (iii) **Non-sufficiency**, the scalar M does not identify which frontiers are moving or in what proportion. None of the three follows from the sign restriction alone, because Condition 1 restricts the effect vector and says nothing about the support of the exposures. Consequence (iii) asks

[2]When frontier exposures are uncorrelated, $\tilde{\lambda}_j = \lambda_j \, \mathrm{Var}(T_j) / \mathrm{Var}(M)$. Correlation between frontiers adds covariance terms, and the exact expression is derived in Online Appendix A.5. Section 3.1 shows that the two frontier measures are close to orthogonal in the data, so the effective weights stay near this benchmark.

the scalar to be non-injective on that support, and if the feasible frontier vectors lie on a line the index parametrises, as they would with perfectly collinear frontier measures, the index recovers the whole vector and non-sufficiency fails. Consequence (ii) needs a non-degenerate index able to reach the masking weight, which fails when the frontier vector is degenerate. Consequence (i) needs the covariance condition, since a coefficient from a regression on one measure carries covariance terms as well as the structural effects, so opposite structural signs are necessary and not sufficient for opposite coefficients. What the sign restriction supplies is the possibility of all three, and the task model below supplies the support variation that makes them bind. Those conditions are maintained wherever the reduced-form statements below are read. The model provides microfoundations for *when and why* Condition 1 holds.

The remainder of this section develops these microfoundations for $N = 2$, the minimal case in which masking is non-vacuous and for which frontier-specific measures are available. The general N-frontier framework, including the general masking result and its proof, is in Online Appendix A.1. The $N = 2$ case provides a lower bound on the dimensionality of the masking set, since with more than two technologies the masking hyperplane expands and the single-index approach becomes even less informative.

2.2 Production Technology

Final output combines two task bundles, physical (P) and cognitive (C), through a Cobb-Douglas aggregator:

$$Y = Y_P^{\mu} \cdot Y_C^{1-\mu} \tag{2}$$

where $\mu \in (0, 1)$.

Each task bundle aggregates a continuum of tasks through a CES structure. Within each sector $j \in \{P, C\}$, three task modules, capital, surviving labor, and reinstated labor,

combine with across-module elasticity of substitution $\rho > 1$:

$$Y_j = \left[Q_{Kj}^{\frac{\rho-1}{\rho}} + Q_{Sj}^{\frac{\rho-1}{\rho}} + Q_{Nj}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}} \tag{3}$$

where Q_{Kj} is the capital task module, Q_{Sj} is the surviving labor task module, and Q_{Nj} is the reinstated labor task module. Tasks within each module are aggregated with elasticity $\sigma > 1$.[3][4]

Three factors: robot capital K_R (supplied elastically at rental rate R_R), AI capital K_A (supplied elastically at rental rate R_A), and labor L (supplied inelastically).

Assumption 1 (Technology–Task Specificity)**.** Robot capital produces only physical tasks. AI capital produces only cognitive tasks. Labor can be allocated to either type.[5]

2.3 Task Allocation and Two Frontiers

Within the physical continuum, tasks $z \in [0, I_R]$ are produced by robots, $z \in (I_R, 1]$ by labor. Within the cognitive continuum, tasks $z \in [0, I_A]$ are produced by AI, $z \in (I_A, 1]$ by labor. Each frontier also admits reinstatement. The stock of reinstated labor tasks is $\delta_j \cdot I_j$, so a frontier advance dI_j creates $\delta_j \, dI_j$ new tasks, where $\delta_j \geq 0$ is the reinstatement rate. Reinstated tasks enter production at the boundary productivity $g_{Lj}(I_j)$. This departs from the Acemoglu and Restrepo (2018b) convention, in which new tasks appear beyond $z = 1$ at the top of the labor-productivity ladder. Entry at the boundary is the conservative assignment, because it credits reinstatement with the lowest labor productivity among surviving tasks,

[3] The calibrated model imposes $\rho = \sigma$, which collapses the nested structure to a single CES over all tasks. See Section 4.5.

[4] Three labels are reused and disambiguated by position. The task-module subscript S (as in Q_{Sj}, $\mathcal{A}_{Sj}$) denotes *surviving* tasks, while the worker-group subscript S (as in L_S, W_S, $\varepsilon_S^{\text{mig}}$) denotes non-college workers. The sector index C (as in Y_C and the productivity superscripts a^C) denotes the cognitive sector, while the superscript C on wages and wage-change coefficients (w^C, $\hat{\beta}^C$) denotes college. The sector index P (as in Y_P) is set in a different type from the calligraphic $\mathcal{P}$ that denotes the skill premium. Sector and task-module labels attach to production objects (Y, Q, $\mathcal{A}$, a), and group and education labels attach to labor and wage objects (L, W, w, ε, $\hat{\beta}$).

[5] Assumption 1 could be relaxed to comparative advantage without affecting qualitative results.

so the estimated reinstatement rate is not inflated by an assumed productivity premium on new tasks.

Robot productivity $g_R(z) = A_R(1-z)^{\alpha_R}$, declining in z. AI productivity $g_A(z) = A_A(1-z)^{\alpha_A}$. Labor productivity $g_{Lj}(z) = B_j z^{\beta_j}$, increasing in z.

Definition 1 (Two Frontiers)**.** *The* physical frontier I_R *and* cognitive frontier I_A *are the thresholds at which capital and labor are equally productive in the marginal task of each continuum. Throughout,* I_j *is treated as an independent feasibility cutoff and the comparative statics differentiate through the integration boundary.*

For sector j, the capital task module aggregates automated tasks using the inner CES:

$$Q_{Kj} = \left[\int_0^{I_j} g_{Kj}(z)^{\frac{\sigma-1}{\sigma}}\, dz\right]^{\frac{\sigma}{\sigma-1}} \tag{4}$$

Capital is supplied elastically at the sector rental rate, and automated tasks employ no labor, so no labor allocation decision arises in this module.

The surviving labor module aggregates tasks $z \in (I_j, 1]$ where labor remains:

$$Q_{Sj} = \left[\int_{I_j}^1 \big(g_{Lj}(z)\cdot \ell_S(z)\big)^{\frac{\sigma-1}{\sigma}}\, dz\right]^{\frac{\sigma}{\sigma-1}} \tag{5}$$

The firm allocates labor L_j^S optimally across surviving tasks. The first-order condition yields $\ell_S(z) \propto g_{Lj}(z)^{\sigma-1}$, and substituting back gives:

$$Q_{Sj} = \mathcal{A}_{Sj}\cdot L_j^S, \qquad \mathcal{A}_{Sj} \equiv \left[\int_{I_j}^1 g_{Lj}(z)^{\sigma-1}\, dz\right]^{\frac{1}{\sigma-1}} \tag{6}$$

where $\mathcal{A}_{Sj}$ is the *task content of surviving labor*, a sufficient statistic for the productivity distribution over the labor domain.

The reinstated module holds the stock $\delta_j \cdot I_j$ of new tasks, each at productivity $g_{Lj}(I_j)$. Because all reinstated tasks share the same productivity, labor is allocated equally across

them:

$$Q_{Nj} = \mathcal{A}_{Nj} \cdot L_j^N, \qquad \mathcal{A}_{Nj} \equiv \left(\delta_j \cdot I_j\right)^{\frac{1}{\sigma-1}} \cdot g_{Lj}(I_j) \tag{7}$$

When $\delta_j = 0$, no reinstated tasks exist and Q_{Nj} is absent from the aggregator.

The firm splits total sector labor $L_j = L_j^S + L_j^N$ between surviving and reinstated modules to maximize Y_j. Since Q_{Sj} and Q_{Nj} are both linear in their respective labor inputs, the first-order condition for the outer CES yields a closed-form allocation:

$$\frac{L_j^N}{L_j^S} = \left(\frac{\mathcal{A}_{Nj}}{\mathcal{A}_{Sj}}\right)^{\rho-1} \tag{8}$$

When $\rho > 1$, the module with higher task-content $\mathcal{A}$ attracts a disproportionate share of labor. When $\rho = \sigma$ and $\delta_j > 0$, this reduces to the standard single-CES allocation.

By the envelope theorem, the marginal product of sector labor evaluated at the optimal split is:

$$\frac{\partial Y_j}{\partial L_j} = Y_j^{1/\rho} \cdot Q_{Sj}^{-1/\rho} \cdot \mathcal{A}_{Sj} \tag{9}$$

This expression captures the nested structure, since the marginal worker's contribution depends on sector output Y_j (the outer CES level), the surviving module's current output Q_{Sj} (the within-module scale), and the task content $\mathcal{A}_{Sj}$ (the direct productivity).

Definition 2 (Reinstatement Threshold)**.** *Define $\bar{\delta}_j$ as the* effective *reinstatement rate at which the marginal product of labor (MPL) is unchanged by a frontier advance:*

$$\left.\frac{\partial MPL_j}{\partial I_j}\right|_{\delta_j = \bar{\delta}_j} = 0 \tag{10}$$

The threshold depends on ρ, σ, and the task productivity parameters, and it is defined as a root. The derivation, together with the mapping between the effective rate and the literal task-creation rate, is in Online Appendix A.8.

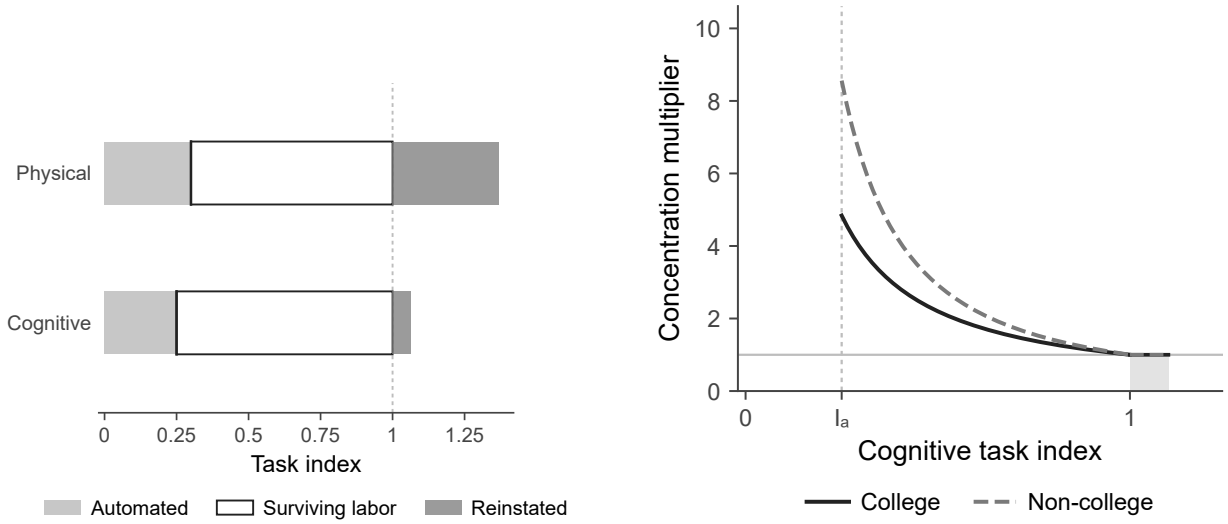
The standard Acemoglu and Restrepo (2022) framework corresponds to $\rho = \sigma$ (or equivalently, reinstated and surviving tasks pooled in a single integral). In that case, the task

content of labor reduces to a single object:

$$\Gamma_j = \left[\int_{I_j}^{1} g_{Lj}(z)^{\sigma-1}\, dz + \delta_j \cdot I_j \cdot g_{Lj}(I_j)^{\sigma-1} \right]^{\frac{1}{\sigma-1}} \tag{11}$$

The nested structure generalizes this by letting the contribution of reinstated tasks be governed by ρ instead of σ. The calibrated model sets $\rho = \sigma$, which returns the single-CES case, so the strength of displacement is governed by the reinstatement rate itself. When $\delta_j \approx 0$, displacement effects are strong.

Figure 1 illustrates the core asymmetry. Robot automation generates a large stock of reinstated tasks, drawn beyond $z = 1$ to display the mass $\delta_R I_R$, while AI extends the continuum by roughly a sixth as much.



(a) Reinstated task mass

(b) Cognitive concentration multipliers

Figure 1: The Two Task Continua and Asymmetric Reinstatement. *Panel A* places both continua on one horizontal scale. Each bar runs from the automated range $[0, I_j]$, through the surviving labor range $(I_j, 1]$, to the effective reinstated mass $\delta_j I_j$ drawn beyond $z = 1$, which enters production at the boundary productivity $g_{Lj}(I_j)$. At the calibrated parameters the physical reinstated mass is about six times the cognitive one. *Panel B* plots the cognitive concentration multipliers, $z^{-\nu}$ for college labor and $z^{-(\nu+\varphi)}$ for non-college, normalized to one at the baseline weights and equal to one on the reinstated block (Section 4.2). Both panels use the parameter values of Section 4: $\hat{\delta}_R = 1.233$, $\hat{\nu} = 1.137$, $\hat{\varphi} = 0.411$, and the implied $\hat{\delta}_A = 0.256$.

Two skill types and equilibrium.

There are L_S non-college and L_H college workers. Within each sector, the two types combine via CES:

$$L_j^{\text{eff}} = \left[\left(a_S^j \cdot \ell_S^j \right)^{\frac{\varepsilon_w - 1}{\varepsilon_w}} + \left(a_H^j \cdot \ell_H^j \right)^{\frac{\varepsilon_w - 1}{\varepsilon_w}} \right]^{\frac{\varepsilon_w}{\varepsilon_w - 1}} \tag{12}$$

with $\varepsilon_w > 1$.[6]

Definition 3 (Comparative Advantage)**.** $\kappa \equiv (a_H^C / a_S^C)/(a_H^P / a_S^P)$. *When* $\kappa > 1$, *college workers have comparative advantage in cognitive tasks.*[7]

In equilibrium each type is indifferent between sectors. The CES aggregator guarantees interior equilibria. The skill premium $\mathcal{P} \equiv \ln W_H - \ln W_S$ is a function of (I_R, I_A) and the endogenous sectoral allocation shares (λ_S, λ_H), where $\lambda_k \equiv \ell_k^P / L_k$ is the fraction of type-k labor working in the physical sector, for $k \in \{S, H\}$. We write $\pi_P \equiv \partial \mathcal{P} / \partial I_R$ and $\pi_C \equiv \partial \mathcal{P} / \partial I_A$. These are the task-model counterparts of θ_j from equation (1). Detailed derivation steps are in Online Appendix A.2.

2.4 Opposite Signs

Definition 4 (Mixed-sign environment)**.** *The economy exhibits* mixed frontier effects *if* π_P *and* π_C *have opposite signs.*

Proposition 1 (Opposite Skill Premium Effects)**.** *Under Assumption 1,* $\varepsilon_w > 1$, $\sigma > 1$, $\rho > 1$, $\kappa > 1$, *and the stability restriction that the sectoral indifference residual is decreasing in the physical-sector labor share (Online Appendix Assumption 1):*

(a) Under pure displacement ($\delta_R = \delta_A = 0$)*:* $\partial \mathcal{P} / \partial I_R > 0 > \partial \mathcal{P} / \partial I_A$.

(b) Sign opposition extends to symmetric reinstatement regimes whenever the two reallocation derivatives keep opposite signs there, which is a condition and not a consequence

[6] Calibrated at 1.5, within the Katz and Murphy (1992) range.

[7] Parametrized symmetrically: $a_S^P = 1$, $a_S^C = 1/\sqrt{\kappa}$, $a_H^P = a_H/\sqrt{\kappa}$, $a_H^C = a_H$. Online Appendix B.7 shows that $\kappa > 1$ is necessary and sufficient for the ordering that signs the premium response to reallocation, and the value used in the calibration (Section 4.5) satisfies it.

of the assumptions above. Which frontier raises the premium can differ from the pure-displacement case.

Proof. See Online Appendix C. Technology–task specificity and $\sigma > 1$ sign the direct effect of each frontier advance on the sectoral indifference residual, in opposite directions across the two frontiers, and the stability restriction signs the denominator of the implicit derivative. Together they deliver opposite reallocation directions (Lemma C.4). Specificity and $\sigma > 1$ alone sign the numerators and not the derivatives. The nested CES structure preserves the monotonic mapping from reallocation to the skill premium, because the additional ρ parameter scales the magnitude of MPL responses without altering their signs (Lemma C.5). $\kappa > 1$ maps reallocation monotonically to the skill premium (Lemmas C.2–C.3). □

Proposition 1 verifies the mixed-sign condition for $N = 2$. The general masking result (Online Appendix A.1) applies by construction. The economic foundations, why opposite signs arise from skill-frontier comparative advantage and asymmetric reinstatement, are in Online Appendix A.3. Under $N = 2$, the automation mix reduces to a scalar $\omega \equiv \Delta I_R/(\Delta I_R + \Delta I_A) \in [0, 1]$.

Corollary 1 (Sign Matrix of Automation Measures)**.** *Measures loading on different frontiers produce opposite-signed skill premium coefficients when the exposure covariance structure does not reverse the ordering, and in the data the correlation between the two frontier measures is close to zero. The exact aggregation bias formulas are in Online Appendix A.5.*

The baseline model is extended by four mechanisms, each an independently documented empirical regularity, stated here informally and formalized in Section 4. *Asymmetric reinstatement*: physical automation creates new labor tasks at a rate δ_R well above the rate at which cognitive automation creates them, and the threshold the cognitive rate would have to clear is itself far higher. *Common cognitive concentration*: tasks near the AI frontier carry more weight for both skill groups. *Differential sorting*: non-college workers in the cognitive sector concentrate in routine tasks adjacent to that frontier. *Spatial pass-through*:

non-college workers are less geographically mobile, so local wage effects are larger for them. Together these determine how each automation channel moves the two wage levels, and therefore which of three wage signatures a measure produces.

Corollary 2 (Wage-Decomposition Diagnostic)**.** *Under the extended model just described (formalized as Proposition 4 in Section 4), a measure's skill-specific wage coefficients place it in one of three signatures. The mapping runs from mechanism to signature, and a measure loading on several channels can produce the same pair of signs from different combinations of loadings and responses, so the signs classify a measure by model consistency without identifying its mechanism. Write* $\hat{\beta}^{NC}$*,* $\hat{\beta}^{C}$*, and* $\hat{\beta}^{SP}$ *for the coefficient on a standardized exposure measure in a regression of, respectively, the non-college wage change, the college wage change, and the skill premium change. The specification is in Section 3.1.* Reinstatement*:* $\hat{\beta}^{NC} > \hat{\beta}^{C} > 0$ *and* $\hat{\beta}^{SP} < 0$*, as net task creation raises both wages, more for non-college, compressing the premium.* Displacement*:* $\hat{\beta}^{NC} < \hat{\beta}^{C} < 0$ *and* $\hat{\beta}^{SP} > 0$*, as net task destruction lowers both wages, more for non-college, widening the premium. In the estimated extended configuration displacement produces this signature on either frontier, on the physical frontier directly through the marginal-product channel and on the cognitive frontier because within-sector sorting directs the displacement onto non-college workers. Under pure displacement the cognitive frontier instead compresses the premium (Proposition 1), and within-sector sorting reverses that sign.* Complementarity*:* $\hat{\beta}^{C} > 0$*,* $\hat{\beta}^{NC} \approx 0$*, and* $\hat{\beta}^{SP} > 0$*, as skill-biased complementarity raises college wages without a detectable non-college effect, widening the premium. A measure with coefficients jointly indistinguishable from zero, or a joint wage pattern that matches none of the three preceding signatures, triggers no signature and is uninformative about mechanism. The frontier follows from the signature together with the measure's task content (Section 3.2). The general formalization is in Online Appendix A.7.*

2.5 The Masking Theorem

Proposition 2 (Distributional Non-Sufficiency of Aggregate Automation)**.** *Under the assumptions of Proposition 1 together with the mixed-sign condition of Definition 4, which part (a) delivers under pure displacement and which part (b) carries as a condition elsewhere, there exist economies E_1 and E_2 with the same aggregate automation index $\mathcal{A} \equiv \Delta I_R + \Delta I_A$ but opposite skill premium dynamics:*

$$\Delta \mathcal{P}^{E_1} > 0 > \Delta \mathcal{P}^{E_2} \tag{13}$$

differing only in the automation mix ω. The scalar $\mathcal{A}$ therefore does not determine the direction of the skill premium response. The two-economy construction is in Online Appendix A.4.

Corollary 3 (Aggregation Bias)**.** *A single-threshold model that assumes an automation mix $\tilde{\omega}$ when the true mix is ω predicts with error, where in continuous time $\omega \equiv \dot{I}_R/(\dot{I}_R + \dot{I}_A)$ is the flow mix:*

$$\left(\frac{d\mathcal{P}}{dt}\right)^{pred} - \left(\frac{d\mathcal{P}}{dt}\right)^{true} = (\dot{I}_R + \dot{I}_A)(\tilde{\omega} - \omega)[\pi_P - \pi_C] \tag{14}$$

See Online Appendix B. □

Proposition 3 (Critical Automation Mix)**.** *There exists a unique $\omega^* \in (0, 1)$ at which the skill premium effect changes sign:*

$$\frac{d\mathcal{P}}{dt} = (\dot{I}_R + \dot{I}_A)\,(\pi_P - \pi_C)\,(\omega - \omega^*), \qquad \omega^* = \frac{\pi_C}{\pi_C - \pi_P}. \tag{15}$$

The sign of $(\pi_P - \pi_C)$ carries the direction. Under pure displacement ($\pi_P > 0 > \pi_C$) the premium falls for $\omega < \omega^*$ and rises above it. Under the calibrated extended model ($\pi_P < 0 < \pi_C$) the direction reverses, so the premium rises for $\omega < \omega^*$ and falls above it, the pattern in Figure 2. Scope conditions and the nesting of single-frontier frameworks are in Online Appendix A.5 and Online Appendix A.6.

Remark 1 (Complementarity and Masking). A minimal extension $I_R^{\text{eff}} = I_R(1 + \gamma I_A)$ introduces a cross-frontier interaction, on the admissible set $\{(I_R, I_A, \gamma) : \gamma \geq 0,\ I_R(1+\gamma I_A) \leq 1\}$ where the effective frontier stays inside the task interval. The interaction is supermodular in (I_R, I_A) only where $F_x + xF_{xx} > 0$, with $x = I_R^{\text{eff}}$, because $\partial^2 Y_P / \partial I_R \partial I_A = \gamma[F_x + xF_{xx}]$ and an increasing concave F can make that negative.[8] If the AI-only and robot-only advances imply opposite-signed skill premium effects, then $\omega^* \in (0, 1)$ exists for all $\gamma \geq 0$, so complementarity shifts the threshold and cannot eliminate masking unless it reverses the endpoint effects. Online Appendix D provides the full analysis.

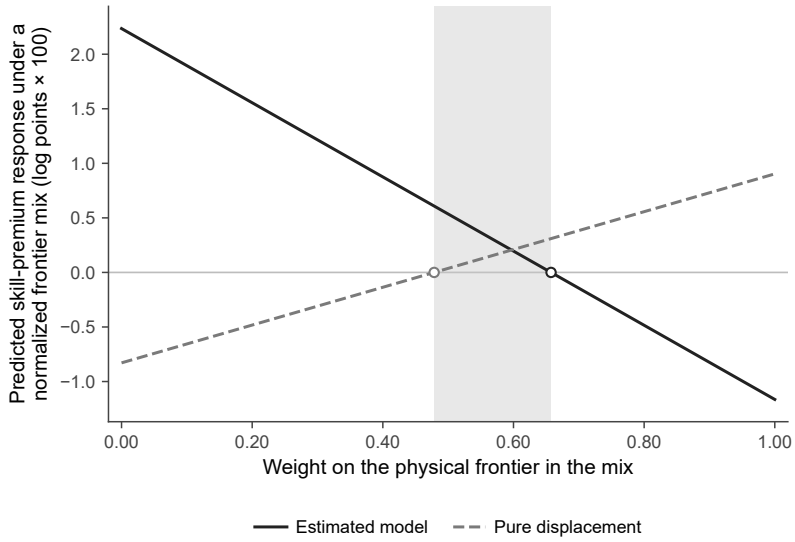


Figure 2: **The Threshold-Shift Diagram.** Each line is the predicted skill-premium response under a normalized frontier mix, $\beta(\omega) = \omega\, m_R + (1-\omega)\, m_A$, where m_j is the model's response to frontier j. Solid line: the calibrated model, crossing zero at $\omega^*_{\text{struct}} = 0.658$. Dashed: pure displacement, crossing at $\omega^*_{\text{pure}} = 0.479$. Gray wedge: the interval between the two crossings. Values from Section 4.7.

The two thresholds sit at $\omega \in [0.479, 0.658]$, so the extensions move the sign boundary by 18 percentage points (Section 4.7). Any design aggregating robot and AI exposure projects a two-dimensional shock onto one dimension. Where the effective ω sits near the operative threshold, the pooled coefficient is attenuated toward zero.

[8] The complementarity parameter γ, and its critical values γ_R^{crit} and γ_A^{crit} in Online Appendix D, are unrelated to the new-task productivity parameters γ_j of the reinstatement module, which always carry a plain sector or skill subscript.

Five thresholds appear in this paper. The pure-displacement threshold $\omega^*_{\text{pure}} = 0.479$ sets both reinstatement rates to zero. The structural threshold $\omega^*_{\text{struct}} = 0.658$ is the mix at which the convex combination of the two single-frontier model premium responses crosses zero at the finite step $\Delta I = 0.08$, with a conditional pairs-bootstrap interval of [0.561, 0.764]. It is the object the estimator computes and the one the pairs bootstrap resamples, so it is the threshold the paper reports. Its first-order counterpart, equation (15) evaluated at the same configuration, is $\omega^*_{\text{FO}} = 0.683$, and the two readings differ by 0.025 in the mix (Online Appendix A.1). The reduced-form target threshold, $\omega^*_{\text{RF}} = 0.669$, is the same finite-step object computed from the empirical premium moments the estimation targets. The empirical masking curve of Section 3.4, built from Webb Robot and Webb AI, crosses zero at $\omega_{\text{curve}} = 0.672$. The interval $[0.479, 0.658]$ between the first two is a *threshold-shift region*, which measures how far the reinstatement, sorting, and mobility extensions move the sign boundary. Both models predict a *positive* effect throughout that interval, by equation (15), and opposite signs outside it, so the interval is not itself a region of near-zero predictions. At $\omega = 0.479$ the extended model predicts +0.61 log points, and its prediction reaches zero only at $\omega^*_{\text{struct}} = 0.658$. Masking is a statement about the estimated coefficient near the true threshold, and the estimated coefficient crosses zero close to the predicted threshold, with the equal-weight index inside the range where it cannot be distinguished from zero.

Testable predictions.

The framework generates four testable predictions: (1) the sign matrix displays systematic variation across measures with different frontier loadings, (2) a pooled index whose effective loading lies near the masking threshold can yield a null even when frontier-specific effects are significant, (3) an OLS–IV sign reversal between occupation-level vulnerability and industry-level deployment, (4) the wage-decomposition diagnostic classifies each measure consistently. All four are tested below.

3 Empirical Evidence

3.1 Data and Measurement

The unit of analysis is the commuting zone (CZ), following the methodology of Autor et al. (2013) and using the 1990 commuting zone definitions of Tolbert and Sizer (1996) with the crosswalks of Autor and Dorn (2013). I construct three outcome variables from the American Community Survey (ACS) 1-year samples (2005–2024): the change in mean log hourly wages for non-college workers ($\Delta \ln w_c^{NC}$), for college workers ($\Delta \ln w_c^{C}$), and in the skill premium ($\Delta \mathcal{P}_c \equiv \Delta \ln w_c^{C} - \Delta \ln w_c^{NC}$, the commuting-zone counterpart of $\mathcal{P}$ from Section 2.3). College workers hold at least a bachelor's degree. Hourly wages are computed as annual wage and salary income divided by the product of weeks worked and usual weekly hours, restricted to workers aged 18–64 with positive wages, deflated to 2019 dollars. The canonical outcome construction uses single-year endpoints (2005 to 2019), and it underlies the sign matrix, the masking analysis, the geography robustness, the IV tables, and the structural moment vector. The composition-adjustment and multi-period analyses (Tables 5 and 9 and the stacked long differences) instead use three-year endpoint windows (2005–2007 average to 2017–2019 average) to smooth endpoint noise in short windows. The two constructions correlate at 0.75 across the commuting zones present in both, population-weighted, and at 0.61 unweighted. The robot coefficient is close but not identical across them, and the notes to Tables 5, 7, 9, and A.3 state the applicable construction.

I analyze three time periods: the robot diffusion period (2005–2019), the post-diffusion period (2019–2024), and the full period (2005–2024). All regressions use long differences. The analysis sample includes the same 722 CZs in every period.[9]

I use seven measures of CZ-level automation exposure, constructed by interacting occupation-

[9]The 1990 CZ partition yields 722 mainland CZs. The analysis sample excludes CZs with fewer than 30 workers per skill-group cell, a restriction that binds in no year. Holding the sample fixed is what makes coefficients comparable across periods, and it requires mapping each survey wave with the PUMA vintage it carries and reading weeks worked from the interval variable in the years when the exact count is not collected (Online Appendix E.1).

or industry-level technology indices with baseline employment composition. All measures are standardized to mean zero and unit standard deviation. Table A.4 summarizes the seven measures.

The two robot measures proxy for physical-frontier automation (ΔI_R) but differ in construction: Webb Robot captures occupation-level *vulnerability* to robot patents, while A&R Robot captures industry-level *deployment* of actual robots. This distinction matters for identification (Section 3.3). The three AI measures proxy for cognitive-frontier automation (ΔI_A) and vary in their definition of AI capability.[10] Routine task intensity and computerization capture broader technology exposure.

Table A.5 reports summary statistics. Two facts emerge from the descriptive statistics in Table A.5. First, the skill premium *compressed* on average, and it kept compressing. Over 2005–2019 non-college wages grew by 18.5 log points against 17.3 for college workers, a population-weighted premium decline of 1.2 log points. From 2019 to 2024 non-college wages grew by 19.6 log points against 17.5, a further decline of 2.1, and over the full window the cumulative compression is 3.3 log points. The sub-period figures sum to the full-period figure up to rounding, because the same 722 commuting zones enter every window. This secular decline in the college premium, running counter to the well-documented increase from 1980 to 2005, is the aggregate pattern that the model's reinstatement mechanism helps explain. Second, there is substantial cross-CZ heterogeneity in all three outcomes. The unweighted p10–p90 range of $\Delta\mathcal{P}$ in 2005–2019 spans 15.6 log points (from -0.123 to $+0.033$), so the average commuting zone experienced compression while a sizable minority saw the skill premium rise. This heterogeneity is the variation that the cross-sectional regressions exploit.

The technology exposure measures also exhibit the correlation structure predicted by the two-frontier model. Webb Robot and Webb AI are only weakly correlated (population-weighted Pearson correlation $r = 0.10$), so the two frontiers generate nearly orthogonal cross-

[10] Acemoglu et al. (2022) use the same Webb and Felten measures to study AI-related hiring at the establishment level, finding that AI-exposed firms reduce non-AI hiring, evidence for the cognitive displacement channel that complements the CZ-level analysis here.

CZ variation. But LMOE and AIOE, despite being constructed as AI measures, correlate at $r = -0.95$ and $r = -0.95$ with Webb Robot, mirroring physical-frontier exposure with opposite sign. In contrast, Webb AI is nearly uncorrelated with LMOE ($r = -0.03$) and AIOE ($r = +0.02$). The correlation between LMOE and AIOE themselves is near-perfect ($r = 0.998$), so the two measure the same underlying construct. These patterns motivate the reclassification analysis in Section 3.2.

Figure A.2 in the appendix documents the relationship across measures. Panel (a) reports pairwise correlations. Panel (b) shows kernel density estimates of the standardized measures.

Unless a table note specifies otherwise, regressions include baseline demographic controls (college share, manufacturing share, female share, nonwhite share, mean worker age) and the initial skill premium SP_0, population weights, and heteroskedasticity-robust standard errors. Throughout, "baseline demographic controls" denotes the five demographic terms and excludes SP_0. Specifications that add or drop SP_0 say so.[11] Every exposure measure is a share-weighted shift-share and both instruments described in Section 3.3 are Bartiks, so the inference concerns raised by Goldsmith-Pinkham et al. (2020), Adão et al. (2019), and Borusyak et al. (2022) apply. The exposure regressions are read under share exogeneity, which the balance tests in Section 3.3 support and for which the commuting-zone robust, state-clustered and Conley standard errors are the right object. The Bartik IV block rests on shock exogeneity instead, in the Borusyak et al. (2022) sense, and there those standard errors do not price dependence across zones loading on the same industry shifter. Nineteen shifters carry that block, with automotive and petrochemicals supplying 0.50 and 0.33 of the identifying variance, so the effective number of shocks is close to three. Under the exposure-robust procedure of Adão et al. (2019) the reduced-form standard error is 0.0014 against 0.0017 for the heteroskedasticity-robust one, but their weak-shock-robust AKM0 confidence set for the same coefficient is unbounded ($p = 0.122$ with the baseline premium among the controls and 0.150 without it, and $p = 0.175$ and 0.422 for the two wage-level reduced

[11] Table OE.4 shows robustness to Conley standard errors (100km, 200km, 500km) and state-level clustering.

forms). Because the robot IV model is just identified, testing a 2SLS coefficient against zero is the same null as testing its reduced form, so the IV block is read as a sign check and not as an independently significant estimate. The Anderson–Rubin sets reported with the IV estimates address first-stage strength and not this shock-level dependence, and the AKM procedures were run on the 722-zone robot IV specifications and not on the 304-zone and 240-zone AI-instrument samples. Online Appendix E.2 reports both procedures.

Each cell of the sign matrix comes from a separate weighted least squares regression

$$Y_c = \alpha + \beta\, M_c + X_c'\Psi + \varepsilon_c, \tag{16}$$

where $Y_c \in \{\Delta \ln w_c^{NC},\ \Delta \ln w_c^{C},\ \Delta \mathcal{P}_c\}$, M_c is one standardized exposure measure, and X_c is the baseline control vector. I write $\hat{\beta}^{NC}$, $\hat{\beta}^{C}$, and $\hat{\beta}^{SP}$ for $\hat{\beta}$ in the three outcome equations. These are conditional projection coefficients. Reading one as a causal effect requires the exposure measure to be exogenous given the controls, which the balance tests of Section 3.3 support without establishing, and the instruments of Section 3.3 identify a deployment effect for a different exposure object and not the effect of these occupation-based measures. The sign matrix is therefore read as a set of conditional associations, and causal language is reserved for the model.

3.2 The Sign Matrix

Corollary 1 predicts that the sign of any automation measure's effect on the skill premium depends on how it loads on the physical versus cognitive frontier, with the two frontiers producing opposite signs. The wage-decomposition diagnostic (Corollary 2) sharpens this. The full pattern of skill-specific wage coefficients places each measure in one of the three signatures (reinstatement, displacement, or complementarity), which the skill premium sign alone cannot. The frontier then follows from the mechanism combined with the measure's task content.

Classifying the AI-labeled measures.

Before examining the full sign matrix, I address a classification puzzle. LMOE (Language Model Occupational Exposure) and AIOE (AI Occupational Exposure) are widely used measures of AI exposure, built on the technical feasibility of automating an occupation's tasks and not on observed adoption (Eloundou et al., 2024; Felten et al., 2021). However, their commuting-zone-level variation is the near-perfect mirror image of robot exposure, and, conditional on baseline composition, tracks routine task intensity instead of AI-specific exposure. Three pieces of evidence support classifying them as displacement measures that load empirically on the physical factor despite their AI labels.

The first piece of evidence is the correlation structure. The population-weighted correlation between LMOE and Webb Robot is $r = -0.95$. CZ-level variation in LMOE is almost entirely the mirror image of robot exposure. AIOE exhibits the same pattern ($r = -0.95$ with Webb Robot). In contrast, Webb AI, constructed using the same natural language processing (NLP) methodology but applied to AI patents, is nearly uncorrelated with LMOE and AIOE ($r = -0.03$ with LMOE, $+0.02$ with AIOE). Both claim to measure AI exposure, yet they share almost no cross-CZ variation.

The second is a variance decomposition. A partial correlation regression decomposes LMOE into its routine-task-intensity (RTI) and AI-specific components:

$$\text{LMOE}_c = 0.128 \cdot \text{RTI}_c + (-0.060) \cdot \text{WebbAI}_c + X_c'\gamma + \varepsilon_c \tag{17}$$

(standardized variables, $R^2 = 0.929$, $N = 722$). Routine task intensity enters with a significant positive coefficient ($t = 5.4$), while Webb AI is statistically insignificant ($t = -1.3$) and economically negligible. Two qualifications keep this regression in its place. The R^2 of 0.929 works largely through the baseline controls, because without them routine task intensity remains the dominant predictor ($\hat{\beta}_{\text{RTI}} = 0.527$, $t = 8.5$, against $\hat{\beta}_{\text{AI}} = -0.066$, $t = -0.5$), but the two variables together account for only about a fifth of LMOE's raw

variation ($R^2 = 0.212$). The Shapley R^2 decomposition is therefore a statement about *incremental* variance beyond the controls, of which 75.7% goes to routine task intensity and 24.3% to Webb AI.[12] The primary fact remains the raw correlation structure of the first paragraph, the -0.95 mirror image with Webb Robot, which no control conditioning produces.

The third is consistency with the sign matrix. If LMOE and AIOE measured AI-specific exposure, they should produce the complementarity signature of Table 1, with positive college wage effects and positive skill premium effects, matching Webb AI. Instead, they produce the displacement signature, with negative coefficients on *both* wage levels, non-college workers hit harder (LMOE: $\hat{\beta}^{NC} = -0.052$, $\hat{\beta}^{C} = -0.031$, AIOE: $\hat{\beta}^{NC} = -0.048$, $\hat{\beta}^{C} = -0.025$), and a positive skill premium coefficient. This is the identical pattern produced by A&R Robot, the canonical physical-frontier displacement measure. The frontier assignment follows from task content and the correlation structure, which place LMOE and AIOE with the physical measures.[13]

For these reasons, I group LMOE and AIOE with the physical-factor measures in the analysis below, on their empirical loading. Their negative coefficients on non-college wages (Table 1) reflect exposure to routine task displacement, a pattern that began in the 1980s with computerization and continued through the robot era, and not recent AI deployment.

The grouping describes empirical loading, not task content. What LMOE and AIOE measure is the automation of routine *cognitive* tasks, so in the structural model they supply the cognitive frontier's displacement moments (Section 4). They co-move with robot exposure, and load on the same factor, because routine-cognitive and routine-manual work concentrate in the same commuting zones. The distinction matters for reading the sign matrix, because these measures behave like the physical-frontier measures in the cross-section while carrying the cognitive continuum's displacement margin in the model, where Webb

[12] AIOE shows the same pattern: $\hat{\beta}_{\text{RTI}} = 0.124$ ($t = 4.7$), $\hat{\beta}_{\text{AI}} = -0.010$ ($t = -0.2$), $R^2 = 0.919$.

[13] This classification is consistent with Acemoglu et al. (2022), who note that AIOE weights occupations by the *technical feasibility* of AI performing their tasks, not by actual AI adoption rates. LMOE is built on the same feasibility logic. Technical feasibility correlates strongly with routine task content at the CZ level because both capture the degree to which occupations involve codifiable, structured activities.

AI, which captures the complementarity margin of the cognitive frontier, enters no moment.

The seven measures.

Figure 3 presents the central empirical result.

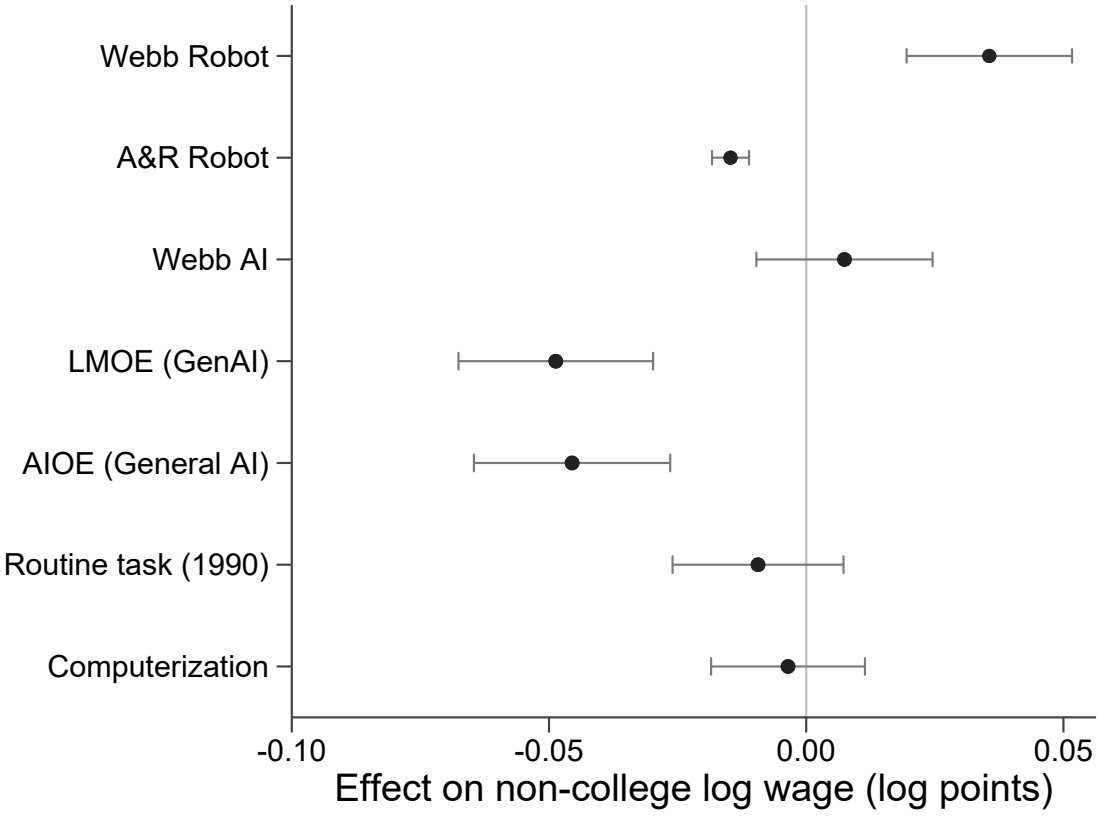


Figure 3: Effect of seven automation measures on $\Delta \ln w^{NC}$ (2005–2019), duplicating the non-college column of Table 1. 95% CIs shown. All regressions include baseline demographic controls, the baseline skill premium, and population weights.

Table 1 extends the analysis to all three outcomes and reports, for each measure, the signature its wage-coefficient pattern triggers under the wage-decomposition diagnostic and the empirical grouping it takes from its cross-CZ factor loading, with the measure's construction used as the fallback for the one measure the factor structure leaves unclassified.

Table 1: The Sign Matrix: Seven Measures × Three Outcomes (2005–2019)

Measure	College wage	Non-college wage	Skill premium	Sign	Signature	Grouping
Panel A: Robot/Physical						
Webb Robot	0.026^{***} (0.006)	0.036^{***} (0.008)	-0.010^{*} (0.006)	↑↑↓	Reinstatement	Physical
A&R Robot	-0.010^{***} (0.002)	-0.016^{***} (0.002)	0.005^{***} (0.002)	↓↓↑	Displacement	Physical
Panel B: AI/Cognitive						
Webb AI	0.018^{***} (0.006)	0.008 (0.009)	0.010^{**} (0.004)	↑—↑	Complementarity	Cognitive
LMOE†	-0.031^{***} (0.009)	-0.052^{***} (0.010)	0.021^{**} (0.009)	↓↓↑	Displacement	Physical†
AIOE†	-0.025^{***} (0.009)	-0.048^{***} (0.010)	0.023^{***} (0.008)	↓↓↑	Displacement	Physical†
Panel C: Other						
RTI (AD)	-0.011^{**} (0.004)	−0.008 (0.006)	−0.002 (0.004)	↓— —	No signature	Physical
Computeriz.	−0.001 (0.003)	0.001 (0.003)	−0.002 (0.003)	— — —	No signature	—
Controls + SP_0		Yes				
Observations		722				

Notes: Each cell reports the coefficient (robust SE) from a separate weighted least squares (WLS) regression of the column outcome on one standardized technology measure, with baseline demographic controls, the initial skill premium, and population weights. Sign columns: ↑ positive, ↓ negative, blank for null. "Signature" is the wage-decomposition mechanism the coefficient pattern triggers under Corollary 2. "Grouping" is the cross-CZ factor the measure loads on (A&R Robot grouped from its construction, Computerization excluded from the PCA). † marks measures built as AI exposure that load on the physical factor while supplying the cognitive displacement margin in the model (Section 3.2). RTI (AD) is the routine-task-intensity index of Autor and Dorn (2013). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

The sign matrix reveals three distinct signatures across the seven measures, each mapping to a prediction of the two-frontier model.

Webb Robot produces positive coefficients on *both* college and non-college wages, with the non-college effect larger ($\hat{\beta}^{NC} = 0.036$, $t = 4.47$, and $\hat{\beta}^{C} = 0.026$, $t = 4.08$). The resulting skill premium effect is negative ($\hat{\beta}^{SP} = -0.010$, $t = -1.69$). This is the reinstatement signature. CZs more exposed to robot-suitable occupations experienced wage gains for both

skill groups, but disproportionately for non-college workers, compressing the skill premium. The pattern is consistent with Proposition 4 under sufficient reinstatement, $\delta_R > \bar{\delta}_R$ in the notation of the structural model (Section 4), under which automation displaces routine tasks but creates new tasks that absorb non-college labor at higher marginal products.

The compression claim needs care. Because the three regressions share the sample, the regressors and the weights, the premium coefficient is exactly the difference of the two wage coefficients, so the ordering $\hat{\beta}^{NC} > \hat{\beta}^{C}$ and the negative premium coefficient are one restriction. That restriction is marginal. It is -0.010 with $t = -1.69$, it loses significance when the baseline premium control is dropped, it is unstable when LMOE replaces Webb AI in the horse race given the $r = -0.95$ collinearity, and no individual premium cell survives the Romano–Wolf stepdown reported below. What separates Webb Robot from every displacement measure is the level pattern, two positive wage coefficients at $p < 0.01$ where A&R Robot, LMOE and AIOE each carry two negative ones. The reinstatement label is read off that level pattern, and the wage levels are what supply the moments that discipline the structural estimation.

A&R Robot, LMOE, and AIOE all produce the opposite pattern, negative coefficients on both wage levels, with non-college workers hit harder (A&R: $\hat{\beta}^{NC} = -0.016$, $t = -8.77$, and $\hat{\beta}^{C} = -0.010$, $t = -5.55$). The skill premium rises. This is the displacement signature. The identifying variation in these measures loads on the supply-driven displacement margin, either industry-level robot deployment (A&R) or routine-task intensity (LMOE/AIOE), with little of the offsetting reinstatement that Webb Robot captures. The wage-level signs of LMOE and AIOE match A&R Robot and not Webb AI, as their empirical grouping with physical-factor measures (Section 3.2) predicts.

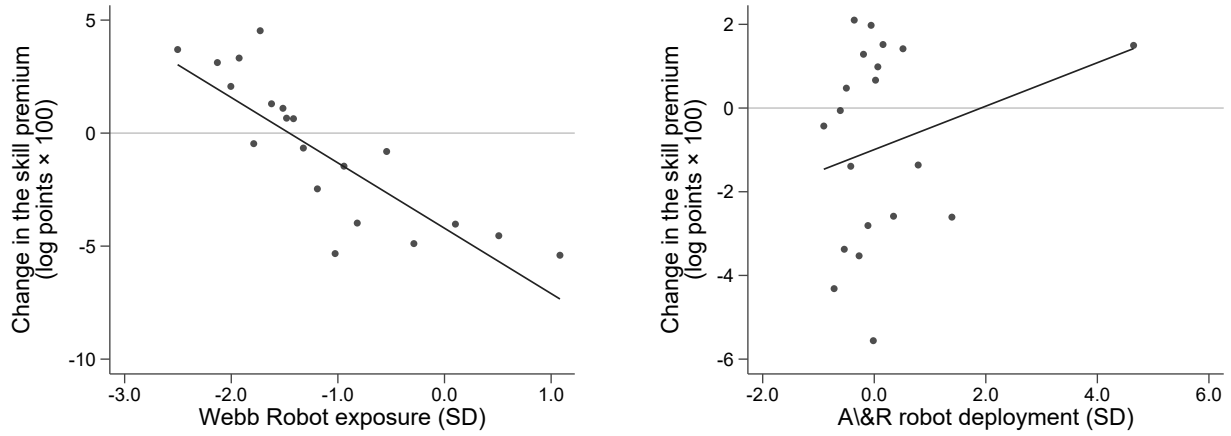
Webb AI produces a positive college wage effect ($\hat{\beta}^{C} = 0.018$, $t = 3.03$) and a non-college wage effect that cannot be distinguished from zero ($\hat{\beta}^{NC} = 0.008$, $t = 0.96$). Non-rejection is not evidence for a zero, and the interval on the non-college coefficient reaches above the college coefficient, so the data bound the asymmetry loosely. The skill premium rises. The

point-estimate pattern is the complementarity one. AI exposure is associated with higher college wages and with no detectable non-college change, which is consistent with a separate AI-complementarity channel outside the structural model, in which cognitive-frontier automation operates through skill-biased technical change instead of task displacement at the CZ level.

Five of the seven measures trigger one of the three signatures, and RTI and Computerization remain unclassified by the wage diagnostic. Only Webb Robot produces the reinstatement pattern. A&R Robot, LMOE, and AIOE produce the displacement pattern. Webb AI produces the complementarity pattern. RTI carries a negative college-wage coefficient with null non-college and premium effects, a pattern that matches none of the three signatures, so under Corollary 2 it triggers no signature, and Computerization is null across all outcomes. With twenty-one coefficients in the matrix, a few marginal stars would arise by chance. The coefficients that define the signatures are wage-level effects at $p < 0.01$, and the largest of them ($|t| > 4$ for the Webb Robot and A&R Robot wage levels and the LMOE non-college effect) survive a Bonferroni correction across all twenty-one cells. The Webb AI cells sit below that threshold, and the cognitive-frontier interpretation is supported alongside them by the AI-IV wage-level results of Section 3.3. A Romano–Wolf stepdown correction across all twenty-one cells (1,000 replications, Online Appendix E.7) controls the familywise error rate while exploiting the dependence across cells, and in these data it is stricter than the $|t| > 4$ shorthand above for the marginal cells. The non-college wage effects survive the stepdown, the college-wage cells are weaker under it (only A&R Robot clears it, and Webb Robot's college cell falls to $p = 0.062$), and no individual skill-premium cell survives. The pooled index (not shown, see Section 3.4) is attenuated toward zero, as opposite-signed components cancel. A scalar model that treats all automation measures as estimates of the same economic object cannot generate this sign structure. Measures of one frontier can still differ when they load on distinct margins.

Figure 4 illustrates the contrast between Webb Robot and A&R Robot. The two mea-

sures both claim to capture robot automation, yet they move the skill premium in opposite directions, and they do the same on the wage levels behind it. The model resolves this. Webb Robot captures occupation-level task content, picking up CZs where reinstatement accompanies automation. A&R Robot captures industry-level deployment, which loads far more heavily on the displacement margin. Both are "correct" measures of robot automation. They identify different economic objects.



(a) Webb Robot

(b) A&R Robot

Figure 4: Webb Robot vs. A&R Robot on $\Delta\mathcal{P}$ (2005–2019). Binscatter with 20 bins, residualized on the baseline controls, in log points $\times$ 100.

Remark 2 (Why Webb AI Enters No Structural Moment). The reduced-form coefficient of Webb AI exposure on the skill premium is positive, one log point per standard deviation of exposure, and modest relative to the robot coefficients. The structural calibration does not use Webb AI coefficients as target moments, because its cognitive moments come from LMOE (Section 3.2), and the measurement model (Online Appendix E.5) shows that about half of Webb AI's cross-sectional variance is idiosyncratic, so after partialling out geographic controls it retains limited identifying power (Online Appendix E.4).

3.3 Identification

The sign matrix establishes a systematic pattern across seven measures. This section brings exogenous variation to bear on it. Evidence on the displacement-dominated local short-run margin rests on the standard A&R deployment IV and its reduced form. The cross-measure variant is a sign diagnostic, the AI IV is weak and complementary, and reinstatement is not IV-identified, so the section is narrower than a full causal account of both frontiers. The theory predicts a specific sign reversal between OLS and IV estimates, because occupation-level vulnerability captures the net effect of displacement and reinstatement, while industry-level deployment recovers a deployment-weighted short-run margin that is less exposed to the non-local and long-run reinstatement margins. Neither instrument identifies a pure channel, and Online Appendix E.8 decomposes the wedges precisely. The OLS wage decomposition supplies a sign pattern consistent with reinstatement and moments that discipline the structural model.

Table 2 previews the argument. The displacement-only model predicts negative wage effects for both skill groups. The full model (with reinstatement) predicts positive effects. The OLS, which captures occupation-level vulnerability and thus the net effect of both channels, should match the full model's signs. The IV, which leans on industry-level deployment and is therefore less exposed to the reinstatement margins, should match the displacement-only model's signs.

Table 2: Displacement-Only vs. Full Model: Predicted Signs and OLS/IV Estimates

	Model (displacement only)	**Model (full)**	**OLS (Webb)**	**IV (A&R)**
Robot $\to \Delta w^{NC}$	$-$	$+$	$+0.036^{***}$	-0.020
Robot $\to \Delta w^{C}$	$-$	$+$	$+0.026^{***}$	-0.010
Robot $\to \Delta \mathcal{P}$	$+$	$-$	-0.010^{*}	$+0.010$
Channel	Displacement	Full extended model	Vulnerability (net)	Deployment (short-run margin)

Notes: Columns 1–2: predicted signs from the two-frontier model under displacement only ($\delta_R = 0$) and under the full extended model of Section 4, which adds reinstatement, common concentration, differential sorting and spatial pass-through together. Reinstatement alone leaves the robot premium response at $+0.57$, and the sign turns negative only once differential mobility enters (Table 14). Column 3: OLS estimates of Webb Robot from Table 1. Column 4: two-stage least squares (2SLS) estimates of A&R Robot deployment instrumented by EU5 Bartik (Online Appendix E.9, Table OE.15, Panel A). All specifications include baseline demographic controls and SP_0. $N = 722$. Stars report conventional inference, see Section 3.1 for exposure-robust inference. *** $p < 0.01$, * $p < 0.10$.

The data match every predicted sign. The OLS columns match the full model: Webb Robot raises both wages, compresses the premium. The IV columns match the displacement-only model: A&R deployment lowers both wages, widens the premium. The sign reversal between columns 3 and 4, the same underlying technology measured at different margins, is the sign pattern predicted by, and consistent with, the reinstatement mechanism.

The OLS estimates in Section 3.2 exploit cross-CZ variation in technology *exposure*, which is predetermined by baseline occupational composition. CZs with different occupational structures may also differ on unobservables, such as workforce adaptability, managerial quality, and local institutions, that independently affect wage dynamics. To address this, I pursue an IV strategy using the two European Bartik instruments standard in this literature.

Instruments.

The robot IV follows Acemoglu and Restrepo (2020): European (EU5) robot adoption by industry from 1993–2014, interacted with 1990 US industry employment shares at the CZ level. This isolates variation in US robot deployment driven by global technology supply shocks and not by local labor market conditions. The AI IV comes from the Babina et al. (2024) replication materials (Babina et al., 2025), a Bartik-style instrument interacting European AI-adoption shifts with local industry composition, the AI analogue of the EU5 robot Bartik. Both instruments are standardized to mean zero and unit standard deviation.

The robot IV is available for all 722 CZs. The AI IV is available for a subset of 304 CZs with nonmissing AI adoption data, of which 240 also have the change in AI worker share (ΔAI adoption). I report results for all available samples, noting where the IV subsample differs.

First Stage.

Instrumenting A&R robot *deployment* with the EU5 Bartik yields a strong first stage. Write $\hat{\pi}^{\text{FS}}$ for the first-stage coefficient, an object unrelated to the structural frontier effects π_P and π_C of Section 2.3. It is 0.729 ($t = 6.28$), with a Kleibergen-Paap rk Wald F of 39.4 (equal to the Sanderson-Windmeijer F in this just-identified single-endogenous-variable case) and Cragg-Donald $F = 1{,}526$, well above any conventional threshold. The gap between the robust and non-robust statistics reflects the population weighting and heteroskedasticity.

Instrumenting *Webb Robot* (occupation-level feasibility) with the same Bartik is mechanically harder because the IV captures industry-level deployment while Webb captures occupation-level vulnerability. Unconditionally, the correlation is near zero ($r = -0.02$, $F = 0.55$). Once I condition on female share, nonwhite share, mean age, and manufacturing share (but *not* college share), the first stage becomes strong, with $\hat{\pi}^{\text{FS}} = -0.132$ ($t = -5.5$, $F = 30.7$). The sign is negative because CZs with higher predicted robot deployment have lower Webb Robot exposure, conditional on industry composition. Adding college share absorbs most of the identifying variation, since college share is the dominant predictor of Webb Robot ($\hat{\pi}^{\text{FS}}_{\text{college}} = -10.3$, $t = -19.9$), reducing the first-stage KP F to 7.63, short of any

conventional strength threshold. I therefore report the specification *without* college share as preferred (F = 30.7) and the specification *with* college share as a robustness check (Online Appendix E.9, Table OE.15, Panel C).

The Babina IV predicts AI adoption strongly without controls ($\hat{\pi}^{\text{FS}} = 0.656$, $F = 12.7$). However, adding the full control vector including college share absorbs most of the variation ($F = 0.68$), because AI adoption and college share are highly correlated ($r = 0.63$). I therefore use a preferred specification with female share, nonwhite share, mean age, and manufacturing share but excluding college share, yielding $\hat{\pi}^{\text{FS}} = 0.459$ ($t = 1.77$, KP $F = 3.14$). The Babina IV also predicts Webb AI exposure with controls ($\hat{\pi}^{\text{FS}} = 0.341$, $F = 22.6$), but this cross-measure first stage should be interpreted cautiously since Babina is industry-based and Webb is occupation-based.

Reduced Form.

Table 3 reports the reduced-form effect of both IVs directly on ΔSP. The reduced form is immune to weak-instrument concerns.

Table 3: Reduced Form: Skill Premium on European Bartik IVs

	Robot IV Only			Both IVs		
	(1)	(2)	(3)	(4)	(5)	(6)
Robot IV (EU5)	0.007	0.012	0.012	0.006***	0.006***	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
AI IV (EU Bartik)				0.023***	0.031***	0.031***
				(0.004)	(0.004)	(0.004)
Baseline SP		−0.287***	−0.287***		−0.405***	−0.405***
		(0.056)	(0.064)		(0.074)	(0.060)
Webb AI	Yes	Yes	Yes	—	—	—
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
SE type	Robust	Robust	State	Robust	Robust	State
R^2	0.203	0.254	0.254	0.259	0.360	0.360
N	722	722	722	304	304	304

Notes: Dependent variable: $\Delta\mathcal{P}$ (2005–2019). All regressions include female share, nonwhite share, mean age, and manufacturing share, and *exclude* college share. Cols. (1)–(3): A&R robot IV with Webb AI exposure as control. Cols. (4)–(6): both IVs entered directly. WLS with population weights. Robust SEs in cols. (1)–(2) and (4)–(5), state-clustered in cols. (3) and (6). Stars report conventional inference, see Section 3.1 for exposure-robust inference, which was run on cols. (1)–(3) only.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Both IVs enter with the *same* positive sign on ΔSP in reduced form. CZs with greater exogenous exposure to European robot adoption experienced larger increases in the skill premium (columns 1–3), and CZs with greater predicted AI adoption from the EU AI Bartik also experienced increases (columns 4–6). The robot IV coefficient is stable at +0.006 to +0.012 across specifications. The AI IV coefficient ranges from +0.023 to +0.031.

The *same-sign* reduced-form result is consistent with the two-frontier model. The robot IV predicts greater SP because the short-run *displacement* margin it weights most heavily hits non-college workers harder, the defining signature of physical-frontier exposure. It does not raise the college wage directly. The AI IV predicts greater SP because AI adoption raises cognitive labor demand disproportionately for college workers. Both channels widen the skill premium, but through different mechanisms on different worker groups.

Two-stage least squares estimates.

I pursue three IV strategies of increasing ambition.

The first strategy instruments the robot side and treats AI as exogenous. I report two variants of the robot IV strategy, which differ in the endogenous variable.

Variant (a): A&R deployment instrumented by EU deployment. Using the A&R robot IV to instrument A&R US deployment ($N = 722$), with Webb AI as an exogenous control and full baseline demographic controls plus SP_0. The first-stage KP $F = 39.4$. Because the model is just identified with one instrument, this equals the Montiel Olea and Pflueger (2013) effective F, whose critical value for 5 per cent worst-case bias is 37.4 in this design, so the first stage clears the strictest of their thresholds. Stock-Yogo critical values are not used, since they are tabulated for the Cragg-Donald statistic under homoskedasticity and do not apply to a robust F. The 2SLS yields:

$$\hat{\beta}^{\text{NC}} = -0.020\ (0.003), \qquad \hat{\beta}^{\text{C}} = -0.010\ (0.002), \qquad \hat{\beta}^{\text{SP}} = +0.010\ (0.003) \tag{18}$$

Robot deployment lowers wages for both groups but hits non-college workers twice as hard,

widening the skill premium. The Anderson-Rubin test rejects the null for all three outcomes ($p < 0.001$, $p = 0.0005$, and $p < 0.001$, reported in Online Appendix E.9, Table OE.15, Panel A). Standard errors in parentheses are heteroskedasticity-robust. The three signs are the content of this exercise.

Variant (b): Webb feasibility instrumented by EU deployment, a reduced-form robustness exercise. The same IV can be pointed at Webb Robot exposure ($N = 722$), with Webb AI as an exogenous control. Because college share is the dominant predictor of Webb Robot and absorbs most of the first-stage variation when included, this specification uses demographics + manufacturing share + SP_0 but *excludes* college share. The first stage is strong in the statistical sense ($\hat{\pi}^{\text{FS}} = -0.132$, $t = -5.5$, $F = 30.7$), and the 2SLS yields:

$$\hat{\beta}^{\text{SP, IV}} = -0.093 \ (0.015), \qquad \hat{\beta}^{\text{AI, OLS}} = +0.011 \ (0.008) \tag{19}$$

This variant does not carry identification weight. The instrument's economic channel is deployment, which is the endogenous variable in Variant (a) and not here, so the -0.093 is mechanically the positive reduced form of the instrument on the premium ($+0.012$, Table 3) rescaled by an incidental negative first-stage correlation between deployment and measured feasibility. What the exercise establishes is narrower and still useful, namely that the compression sign on Webb Robot survives when the variation comes from European deployment shocks instead of cross-CZ occupation composition. The estimate is stable across European adoption windows (1993–2014, 2000–2014, 1993–2007), and state clustering leaves it at $\hat{\beta} = -0.093$ (SE $= 0.016$). The magnitude, a first-stage-dependent Wald rescaling, is not comparable to the OLS coefficient, and the formal reconciliation of the two variants is in Online Appendix E.8. Identification of the displacement channel rests on Variant (a) and the reduced forms in Table 3.

Adding college share to the controls weakens the first stage (KP $F = 7.63$) but amplifies the 2SLS magnitude to $\hat{\beta}^{\text{SP}} = -0.131$ (SE $= 0.050$), with an AR confidence set of

$[-0.327, -0.074]$ that excludes zero (Online Appendix E.9, Table OE.15, Panel C).

The second strategy instruments AI adoption and treats robot exposure as exogenous. Using the Babina AI IV on the AI subsample ($N = 240$) with Webb Robot treated as exogenous:

$$\hat{\beta}^{\text{AI, IV}} = +0.037^{*} \quad (\text{SE} = 0.020), \qquad \hat{\beta}^{\text{Robot, OLS}} = -0.024^{*} \quad (\text{SE} = 0.013) \tag{20}$$

The point estimates keep opposite signs under instrumentation. The IV estimate on AI adoption is positive, and its state-clustered standard error is 0.012. The cluster-robust Kleibergen-Paap first-stage F with the preferred demographic controls is 3.14, far short of any threshold that would license Wald inference. I therefore report Anderson-Rubin (AR) confidence sets that are robust to weak instruments. The AR test rejects the null of $\beta^{\text{AI}} = 0$ at $p = 0.0009$, and the Stock-Wright LM statistic rejects at $p = 0.003$ (Online Appendix E.9, Table OE.15, Panel B).

At the wage level the point estimates are +0.052 on college wages and +0.015 on non-college wages, and the Anderson-Rubin test rejects a zero coefficient for the first ($p = 0.001$) and not for the second ($p = 0.209$). Rejecting zero is weaker than signing the coefficient. Inverting the same test over candidate values leaves an unbounded confidence set that includes negative values, because the first stage does not reject at conventional levels ($\chi^2(1) = 3.25$, $p = 0.072$) and the Anderson-Rubin statistic converges to that first-stage value as the candidate grows without bound (Online Appendix E.9). The pattern is the complementarity signature, read as a sign check, and subject to the baseline-imbalance caveat in Section 3.3.

The third strategy instruments both exposures. The most demanding specification instruments both robot exposure and AI adoption simultaneously using both Bartik IVs (just-identified, $N = 240$). With controls and baseline SP:

$$\hat{\beta}^{\text{Robot, IV}} = -0.092^{***} \quad (\text{SE} = 0.016), \qquad \hat{\beta}^{\text{AI, IV}} = -0.029 \quad (\text{SE} = 0.022) \tag{21}$$

The robot coefficient remains large and carries a small conventional standard error. The AI adoption coefficient is negative, a sign flip relative to Strategy 2, and imprecisely estimated. This reflects the weak-instrument problem inherent in just-identified systems with two endogenous variables, since the Cragg-Donald F is 2.64, below the Stock and Yogo (2005) 25% maximal-size value of 3.63, and the cluster-robust Kleibergen-Paap F is 2.05. The Anderson-Rubin test, which is robust to weak instruments, rejects the joint coefficient null at $p < 0.001$, even though the conventional 2SLS point estimates remain unreliable.

I therefore interpret the dual-IV specification as a diagnostic and not a preferred estimate. The reduced form (Table 3, columns 4–6) is the more reliable summary of the dual-IV evidence. The full dual-IV system, with its weak-identification diagnostics, is reported in the notes to Table OE.15.

Exclusion Restriction.

For both IVs, the key identifying assumption is that European technology adoption affects US CZ-level wages only through the corresponding automation channel.

The robot IV satisfies a clean placebo, since its population-weighted correlation with the baseline skill premium is near zero ($r(\text{Robot IV}, \text{SP}_{2005}) = -0.06$), and the balance regression of SP_{2005} on the instrument is insignificant ($t = -0.8$, population-weighted, robust). CZs with greater predicted robot exposure had nearly identical baseline skill premia, consistent with the instrument being orthogonal to pre-existing labor market structure.[14]

The AI IV is more problematic, because it is strongly correlated with the baseline skill premium ($r(\text{AI IV}, \text{SP}_{2005}) = 0.40$, population-weighted, balance regression $t = 5.77$). CZs with greater predicted AI adoption had substantially higher baseline skill premia, reflecting the correlation between AI adoption and human capital concentration. Controlling for baseline SP does not eliminate the AI IV reduced-form coefficient, which rises from $\hat{\beta} = +0.023$ in column 4 of Table 3 to +0.031 in column 5. I therefore treat the AI IV results as complementary evidence and not as definitive identification, and rely on the reduced form with

[14]This is the same balance test reported in Acemoglu and Restrepo (2020, Table 3). Reported t-statistics throughout this subsection come from the population-weighted balance regression, not from a raw correlation.

baseline SP as the preferred specification.

Aggregation Under IV.

The masking theorem is a mixed-sign result, so its IV prediction is about the pooled coefficient and not about fit. Both instruments enter the reduced form with the same positive sign (Table 3), because the robot IV places less weight on the nonlocal and long-run reinstatement margins instead of recovering the net-of-reinstatement effect, so a nonnegatively weighted index of the two cannot return a null. It does not. On the 304-zone common-instrument sample, which is larger than the 240-zone simultaneous dual-IV system because it does not require the AI adoption measure, the pooled standardized index is significant on conventional inference with demographic controls, college share and baseline SP ($\hat{\beta} = +0.011$, $t = 4.6$, Table OE.28, Panel B). With college share added, the two reduced-form coefficients of Table 3, columns (5)–(6), move to +0.008 and +0.005, and separated and pooled R^2 are both 0.497 there, an equality that is a property of these data and this control vector and not a prediction of the theorem. The instrument diagnostics are in Online Appendix E.9 and the table itself sits with the auxiliary specifications. The opposite-signed objects in the IV exercise are the OLS and IV estimates of the same robot technology (Table 4), not the two instruments.

Table 4 collects the key estimates across estimation strategies. Each row is a distinct estimation exercise, and columns report the robot and AI coefficients on ΔSP from the preferred specification.

Three patterns emerge. First, the deployment IV carries the displacement sign under exogenous variation. The standard IV (Panel B, row 1) shows that instrumented A&R deployment *raises* the skill premium (+0.010), because robot installation lowers both wages, with non-college workers hit harder. The cross-measure exercise (Panel B, row 2) has the same sign as the Webb Robot compression (−0.093), and the sign difference between the two rows reflects the different endogenous variables (deployment vs. feasibility), not contradictory evidence. With college share included as a control, the cross-measure coefficient moves to

Table 4: Robot and AI Effects on $\Delta\mathcal{P}$: Estimation Strategies and Sign Diagnostics

Strategy	Robot $\hat{\beta}$	AI $\hat{\beta}$	N	1st-stage F	AR p
Panel A: Baseline					
OLS (Webb × Webb)	-0.037^{***}	$+0.020^{***}$	722	—	—
	(0.003)	(0.004)			
Panel B: Robot IV strategies					
Standard IV (A&R ← Bartik)	+0.010	+0.011	722	39.4	< 0.001
	(0.003)	(0.006)			
Cross-measure IV (Webb ← Bartik)	−0.093	+0.011	722	30.7	< 0.001
	(0.015)	(0.008)			
Panel C: AI IV strategy					
AI IV (Babina ← Bartik)	-0.024^{*}	$+0.037^{*}$	240	3.1	0.001
	(0.013)	(0.020)			

Notes: Dependent variable: $\Delta\mathcal{P}$ (2005–2019). Each row is a separate regression. OLS: Webb Robot + Webb AI, with female share, nonwhite share, mean age, and manufacturing share but excluding college share (comparable to the cross-measure IV). Standard IV: A&R robot deployment instrumented by EU5 Bartik, Webb AI exogenous, full controls including college share. Cross-measure IV: Webb Robot instrumented by the same EU5 Bartik, Webb AI exogenous, controls exclude college share. AI IV: Δ AI adoption instrumented by Babina EU Bartik, Webb Robot exogenous, controls exclude college share. All specifications include female share, nonwhite share, mean age, manufacturing share, and baseline SP, with college share added only in the specifications noted above. 1st-stage F: Kleibergen-Paap cluster-robust Wald rk F-statistic. AR p: Anderson-Rubin weak-instrument-robust p-value for the instrumented variable. WLS with population weights. Cluster-robust SEs. Stars report conventional inference, see Section 3.1 for exposure-robust inference.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

-0.131 (AR $p < 0.001$, Online Appendix E.9, Panel C).

Second, the AI coefficient is positive in every preferred specification summarized in Table 4. In the dual-IV diagnostic of equation (21) it is negative and imprecise (-0.029, SE $= 0.022$). The AI IV and the OLS estimate different endogenous objects, AI adoption versus Webb AI exposure, so their magnitudes are not directly comparable. What carries over is the sign of the point estimate (AR $p = 0.001$).

Third, the robot and AI coefficients carry opposite signs in three of the four strategies. The exception is the standard IV, which instruments robot *deployment* instead of occupation-level feasibility. There the robot coefficient turns positive ($+0.010$), as the displacement-only reading predicts when the estimand places less weight on nonlocal and longer-run reinstatement margins, so it shares the sign of the AI coefficient. The two-frontier structure is not an artifact of any single estimation exercise.

The employment margin.

The baseline estimates use CZ-level mean log wages by education group. The wage effects could instead reflect compositional shifts, changes in *who* lives and works in each CZ, and not price effects on incumbent workers. A second concern is that the skill premium effects may be driven by selective migration of college workers instead of technology-induced changes in labor demand. Three tests address these concerns.

The first test uses composition-adjusted wages. I construct Mincer residual wages by regressing individual log wages on age, age squared, gender, race, and ethnicity, separately by year, then averaging the residuals by commuting zone and education group. The regression carries no commuting-zone effect, because a zone effect fitted inside each year would absorb the entire zone-year wage level and leave the two group means as mirror images of one another, and the wage columns below are read as levels. Education is excluded from the residualization, since the college gap in the residuals is the object of interest. The residual wage change purges the mechanical effect of shifts in age, gender, race, and ethnicity composition within each education group. It does not hold worker identities fixed, so any change in

occupation, industry, unobserved skill or selection into employment stays inside the residual, and reading the adjusted change as a price effect requires that whatever composition remains be uncorrelated with exposure, which is maintained here and not tested.

Table 5 reports the results. For robot exposure the premium effect moves from -0.015 ($t = -4.16$) raw to -0.011 ($t = -3.40$) adjusted, and for AI from $+0.011$ ($t = 4.02$) to $+0.008$ ($t = 3.30$). The wage levels behind the robot premium survive the adjustment. Non-college wages rise by $+0.033$ ($t = 4.31$) raw and $+0.026$ ($t = 3.71$) adjusted, college wages by $+0.018$ ($t = 2.73$) and $+0.014$ ($t = 1.92$). Signs are unchanged throughout, and the attenuation of 26 to 27 percent on the two premium coefficients is the part of the baseline result that the adjustment removes. Because the Mincer coefficients are re-estimated in every year, that gap mixes a change in who is counted with a change in what their characteristics pay, so it is not the composition channel alone.

Table 5: Raw vs. Composition-Adjusted Wages

	College Wages		Non-College Wages		Skill Premium	
	Raw	Adjusted	Raw	Adjusted	Raw	Adjusted
	(1)	(2)	(3)	(4)	(5)	(6)
Robot exposure	0.018***	0.014*	0.033***	0.026***	−0.015***	−0.011***
	(0.006)	(0.008)	(0.008)	(0.007)	(0.004)	(0.003)
AI exposure	0.010*	0.005	−0.002	−0.003	0.011***	0.008***
	(0.005)	(0.006)	(0.006)	(0.005)	(0.003)	(0.003)
Controls + SP_0	Yes	Yes	Yes	Yes	Yes	Yes
N	722	722	722	722	722	722

Notes: "Raw" columns use CZ-level mean log wages. "Adjusted" columns use Mincer residual wages, constructed by regressing individual log wages on age, age squared, gender, race, and ethnicity separately by year, with neither education nor a commuting-zone effect among the regressors, then averaging residuals by commuting zone and education group. All regressions include baseline demographic controls and the baseline skill premium. $\Delta\mathcal{P}$ is computed from three-year endpoint windows (2005–2007 average to 2017–2019 average), and the baseline skill premium is the 2005–2007 window value. WLS with population weights. CZ-clustered robust SEs. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The second test addresses college outmigration. If robot-exposed CZs experience selective outmigration of college workers, the resulting decline in local college labor supply could mechanically compress the skill premium, mimicking the pattern attributed to reinstatement. The channel has to start somewhere, and the first place to look is whether exposure moves

the college share of employment at all. Regressing that share change on the two exposures, with the same controls, weights and clustering the wage regressions use, gives −0.0024 (SE 0.0020) for robot exposure and −0.0007 (SE 0.0019) for AI. The robot point estimate carries the sign the migration story needs, with the college share falling where robots arrive, but it is a tenth of a standard deviation of the share change and is not distinguishable from zero. This step conditions on nothing that exposure could have caused.

Conditioning on the share change directly points the same way, and is the weaker exercise. Adding Δs_c^C to the non-college wage regression *increases* the robot coefficient from +0.0326 ($t = 4.25$) to +0.0334 ($t = 4.45$), a 2.6% increase (percentages here are computed from the unrounded coefficients). For college wages, the coefficient rises from +0.0190 ($t = 3.00$) to +0.0201 ($t = 3.31$, +5.9%). For the skill premium, the coefficient moves from −0.0136 ($t = -3.59$) to −0.0133 ($t = -3.45$), an attenuation of 2.0 percent, and remains significant. The share change is measured over the outcome window and is itself a candidate response to exposure, so this comparison is not a decomposition into direct and migration-mediated parts. What it shows is that the wage pattern does not depend on holding the college share fixed.

The third test separates demand from supply. The joint behavior of wages and employment quantities distinguishes demand from supply explanations. A positive demand shock raises wages and employment. A positive supply shock raises employment but lowers wages. Robot exposure produces a distinctive combination, with non-college wages *rising* (+0.033, $t = 4.92$) while non-college employment *falls* (−0.028, $t = -3.19$). College workers show the same pattern, with wages up (+0.023, $t = 4.02$) and employment down (−0.032, $t = -3.16$).[15]

This combination, wages up and employment down, is inconsistent with a single positive demand shift or a single positive supply shift, though a negative labor-supply shift would

[15]These wage coefficients come from the joint wage-employment sample. The reference single-measure estimates, quoted throughout, are +0.036 (non-college) and +0.026 (college) from the sign matrix (Table 1). The mediation analysis below restricts to 470 commuting zones with task-displacement data and reports +0.039. The qualitative pattern is identical across samples.

produce it as well. The joint pattern is consistent with displacement and reinstatement operating simultaneously. Employment lies outside the model, which maps displacement into wages through marginal products and moves employment only through migration, in the same direction as the local wage. A joint test over the four wage moments and the four education-specific employment moments rejects that co-movement ($J = 47.0$, five degrees of freedom, $p = 5.6 \times 10^{-9}$, Online Appendix D.7). The rejection is a scope result, and the estimate $\hat{\delta}_R = 1.233$ comes from the wage moments alone.

For AI, the employment effects are positive but statistically insignificant (+0.010–0.011, $t \approx 1.3$–1.5), consistent with Webb AI capturing a complementarity margin operating through college-skill returns.

The Reinstatement Channel.

The sign matrix establishes *that* the reference Webb Robot and Webb AI measures produce opposite skill-premium coefficients. The extended model attributes the pattern to asymmetric reinstatement, in which physical automation displaces manual tasks but simultaneously creates new tasks that absorb non-college workers, while cognitive automation displaces cognitive tasks without a comparable reinstatement channel. Direct evidence on this mechanism comes from task-content data.

I use the displacement–reinstatement decomposition of Acemoglu and Restrepo (2022), aggregated to CZs via industry Bartik shares. For each industry j, the decomposition separates employment change into a *reinstatement* component (new task creation) and a *displacement* component (task automation). I construct CZ-level measures as $\text{TD}_c = \sum_j s_{jc,0} \times \text{td}_j$, where $s_{jc,0}$ is the baseline industry employment share and td_j is the national industry-level component from the 2000–2016 decomposition. Table 6 reports the results.

The Net Reinstatement column carries the comparison. Webb Robot is the *only* measure for which the coefficient is positive, at $\hat{\beta} = +0.127$ ($t = 0.94$). The point estimate is not distinguishable from zero, and because the outcome is standardized the slope carries no information about the level of the net balance, so the interval admits gradients of either sign

Table 6: Task Displacement Decomposition: Automation Measures and the Reinstatement–Displacement Balance

	Reinstatement		Displacement		Net Reinstatement	
	$\hat{\beta}$	t	$\hat{\beta}$	t	$\hat{\beta}$	t
Panel A: Physical-frontier measures						
Webb Robot	−0.415	−3.43***	−0.316	−2.54**	+0.127	+0.94
Routine Task Index	+0.728	+7.53***	+0.648	+6.44***	−0.451	−4.27***
Computerization	+0.949	+22.75***	+0.874	+16.30***	−0.661	−9.42***
Panel B: Cognitive-frontier measures						
Webb AI	−0.312	−2.39**	−0.188	−1.51	−0.024	−0.22
LMOE	+0.975	+6.33***	+0.835	+5.82***	−0.524	−3.64***
AIOE	+0.825	+6.02***	+0.719	+5.51***	−0.475	−3.48***

Notes: Each cell is a separate CZ-level regression of the Bartik task-displacement component (standardized) on the indicated automation measure (standardized). Industry Bartik approach: national industry-level decomposition from Acemoglu and Restrepo (2022), 2000–2016, aggregated via baseline industry employment shares. Reinstatement = new tasks created (comp). Displacement = tasks automated (subs). Net = Reinstatement − Displacement, constructed at the industry level before standardization, and each component is standardized separately. All regressions include baseline demographic controls (college share, manufacturing share, female share, nonwhite share, mean age), baseline skill premium, population weights, CZ-clustered robust standard errors. $N = 470$. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

and the exercise does not detect one. The wage gain runs through the marginal product of the reinstated tasks and not their count, because new tasks enter near the frontier's boundary productivity, so robot exposure raises the marginal product of labor and both skill groups' wages even when net task creation merely offsets destruction, and a flat net-balance gradient and reinstatement-dominant wages are consistent. Both separately standardized components covary negatively with robot exposure (displacement $\hat{\beta} = -0.316$, $t = -2.54$, creation $\hat{\beta} = -0.415$, $t = -3.43$). Because each component is standardized on its own scale, their difference is not the net coefficient, so the comparison is one of direction and not of magnitude.[16]

Every other measure tells a different story. For Routine (-0.451, $t = -4.27$), Computerization (-0.661, $t = -9.42$), LMOE (-0.524, $t = -3.64$), and AIOE (-0.475, $t = -3.48$),

[16]The negative signs on both reinstatement and displacement for Robot reflect the Bartik construction, since CZs with high robot exposure tend to have industry compositions where *national* reinstatement was lower. The net gradient is indistinguishable from zero, with displacement and creation moving in the same direction.

the net-balance gradient is strongly negative, so the measure falls where exposure is higher. Webb AI shows a near-zero net (-0.024, $t = -0.22$) for a different reason. Its reinstatement coefficient is negative and significant (-0.312, $t = -2.39$) while its displacement coefficient is negative but insignificant (-0.188, $t = -1.51$), consistent with its cognitive-frontier nature operating through price and composition channels instead of task reallocation.

The pattern survives a horse-race specification. When Robot and AI exposure enter jointly, both carry negative reinstatement coefficients, predicting lower measured reinstatement (Robot: $\hat{\beta} = -0.353$, $t = -2.57$, AI: $\hat{\beta} = -0.269$, $t = -2.20$, $R^2 = 0.193$), but neither predicts net reinstatement (Robot: $\hat{\beta} = +0.136$, $t = 0.99$, AI: $\hat{\beta} = -0.041$, $t = -0.38$, $R^2 = 0.135$). The physical-frontier point estimate is positive and the cognitive-frontier point estimate is negative, but neither net coefficient is statistically distinguishable from zero in the joint specification.

A pre-period placebo supports temporal validity. Regressing 1987–2000 task displacement on 2005–2019 automation exposure, which should be pre-determined, produces null coefficients for Robot on all three components (reinstatement: $t = -1.30$, displacement: $t = -0.87$, net: $t = +0.27$).

Finally, the task-displacement components are added to the baseline wage regression on the 470-CZ Bartik sample. As with the college-share comparison above, the components are candidate responses to exposure measured over the outcome window, so the exercise is descriptive. Net reinstatement enters the non-college wage equation insignificantly ($\hat{\beta}_{\text{net}} = -0.0032$, $t = -0.87$), and the Robot coefficient is virtually unchanged ($+0.0395$ without vs. $+0.0399$ with the mediator). The skill premium result is the same (-0.0148 without, -0.0147 with). The measure may also be too noisy to move anything in a sample of 470 CZs.

3.4 Masking

The sign matrix and identification results establish that the reference Webb Robot and Webb AI measures produce opposite skill-premium effects through different channels. Proposition 2 predicts a direct consequence, that aggregating the two into a single automation index will destroy the signal. This section tests that prediction.

Pooled against decomposed.

Table 7 tests the masking prediction directly. Column (1) regresses $\Delta\mathcal{P}$ on a single aggregate automation index, the average of Webb Robot and Webb AI, both standardized. Column (2) enters Robot alone. Column (3) enters AI alone. Column (4) enters both.

Table 7: The Masking Test: Pooled vs. Decomposed Automation Effects on the Skill Premium

	No Controls				With Controls			
	(1) Pooled	(2) Robot	(3) AI	(4) Both	(5) Pooled	(6) Robot	(7) AI	(8) Both
Agg. automation ($\bar{A}$)	-0.019^{***}				$+0.009$			
	(0.006)				(0.006)			
Robot exposure		-0.029^{***}		-0.031^{***}		-0.009		-0.014^{**}
		(0.003)		(0.003)		(0.006)		(0.007)
AI exposure			$+0.019^{***}$	$+0.023^{***}$			$+0.012^{**}$	$+0.015^{***}$
			(0.005)	(0.005)			(0.005)	(0.005)
R^2	0.041	0.220	0.061	0.308	0.364	0.365	0.379	0.390
N	722	722	722	722	722	722	722	722

Notes: Dependent variable: $\Delta\mathcal{P}_c$ (2005–2019), computed from single-year endpoints (2005 to 2019). Aggregate automation is the average of standardized Webb Robot and Webb AI, $\bar{A}_c = 0.5R_c + 0.5A_c$. Baseline controls include college share, manufacturing share, female share, nonwhite share, and mean age, without the baseline skill premium. Population weights, robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The masking prediction holds precisely. Without controls, the pooled index is significantly negative ($\hat{\beta} = -0.019$, $t = -3.02$) but explains only 4.1% of the variation, compared with 30.8% for the two-index model. The scalar index has a negative net association because the Robot coefficient is larger in absolute value than the AI coefficient, so the net correlation is negative, but attenuated. Decomposition reveals the hidden structure, with $\hat{\beta}_R = -0.031$ ($t = -10.63$) and $\hat{\beta}_{AI} = +0.023$ ($t = 5.03$), and the two-index R^2 more than seven times

higher.

With controls, the pooled coefficient *flips sign* to +0.009 ($t = 1.46$, insignificant), as the college-share gradient, which is positively correlated with both AI exposure and the skill premium, absorbs the variation that previously identified the Robot effect. Meanwhile, the two-index model retains individually significant, opposite-signed coefficients ($\hat{\beta}_R = -0.014$, $t = -2.18$, and $\hat{\beta}_{AI} = +0.015$, $t = 2.77$) and achieves $R^2 = 0.390$, a 7.1% improvement over the scalar model's $R^2 = 0.364$.

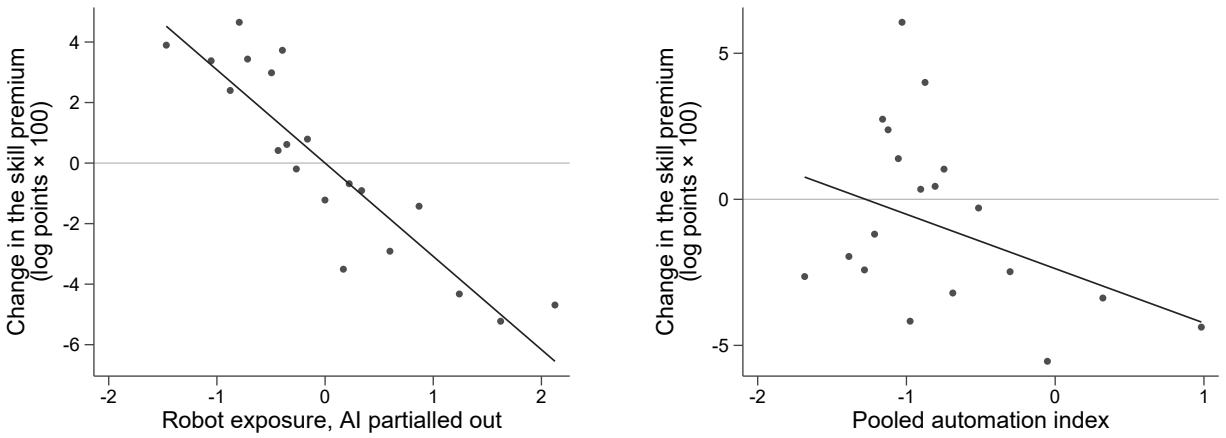
The result is not specific to Webb Robot. For the routine occupation share (1990), the signed PCA rotation produces cancellation because its loadings orient the two measures oppositely, which is a property of the rotation and not of the nonnegative pooling of orientation-aligned channels that the masking theorem describes. The two measures are close to orthogonal in the population-weighted correlation matrix ($r = -0.011$), so the leading component is the *difference* and not the sum, and since both carry nearly equal positive associations (+0.018 and +0.019 in the no-controls specification), differencing them returns approximately zero ($\hat{\beta} = -0.001$, $t = -0.20$, $R^2 \approx 0$). For Computer adoption (Babina), masking is less severe because both channels push the skill premium in the same direction, but the scalar index still inflates the apparent effect (+0.031 vs. +0.019 when AI is separated). The pooled-versus-decomposed results for all three measures are in Online Appendix E.13, and the PCA-index specification is in Table OE.28, Panel A.

Figure 5 illustrates the contrast visually.

The Masking Curve.

The pooled-vs.-decomposed test compares two points, equal weights ($\omega = 0.5$) and full decomposition. This subsection traces the coefficient continuously as the index's frontier loading varies. For any weight $\omega \in [0, 1]$, define a pooled automation index

$$M_c(\omega) \equiv \omega R_c + (1 - \omega) A_c, \tag{22}$$



(a) Separated: robot, partialling out AI

(b) Pooled: total automation

Figure 5: **Masking Seen Directly.** Population-weighted binscatters of the change in the skill premium, in log points × 100, each horizontal axis in units of its own regressor. *Panel A* partials AI exposure out of both axes and plots the robot residual. *Panel B* plots the raw outcome against the pooled index $\bar{A}_c = 0.5R_c + 0.5A_c$. The line in each panel is the weighted least-squares fit on the 722 commuting zones, with slopes of -3.1 in Panel A and -1.9 in Panel B, in log points per unit of that panel's regressor. The residualized robot exposure of Panel A has standard deviation 0.96 and the pooled index of Panel B has 0.83, so per standard deviation the slopes are -3.0 and -1.5. Twenty weighted quantile bins were requested, and Panel B realizes nineteen because two cut-points coincide.

where R_c is standardized robot exposure (Webb Robot) and A_c is standardized AI exposure (Webb AI). For each ω on a fine grid, estimate

$$\Delta\mathcal{P}_c = \alpha(\omega) + \beta(\omega)M_c(\omega) + X_c'\gamma + u_c, \tag{23}$$

using population weights and the baseline controls X_c. The mapping $\omega \mapsto \beta(\omega)$ is the *masking curve*.

Table 8 reports coefficients at representative loadings.

Table 8: Masking Curve: Skill Premium Coefficient at Selected Robot Loadings

Robot Loading ω	$\hat{\beta}(\omega)$	Robust S.E.	t-stat
0.00 (AI only)	+0.0102	(0.0045)	2.27
0.25	+0.0097	(0.0053)	1.83
0.50 (Equal weight)	+0.0057	(0.0059)	0.96
0.75	−0.0030	(0.0062)	−0.47
1.00 (Robot only)	−0.0101	(0.0060)	−1.69
Zero crossing ω^*		≈ 0.67	
Controls		Yes	
Observations		722	

Notes: Each row reports the coefficient from equation (23) using $M_c(\omega) = \omega R_c + (1-\omega)A_c$. Population weights and baseline demographic controls plus the baseline skill premium included. Robust standard errors in parentheses. The zero crossing $\omega_{\text{curve}} = 0.672$ is interpolated on the finer grid plotted in Figure 6.

The masking curve traces a monotone decline from +0.010 (pure AI, $\omega = 0$) to −0.010 (pure Robot, $\omega = 1$). The coefficient crosses zero at $\omega_{\text{curve}} = 0.672$, close to the structural threshold $\hat{\omega}^*_{\text{struct}} = 0.658$ of Section 4. The curve's cognitive arm is Webb AI while the structural estimation targets the LMOE moments, so the estimator does not target the curve, but both are built from the same wage data, so the comparison is a non-targeted empirical one and not statistically independent evidence. The one-standard-error band, the range of loadings where $|\hat{\beta}(\omega)| < \text{SE}$, equivalently $|t| < 1$, spans $\omega \in [0.50, 0.84]$, covering the natural equal-weight combination and robot-tilted loadings through roughly $\omega = 0.84$. Outside that band the coefficient exceeds its standard error, though it need not reach conventional significance. At $\omega = 0.25$ the t-statistic is 1.83 and at the pure-robot end $\omega = 1$ it is −1.69,

neither significant at the 5 percent level, while the pure-AI end $\omega = 0$ reaches $t = 2.27$ and is significant at the 5 percent level. The equal-weight point sits at +0.006 here versus +0.009 in Table 7 because the curve conditions on the baseline skill premium in addition to the baseline controls. A pooled scalar index whose effective loading lies near the masking threshold will produce a coefficient that cannot be distinguished from zero. That covers the equal-weight index a researcher would reach for by default, though not every weighting, since the pure-AI index is significant.

Figure 6 plots the fitted curve for Webb Robot against the two quantities the theory and the literature supply, the structural threshold ω^*_{struct} of Section 4.7 and the equal-weight loading an aggregate automation index implicitly adopts. The curve crosses zero close to the structural threshold, and the equal-weight index sits inside the range where the coefficient cannot be distinguished from zero.

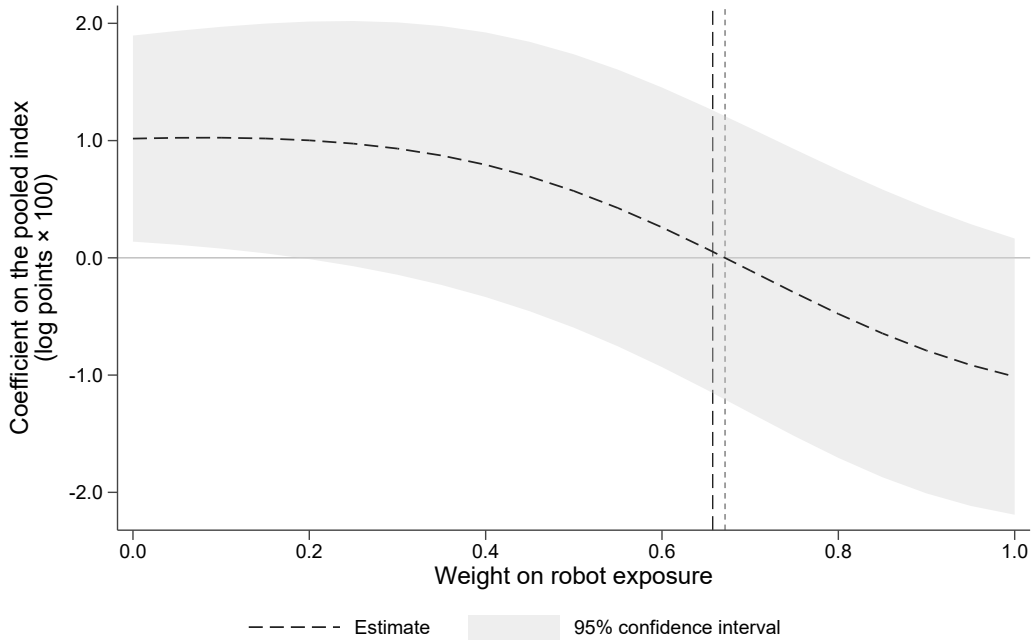
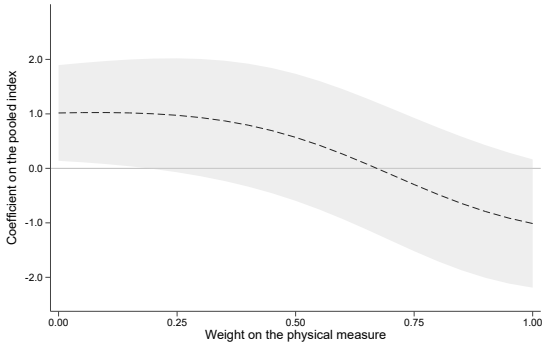


Figure 6: **The Masking Curve.** Estimated coefficient $\hat{\beta}(\omega)$ from regressing $\Delta\mathcal{P}_c$ on the pooled index $M_c(\omega) = \omega R_c + (1-\omega)A_c$, with baseline demographic controls, the baseline skill premium, and population weights. $N = 722$. The shaded band is the pointwise 95 percent confidence interval. Here ω is the researcher-chosen index loading. The longer-dashed vertical line is the structural threshold $\omega^*_{\text{struct}} = 0.658$ of Section 4.7, and the shorter-dashed line is the empirical crossing $\omega_{\text{curve}} = 0.672$.

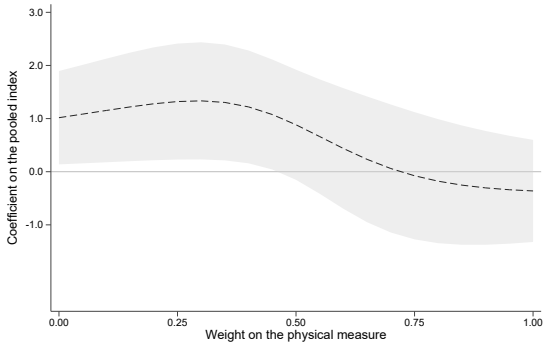
Figure 7 repeats the exercise for all three physical measures, with pointwise ± 1 standard-error and 95% confidence bands.

A falsification.

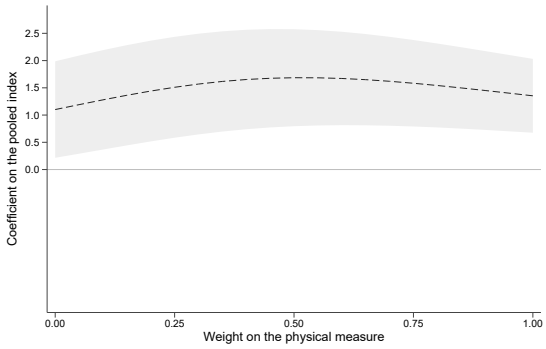
Masking is a mixed-sign phenomenon, which makes it refutable. It must appear where the two channels carry opposite signs, and with nonnegative, orientation-aligned index weights it must be absent where they do not. The other two physical measures supply the test (Figure 7). The routine occupation share is the weaker case. The distinct AD routine-task-intensity measure in Table 1 has an uninformative wage signature, and the 1990 routine occupation share used here is evaluated through this masking-curve exercise. Its zero crossing shifts out to $\tilde{\omega}_0 \approx 0.76$, against 0.73 for Webb Robot on the same specification, because a weaker negative effect needs more physical loading to offset the AI channel. Computer



(a) Webb Robot



(b) Routine occupation share (1990)



(c) Computer adoption (Babina)

Figure 7: **Masking curve.** Estimated coefficient $\beta(\omega)$ from regressing $\Delta\mathcal{P}_c$ on the pooled index $M_c(\omega) = \omega Z_c^P + (1-\omega)A_c$ for each of three physical-frontier measures Z_c^P, with baseline demographic controls only, without the baseline skill premium, so the crossings sit to the right of the $\omega^* \approx 0.67$ of Figure 6. Panel (a): Webb Robot, zero crossing at $\tilde{\omega}_0 \approx 0.73$. Panel (b): routine occupation share, 1990, $\tilde{\omega}_0 \approx 0.76$. Panel (c): computer adoption, the share of workers in computer occupations from the Babina et al. (2024) data ($N = 677$), with no crossing because both channels are positive. Shaded regions indicate where $|\hat{\beta}| <$ SE.

adoption supplies the second case. The adoption measure loads +0.857 on the physical factor against Webb Robot's −0.951 (Table OE.8), so it is an inversely oriented measure of the same frontier, and its skill premium effect is positive.[17] Pooling it with AI exposure therefore combines two channels that push the same way, so the theorem predicts no null. None appears. The estimated coefficient stays positive and significant at every loading, peaks near equal weights, and never crosses zero. At $\omega = 0.5$ the coefficient is +0.017 ($t = 3.19$ on the control set of Figure 7, which omits the baseline skill premium). Aggregation amplifies the signal instead of destroying it.

The instrumental variables estimates carry the same discipline. Both instruments enter the reduced form with the same positive sign (Table 3), so no nonnegatively weighted index of the two can return a null. None does. The pooled standardized index is +0.011 ($t = 4.6$, Table OE.28, Panel B). The pathology appears where the theory says it should and is absent where the theory says it should not. That is what separates masking from ordinary attenuation, which would shrink the coefficient wherever measures are pooled.

Cross-frontier interaction.

The masking theorem is a statement about linear aggregation. A separate question is whether the two frontiers *interact*, so that the joint effect of dual exposure departs from the sum of the two linear channels. I test this by adding the interaction $Z_c^P \times A_c$ to the horse-race specification, writing Z_c^P for physical-frontier exposure to keep it apart from the skill premium $\mathcal{P}$:

$$\Delta\mathcal{P}_c = \alpha + \beta_P Z_c^P + \beta_A A_c + \xi\,(Z_c^P \times A_c) + X_c'\gamma + u_c. \tag{24}$$

A negative $\hat{\xi}$ means that in CZs exposed to both frontiers the premium change falls *below* the linear sum of the two channels. With $\hat{\beta}_P < 0$ and $\hat{\beta}_A > 0$, it deepens the physical-frontier

[17] This is the share of workers in computer occupations from the Babina et al. (2024) data ($N = 677$). It is a different measure from the Acemoglu and Restrepo (2020) computerization-intensity measure that enters the sign matrix, which is null.

compression where cognitive exposure is high, so the frontiers combine sub-additively instead of adding up. Table A.1 reports $\hat{\xi}$ across measures and specifications.

For Webb Robot and Computer adoption, the interaction is consistently negative across all specifications (Table A.1, Panel A), indicating that in dual-exposed CZs the premium change falls below the linear sum of the two channels. The interaction survives every control set for Computer adoption, without controls and in the kitchen-sink specification alike. Within college-share terciles (Panel C), the Webb Robot interaction is negative in all three groups, with the largest magnitude in the high-college tercile ($\hat{\xi} = -0.038$, $t = -2.89$), where both frontiers have the most scope to operate simultaneously.

The routine occupation share interaction is *positive* throughout, the reverse of Robot and Computer adoption. This is interpretable. Routine tasks are displaced by both physical automation and early software, so the Routine×AI term captures complementarity between two partially overlapping constructs, not cross-frontier substitution. The positive sign strengthens in the high-college tercile (+0.023), consistent with sorting amplifying within-frontier overlap for cognitively intensive occupations.

Figure A.1 plots the within-tercile relationship between $\Delta\mathcal{P}_c$ and each physical-frontier measure after partialling AI exposure out of both axes. It shows the physical-exposure gradient within each community type, not the interaction term itself.

The full battery of interaction robustness tests, including geography-plus-complementarity models, triple interactions with college share, and within-tercile residualized interactions, is reported in Online Appendix E.13. The headline result is that the cross-frontier interaction survives every robustness check for Computer adoption, holding under all four residualization methods in Panel C and in the kitchen-sink specification, and is consistently negative for Webb Robot, though statistically weaker once controls absorb the collinear variation. The routine occupation share interaction is positive in all full-sample specifications and in the mid- and high-college terciles, though negative and statistically insignificant in the low-college tercile, in line with within-frontier overlap driving the opposite sign for this measure.

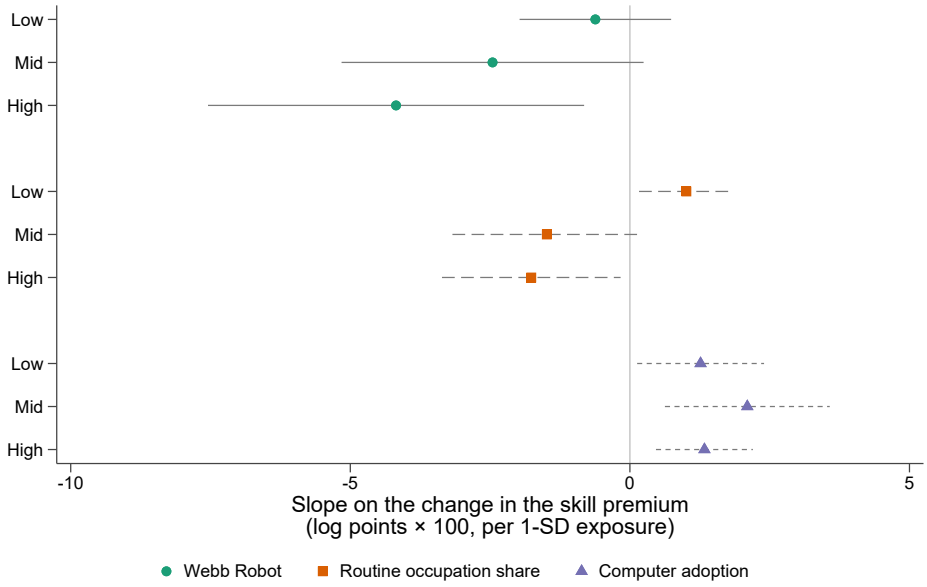


Figure 8: **Tercile-Specific Slopes.** Slope of the change in the skill premium on each physical-frontier measure, estimated separately within population-weighted terciles of the baseline college share, with baseline demographic controls and the baseline skill premium, one measure at a time. Horizontal bars are 95% confidence intervals. The three groups of three are Webb Robot, the routine occupation share, and Computer adoption. Webb Robot is negative in every tercile and steepens with college share, a pattern that does not survive the horse-race specification of Table A.3, where the middle tercile is the steepest. $N = 722$ for Webb Robot and the routine occupation share, and $N = 677$ for Computer adoption, which is missing for the commuting zones absent from the Babina et al. (2024) data.

3.5 Temporal Dynamics

The structural model of Section 4 is estimated entirely on 2005–2019 cross-sectional variation. This subsection describes what happened to the two channels afterward. The model's attenuation channel implies that robot-driven compression weakens as the flow of new task creation slows, because the compression is generated by that flow. What the data show after 2019 is stronger than that, because the physical channel changes sign, and the change is large enough that equality with the pre-2019 coefficient is rejected. The post-2019 evidence is descriptive and not an out-of-sample test of the model, because any specification whose effect scales with the size of the diffusion shock implies the same attenuation, and because the sign change asks for a movement in the reinstatement rate that nothing here measures directly.

Within the baseline period.

The first test decomposes the baseline window into two sub-periods (2005–2010 and 2010–2019) using a stacked long-difference design. Each CZ appears twice in a balanced panel ($N = 1{,}444$), with period fixed effects absorbing aggregate time trends and standard errors clustered on the commuting zone. Interaction terms Robot $\times$ Late and AI $\times$ Late test whether the effect of each technology changed across windows.

Table A.2 reports five specifications: separate sub-period OLS (columns 1–2), pooled with period fixed effects (column 3), the two exposures interacted with the period while the controls stay pooled (column 4), and every regressor interacted with the period (column 5).

Both channels are present in both sub-periods and neither changes detectably. Robot exposure compresses the premium by 0.011 ($t = -4.11$) in 2005–2010 and by 0.009 ($t = -2.61$) in 2010–2019. Cognitive exposure widens it by 0.005 ($t = 2.64$) and then by 0.009 ($t = 3.53$). Stacking the two windows with every regressor interacted with the period (column 5 of Table A.2) makes the interaction on each exposure the difference between the two windows' own coefficients. The robot difference is +0.002 (SE 0.005) and the cognitive difference is +0.004 (SE 0.003), so equality of each channel across the two halves of the

baseline window is not rejected.

Holding the demographic controls to a common coefficient across periods reverses that answer, and the reversal is a property of the control set and not of the exposures. Under pooled controls, column (4) of the same table, the robot difference is −0.011 and the cognitive difference is +0.006. Robot exposure correlates −0.865 with baseline college share, so a control coefficient forced to be equal in the two windows absorbs part of the cross-period level difference and loads it onto the interaction. The pattern repeats with LMOE as the cognitive regressor, where the pooled-control difference is +0.017 and the period-specific one is −0.010. Letting the controls move is what isolates the exposures, and on that comparison the structure is stable inside the baseline window.

The robot coefficient in the LMOE specification shows no detectable period shift either, so the two specifications agree that the structure inside the baseline window is stable.

Across periods.

The second approach extends the sample into the generative AI era. Table 9 reports the skill premium regression for two non-overlapping subperiods and the full 2005–2024 window using ACS data through 2024.

The 2019–2024 window breaks the physical channel. The robot coefficient moves from −0.015 in 2005–2019 to +0.010 in 2019–2024. Setting one window's significance verdict beside the other's is not a test that the coefficient changed, so the two windows are stacked with every regressor interacted with the period, which makes the interaction on robot exposure the difference between the two windows' own coefficients and gives that difference a standard error. It is +0.025 (SE 0.005, $t = 4.72$), clustered on the commuting zone, which appears in both windows. Equality is rejected. The cognitive channel moves the other way, from +0.011 to +0.002, a difference of −0.009 (SE 0.004, $t = -2.51$), and equality is rejected there as well. The two channels then behave differently over the full window. The robot coefficient offsets almost exactly across the subperiods, −0.015 and then +0.010, and leaves +0.000 over 2005–2024. The cognitive coefficient does not offset. It is +0.011 before 2019,

Table 9: Technology Effects on $\Delta\mathcal{P}$ Across Periods

	Webb AI specification			LMOE specification		
	(1) 2005–19	(2) 2019–24	(3) 2005–24	(4) 2005–19	(5) 2019–24	(6) 2005–24
Robot (Webb)	-0.015^{***}	$+0.010^{***}$	$+0.000$	$+0.009$	$+0.008$	$+0.011$
	(0.004)	(0.004)	(0.004)	(0.008)	(0.006)	(0.008)
AI / LMOE	$+0.011^{***}$	$+0.002$	$+0.009^{***}$	$+0.032^{***}$	-0.005	$+0.011$
	(0.003)	(0.002)	(0.003)	(0.011)	(0.009)	(0.012)
Baseline SP	-0.394^{***}	-0.301^{***}	-0.444^{***}	-0.407^{***}	-0.287^{***}	-0.442^{***}
	(0.039)	(0.036)	(0.042)	(0.038)	(0.039)	(0.043)
N	722	722	722	722	722	722
R^2	0.594	0.294	0.669	0.585	0.293	0.659

Notes: Dependent variable: $\Delta\mathcal{P}$ for the indicated period. Columns (1)–(3): Webb Robot + Webb AI. Columns (4)–(6): Webb Robot + LMOE. All specifications include baseline (2005) demographic controls and period-start skill premium. Every $\Delta\mathcal{P}$ is computed from three-year endpoint windows, so 2005–2019 runs from the 2005–2007 average to the 2017–2019 average and 2019–2024 from 2017–2019 to 2022–2024, and the period-start skill premium is the window value. WLS with 2005 population weights. Robust SEs. All six columns hold the same 722 commuting zones, which is what makes the coefficients comparable across periods. The commuting-zone assignment uses the PUMA vintage each American Community Survey wave actually carries: 2000-vintage definitions for 2005–2011, 2010-vintage for 2012–2021 and 2020-vintage for 2022 onward (Online Appendix E.1).
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

+0.002 after, and +0.009 over the full window, so the premium-widening cognitive channel is a feature of the whole period and not of the diffusion window alone. The reversal is specific to the Webb AI horse race. With LMOE as the cognitive control the robot coefficient is +0.009 before 2019 and +0.008 after, neither distinguishable from zero, which is what the $r = -0.95$ collinearity between LMOE and Webb Robot documented in Section 3.1 implies for a specification that carries both, and over the full window neither coefficient in that specification is distinguishable from zero. Explained variation falls from $R^2 = 0.594$ in 2005–2019 to 0.294 in 2019–2024, the shorter of the two windows.

The wage levels locate the reversal. Over 2005–2019 robot exposure raises both, non-college wages by 0.033 ($t = 4.31$) and college wages by 0.018 ($t = 2.73$), which compresses the premium and is the reinstatement signature of Corollary 2. Over 2019–2024 the same measure lowers non-college wages by 0.014 ($t = -4.74$) while the college effect is −0.004 and imprecise, so the premium widens because the lower group loses ground and not because

the upper group gains it. That is the displacement signature, produced by the measure that produced the reinstatement signature in the earlier window. The quantity margin does not turn with the price margin. Robot exposure reduces employment for both groups in both windows, by 0.035 and 0.042 before 2019 and by 0.022 and 0.041 after, so the wage reversal happens while employment keeps moving the same way.

The pandemic window.

The post-2019 movement could reflect a discrete COVID-19 labor market shock instead of a change in the pace of new-task creation. Table 10 tests this by decomposing 2019–2024 into a COVID window (2019–2022) and a post-COVID recovery (2022–2024).

Table 10: COVID Sub-Period Decomposition

	Robot (Webb)				AI (Webb)	
Period	$\hat{\beta}^{C}$	$\hat{\beta}^{NC}$	$\hat{\beta}^{SP}$	t^{SP}	$\hat{\beta}^{SP}$	t^{SP}
2019–2022	+0.001	−0.000	+0.001	0.22	+0.009***	4.28
2022–2024	−0.005**	−0.012***	+0.007**	2.21	−0.003	−1.47
2019–2024	−0.004	−0.014***	+0.010***	2.73	+0.002	1.06
N		722				

Notes: Dependent variables: $\Delta \ln w^{C}$ (college), $\Delta \ln w^{NC}$ (non-college), and $\Delta \mathcal{P}$ (skill premium) for the indicated sub-period. All specifications include baseline (2005) demographic controls, period-start SP, and population weights. CZ-clustered SEs. Every sub-period holds the same $N = 722$ commuting zones, and all three windows use three-year endpoint averages. $\hat{\beta}^{SP} = \hat{\beta}^{C} - \hat{\beta}^{NC}$ before rounding, so printed values may differ in the final digit.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

The premium effect of robot exposure sits in the recovery and not in the pandemic. In 2019–2022 it is +0.001 ($t = 0.22$). In 2022–2024 it reaches +0.007 ($t = 2.21$), carried by a non-college wage effect of −0.012, and over the whole 2019–2024 window it is +0.010 ($t = 2.73$). The difference between the two sub-windows is +0.006 (SE 0.005) and is not distinguishable from zero. What the sub-periods do show is that the compression is already gone by 2019–2022, before any of the positive coefficient appears.

The cognitive channel does move across the two sub-windows. It widens the premium by 0.009 ($t = 4.28$) during the pandemic window and sits at −0.003 ($t = -1.47$) in the

recovery, and the difference between the two is −0.012 (SE 0.003, $t = -3.56$). Remote work accelerating adoption in college-intensive occupations is one reading of the pandemic-window spike. The model does not predict it either way, because Webb AI carries the complementarity margin and not the displacement margin the cognitive frontier is estimated on.

Composition-adjusted wages carry the same pattern. The Mincer residuals of Section 3.3 give a robot premium coefficient of −0.011 ($t = -3.40$) over 2005–2019 and +0.008 ($t = 2.08$) over 2019–2024, against raw values of −0.015 and +0.010, and the cognitive coefficient moves from +0.008 ($t = 3.30$) to zero ($t = 0.11$). Both the compression before 2019 and the widening after it therefore survive adjustment for observed worker characteristics.

The reversal of the robot-compression channel does not reflect a fall in robot adoption. International Federation of Robotics data show that U.S. manufacturing robot density rose from roughly 100 robots per 10,000 employees in 2005 to about 200 in 2017 and continued climbing to roughly 295 by 2023 (International Federation of Robotics, 2024). The aggregate stock kept rising, so whatever changed operates through the composition of adoption and not its level. The wage-compressing reinstatement was concentrated in the high-suitability occupations that *Webb Robot* exposure captures, where adoption matured as firms exhausted high-return applications (automotive, electronics, metals), while investment reallocated toward AI, which Babina et al. (2024) show accelerating after 2018. In the model the compression channel is the flow of new-task creation and its wage incidence, not the robot stock, so a rising aggregate density is consistent with a compression channel that weakens. A change of sign asks for more than that, since within the model it requires the effective reinstatement rate to fall below the threshold of Definition 2, at which point displacement dominates and the same measure widens the premium it used to compress.

Within community type.

Because robot-intensive occupations concentrate in manufacturing, robot exposure correlates strongly with baseline college share ($r = -0.865$ in the full sample). The concern is

that the opposite-sign result reflects divergent trajectories of high- versus low-education communities and not automation per se. I address this through four complementary strategies: within-tercile estimation, formal interaction tests, continuous heterogeneity specifications, and flexible control functions. Throughout, I extend the analysis beyond Webb Robot to two additional physical-frontier measures, the routine occupation share (1990) and Babina et al. computer adoption, which have very different correlations with college share ($r = +0.490$ and $r = +0.811$, respectively).

I partition CZs into population-weighted terciles of baseline college share, so each tercile holds one-third of employment (low: $\leq 27.1\%$ with $N = 579$ CZs, mid: 27.1–33.1% with $N = 100$, high: $\geq 33.1\%$ with $N = 43$), and re-estimate the horse-race specification within each group. Table A.3 reports the results.

The robot coefficient is negative in every tercile under every specification, and nine out of nine cells carry the predicted sign. Without controls (Panel A), all three tercile-specific estimates are individually significant ($t = -3.40, -2.76, -2.30$). The mid-college tercile shows the *largest* point estimate (-0.058), not the smallest as the geography critique would predict. With baseline controls (Panel B), the low-college estimate attenuates to near zero (-0.002, $t = -0.26$), but this reflects absorption by the college share control within an already narrow range. The mid-college estimate remains significant (-0.033, SE $= 0.015$), while the high-college estimate remains negative but is not statistically significant (-0.024, SE $= 0.017$). Adding baseline skill premium (Panel C) preserves the pattern, with a high-college coefficient of -0.029 ($t = -2.07$), so the negative sign is present in the tercile where manufacturing communities are least represented. The AI coefficient is positive in all nine cells, significant in seven.

The within-tercile correlations show that the tercile partition removes substantial confounding. Within the low-college tercile the Robot–college correlation drops from -0.865 to -0.779, within mid-college to -0.461, and within high-college to -0.550. Yet the sign persists. The pattern replicates across physical-frontier measures. Computer adoption, which

correlates *positively* with college share ($r = +0.811$), shows positive skill-premium effects in all three terciles without controls ($+0.018^{***}$, $+0.012$, $+0.004$), showing that the sign is not mechanically tied to a negative college-share correlation.

A formal interaction specification regresses $\Delta\mathcal{P}$ on Robot, AI, tercile dummies, and Robot $\times$ tercile interactions. The F-test for equal robot slopes across terciles fails to reject equality, with $F(2, 715) = 1.87$ ($p = 0.155$) without controls, $F(2, 710) = 1.35$ ($p = 0.260$) with controls, and $F(2, 709) = 0.44$ ($p = 0.644$) with controls and baseline skill premium. The routine occupation share and Computer adoption give the same answer under all three control sets. The robot effect does not vary detectably with community education level.

Instead of imposing tercile boundaries, I interact each physical measure with demeaned college share, so the main effect is identified at the sample mean. For Webb Robot with controls, the main effect is -0.016 ($t = -2.43$) and the interaction is -0.018 ($t = -0.52$), so no attenuation in more educated communities is detectable at this sample size. Implied effects range from -0.014 at the 10th percentile of college share to -0.018 at the 90th, a narrow band. With baseline skill premium, the interaction remains insignificant ($\delta = -0.014$, $t = -0.45$).

As a further check, I progressively add college share^2, college$\times$manufacturing, a cubic polynomial, a median spline, tercile fixed effects, and a kitchen-sink specification combining all terms. Across seven specifications, the Webb Robot coefficient ranges from -0.014 to -0.015 without baseline SP and from -0.014 to -0.017 with baseline SP. The coefficient is stable across the control battery. Most specifications move it farther from zero (the largest amplifications are -7.0% and -15.7%), while the tercile-fixed-effects specification with baseline SP attenuates it modestly, from -0.0144 to -0.0136. Computer adoption is even more stable, at approximately $+0.010$ without the baseline premium and $+0.012$ to $+0.013$ with it. The opposite-sign pattern is robust to the paper's observed geography controls and heterogeneity diagnostics.

The temporal and geographic evidence supports four conclusions. First, inside the base-

line window the two channels are stable. Robot exposure compresses the premium in 2005–2010 and again in 2010–2019, cognitive exposure widens it in both, and once the controls are allowed to differ by period neither difference is distinguishable from zero. Second, the 2019 boundary is a break the data can carry. The robot coefficient moves by +0.025 (SE 0.005) and the cognitive coefficient by −0.009 (SE 0.004), and equality is rejected for both. The reversal runs through non-college wages, which robot exposure raises before 2019 and lowers after. Third, the COVID decomposition separates the two channels without dating a common break. The cognitive channel changes across the pandemic and recovery windows (-0.012, $p < 0.001$), while the robot difference between those same windows is +0.006 and not distinguishable from zero, so the sub-periods do not locate the physical break more finely than the 2019 boundary does. Fourth, the opposite-sign pattern survives within every college-share tercile, is robust to flexible control functions (the coefficient moving by at most about 7% without the baseline premium control and 16% with it, almost always as amplification), and cannot be rejected as homogeneous across community types ($p > 0.15$ in all interaction F-tests). The full geography robustness battery, including binscatter visualizations, continuous interactions, and kitchen-sink specifications for all three physical-frontier measures, is reported in Online Appendix E.10. Whether the generative AI era will produce its own two-frontier pattern, with large language models (LLMs) as the new displacement technology, remains an open question that the current data cannot resolve.

4 A Calibrated Structural Model

The reduced-form evidence matches Proposition 1. The model-targeted Webb Robot and LMOE moments produce opposite-signed skill-premium effects. The *direction*, however, reverses the baseline model's prediction. Under pure displacement the model predicts that robots raise the skill premium and AI lowers it, and Section 3 shows the reverse. Table 11 records the mismatch. This section develops the three extensions the data demand, calibrates

the model, and quantifies the masking.

Table 11: Baseline Model Predictions vs. Empirical Signs

	Baseline (Pure Displ.)		Empirical (Section 3)	
	Robot	AI	Robot	LMOE
$\Delta \ln W_S$ (non-college)	−	−	+	−
$\Delta \ln W_H$ (college)	−	−	+	−
$\Delta \mathcal{P}$ (skill premium)	+	−	−	+
Signs opposite?		✓		✓
Direction match?		× (reversed)		

Notes: Columns 1–2 report predicted signs from the baseline model under pure displacement ($\delta_R = \delta_A = 0$). Columns 3–4 report empirical signs from the reduced-form regressions in Section 3, with LMOE carrying the cognitive displacement margin (Section 3.2). The baseline model correctly predicts sign opposition between frontiers (Proposition 1) but gets the direction wrong. Four mechanisms reverse it: asymmetric reinstatement, common cognitive concentration, differential sorting, and spatial pass-through (Table 14).

Four mechanisms reverse the direction. Differential sorting temporarily breaks sign opposition along the way, and spatial equilibrium restores it at the calibrated parameters. Asymmetric reinstatement requires that robot automation create enough new tasks for both wages to rise, and the wage pattern that primarily disciplines it is that robot exposure raises both wages. Common cognitive concentration raises the weight of tasks near the AI frontier for both skill groups and so governs how far both cognitive wages fall. Differential sorting requires that non-college workers in the cognitive sector be concentrated near that frontier, disciplined by $|\Delta \ln W_S| > |\Delta \ln W_H|$ under AI exposure. Spatial pass-through requires that non-college workers be less mobile, disciplined by $\Delta \ln W_S > \Delta \ln W_H$ under robot exposure at the CZ level. Table 14 adds them in six stages, because pass-through separates into a common and a differential part.

4.1 Frontier-Specific Reinstatement

Following Acemoglu and Restrepo (2019), frontier advances create new labor tasks. The main text adopts the normalization $\gamma_j = 1$, under which the effective rate δ_j coincides with the literal task-creation rate. Online Appendix A.8 separates the two objects. The stock of reinstated physical tasks is $\delta_R \cdot I_R$, so an advance dI_R adds $\delta_R\, dI_R$ new tasks at the frontier boundary, each with productivity $g_{LP}(I_R)$. Similarly, the cognitive continuum carries a reinstated stock $\delta_A \cdot I_A$ at productivity $g_{LC}(I_A)$.

In the nested CES framework, reinstated tasks form a distinct production module Q_{Nj} instead of being pooled with surviving tasks. The task content of the reinstated module (equation 7) is:

$$\mathcal{A}_{Nj} = (\delta_j \cdot I_j)^{\frac{1}{\sigma-1}} \cdot g_{Lj}(I_j) \tag{25}$$

This enters the sector production function through the across-module CES aggregator with elasticity ρ. The separation of reinstated from surviving tasks captures the economic reality that new tasks and complementary roles, such as prompt engineering, robot maintenance, and AI training, are qualitatively different from the tasks they sit alongside. These are often created within existing occupation categories and not as new occupation titles, which is what Online Appendix E.12 finds in the data.

A frontier advance works through two economic channels, displacement and reinstatement, which enter sector output as three chain-rule components:

$$\begin{aligned}\frac{dY_j}{dI_j} = & \underbrace{Y_j^{1/\rho} \cdot Q_{Kj}^{-1/\rho} \cdot \frac{dQ_{Kj}}{dI_j}}_{\text{displacement (expands capital module)},\ >0} \\ & + \underbrace{Y_j^{1/\rho} \cdot Q_{Sj}^{-1/\rho} \cdot \frac{dQ_{Sj}}{dI_j}}_{\text{contraction of surviving labor},\ <0} \\ & + \underbrace{Y_j^{1/\rho} \cdot Q_{Nj}^{-1/\rho} \cdot \frac{dQ_{Nj}}{dI_j}}_{\text{reinstatement (expands new module)},\ \geq 0}\end{aligned} \tag{26}$$

The surviving-labor term carries the full derivative of the module. Because $Q_{Sj} = \mathcal{A}_{Sj} L_j^S$, it has two parts,

$$\frac{dQ_{Sj}}{dI_j} = \mathcal{A}_{Sj} \frac{dL_j^S}{dI_j} + L_j^S \frac{d\mathcal{A}_{Sj}}{dI_j}, \tag{27}$$

the labor withdrawn from the module and the change in its task content as the least productive surviving tasks are automated away. Both must be kept, because writing the term as $\mathcal{A}_{Sj}\, dL_j^S/dI_j$ alone drops the task-content margin. The signs annotated in (26) are the technology response at fixed total sector labor. Writing s_{Nj} for the reinstated module's labor share, $d\ln L_j^S/dI_j = d\ln L_j/dI_j - s_{Nj}(\rho - 1)(d\ln \mathcal{A}_{Nj}/dI_j - d\ln \mathcal{A}_{Sj}/dI_j)$, so a full derivative also carries the general-equilibrium term $d\ln L_j/dI_j$, which nothing here bounds. Proposition 4 therefore establishes the equilibrium sign of each sector's marginal-product response numerically. Displacement and reinstatement operate through *different modules*, since displacement expands the capital module and contracts the surviving-labor module, while reinstatement expands the new-task module. The net effect on wages depends on which channel dominates.

Whether a frontier advance raises or lowers the marginal product of labor turns on the reinstatement threshold $\bar{\delta}_j$ of Definition 2.

The effective reinstatement rates, $\hat{\delta}_R = 1.233$ calibrated and $\hat{\delta}_A = 0.256$ implied by it through the external task-rate ratio, carry a sharp asymmetry. At the reference calibration, effective physical reinstatement clears the displacement threshold $\bar{\delta}_R$, so a frontier advance raises the marginal product of labor. Cognitive reinstatement is finite, at about a fifth of the physical rate, and does not clear its own threshold, which is far higher. Cognitive displacement destroys tasks whose weights are raised by concentration near the frontier while cognitive reinstatement returns tasks carrying the baseline weights, so the cognitive output boundary sits at 4.14 against 2/3 in the physical sector, which has no sorting (Online Appendix A.10). AI displacement therefore feeds predominantly into reduced labor demand.

4.2 Concentration and Sorting in the Cognitive Sector

Two things happen to cognitive task weights near the AI frontier, and the model keeps them apart. Tasks close to the frontier are more valuable to both skill groups, and non-college workers are disproportionately concentrated among them. Parametrize the first with a common concentration $\nu \geq 0$ and the second with a differential sorting $\varphi \geq 0$:

$$a_H^C(z;\nu) = a_H \cdot g_{LC}(z) \cdot z^{-\nu}, \tag{28}$$

$$a_S^C(z;\nu,\varphi) = \frac{1}{\sqrt{\kappa}} \cdot g_{LC}(z) \cdot z^{-(\nu+\varphi)}. \tag{29}$$

The weights apply on the labor domain $z \in (I_A, 1]$, bounded away from zero for any interior frontier, so the divergence at the origin never enters the aggregates. Reinstated cognitive tasks carry the baseline weights a_H and $1/\sqrt{\kappa}$ with neither exponent, because a task created at the frontier is new work and inherits no position in the existing sorting order.

The two exponents do different work, and that is what makes them separable. Only φ survives in the comparative-advantage ratio, $a_S^C/a_H^C \propto z^{-\varphi}$, so within this block ν raises both skill weights near the frontier while φ tilts them against non-college workers.[18] The calibrated values are $\hat{\nu} = 1.137$ and $\hat{\varphi} = 0.411$ (Section 4.5). Positive sorting is consistent with the concentration of non-college workers in routine cognitive tasks adjacent to the AI frontier. Online Appendix A.9 documents the occupational evidence from ACS data and Autor et al. (2003), whose task model explains about 60 per cent of the relative demand shift favouring college labour over 1970–1998, of which changes within nominally identical occupations are close to half.

[18] Setting $\varphi = 0$ removes the differential tilt but leaves $z^{-\nu}$ in place. The flat baseline is recovered only at $\nu = \varphi = 0$.

4.3 Spatial Equilibrium

Local wages satisfy the reduced-form incidence relation of Moretti (2011) and Notowidigdo (2020), combining the demand condition (wage equals marginal product) with an upward-sloping inverse local labor supply curve whose amenity and moving-cost shifters are absorbed into CZ constants:

$$\Delta \ln W_k^c = \eta_k \, \Delta \ln \mathrm{MPL}_k^c, \qquad \eta_k \equiv \frac{\varepsilon_D}{\varepsilon_D + \varepsilon_k^{\mathrm{mig}}} \in (0, 1], \tag{30}$$

combining the labor-demand curve, with local demand elasticity ε_D, and the inverse local supply curve, with migration elasticity $\varepsilon_k^{\mathrm{mig}}$, into the equilibrium pass-through η_k. The derivation is in Online Appendix A.10, Lemma A.2.

Assumption 2 (Differential Mobility)**.** Non-college workers are less geographically mobile, $\varepsilon_S^{\mathrm{mig}} < \varepsilon_H^{\mathrm{mig}}$ (Bound and Holzer, 2000; Notowidigdo, 2020; Hornbeck and Moretti, 2024).

The maintained migration elasticities, $\varepsilon_S^{\mathrm{mig}} = 0.78$ and $\varepsilon_H^{\mathrm{mig}} = 3.12$, imply that non-college labor supply is about one-quarter as responsive to local wage shocks as college labor supply. Both are maintained and not estimated, because the literature supplies the ordering and not the levels (Online Appendix D.3.1), and every structural quantity below is conditional on them. This spatial channel is what turns a common national movement into a differential local one. When robot automation raises both marginal products, the labor inflow of college workers compresses their wage gains relative to less-mobile non-college workers. That compression alone does not order the two local responses. Equation (30) gives $\eta_k u_k$ with $u_k \equiv \Delta \ln \mathrm{MPL}_k$, and at the external elasticities $\eta_S = 0.794$ and $\eta_H = 0.490$, so $\Delta \ln W_S > \Delta \ln W_H > 0$ at the CZ level requires $u_S/u_H > \eta_H/\eta_S = 0.618$, a stronger condition than $u_S, u_H > 0$, since with $u_S = 1$ and $u_H = 2$ both marginal products rise and the ordering reverses. Proposition 4 carries the equivalent bracket condition $\Delta_R + m\, u_H^R > 0$ and verifies it on the calibrated configuration.

4.4 Extended Comparative Statics

Proposition 4 (Direction Under Asymmetric Reinstatement)**.** *Maintain $\rho > 1$ and Assumption 2, and let $\delta_R > \bar{\delta}_R$, $\delta_A < \bar{\delta}_A$, $\nu > 0$ and $\varphi > 0$. Parts (a) and (b) are verified numerically on the benchmark and the sensitivity domain of Online Appendix A.10. Under these conditions:*

(a) Robot automation raises both wages, and at the CZ level $\Delta \ln W_S > \Delta \ln W_H > 0$, so the local skill premium falls.

(b) AI automation lowers both wages, and at the CZ level $|\Delta \ln W_S| > |\Delta \ln W_H|$, so the local skill premium rises.

(c) The two local premium responses carry opposite signs, $\operatorname{sgn}(\Delta \mathcal{P}_R^{CZ}) = -\operatorname{sgn}(\Delta \mathcal{P}_A^{CZ})$, in 2,000 of 2,000 conditional pairs-bootstrap re-estimates.

Proof. Each part splits into a level condition on the two marginal products and an ordering condition on a single scalar. The split is an exact identity. Writing u_k^j for the national marginal-product response and $\Delta_j = u_S^j - u_H^j$, the local premium response is $\Delta \mathcal{P}_j^{\mathrm{CZ}} = \eta_H u_H^j - \eta_S u_S^j = -\eta_S[\Delta_j + m\, u_H^j]$ with $m = 1 - \eta_H/\eta_S$ (Lemma A.2). The leading minus sign is the paper's convention $\mathcal{P} = \ln W_H - \ln W_S$, under which the bracket is wage compression, and the premium moves against it.

Level. Each reinstatement rate is compared with its own threshold. In the physical sector, which carries no sorting, the output boundary is $\bar{\delta}_R^Q = 1/(1 + (\sigma - 1)\beta_P) = 2/3$ exactly. In the cognitive sector, displacement destroys a task whose weights are raised by concentration while reinstatement returns one carrying the baseline weights, and the cognitive envelope gives the boundary in closed form. At fixed atom share it is strictly increasing in ν and in φ, and equals 4.14 at the estimates. The corresponding marginal-product boundaries are not available in closed form, so (a) and (b) are verified numerically at the benchmark and across the sensitivity domain of Online Appendix A.10.

Ordering. Given the level signs, $\Delta_R > 0$ and $\Delta_A < 0$ are each sufficient for the stated ordering, and neither is necessary, because the mobility term can carry the sign on its own. Both orderings are verified numerically over the same domain.

(c) Opposite signs hold in 2,000 of 2,000 conditional pairs-bootstrap re-estimates. The identity, the envelope and both closed forms are proved in Online Appendix A.10, which also states which claims are analytic and which are numerical. □

Proposition 5 (Cross-Frontier Spillovers)**.** *Robot automation shifts p_P/p_C and triggers cross-sector reallocation. Under asymmetric reinstatement the two reallocation flows need not be symmetric. At the calibrated configuration, robot reinstatement and cognitive displacement can push the physical-sector labor share in the same direction.*

Since opposite signs hold at the calibrated parameters (Proposition 4(c)), the masking theorem (Proposition 2) and aggregation bias (Corollary 3) apply.

4.5 Calibration and Estimation

Table 12 lists the calibrated baseline parameters, held fixed throughout.

Three parameters are matched to four independent wage moments. The two skill-premium moments are linear combinations of the wage levels, since $\Delta\text{SP} \equiv \Delta w^C - \Delta w^{NC}$, so they are predictions and not targets:[19] In log points × 100, ordered as Robot (NC, C, SP) followed by LMOE (NC, C, SP),

$$\begin{aligned} \hat{m} = 100 \cdot \big(&\hat{\beta}_R^{NC},\ \hat{\beta}_R^{C},\ \hat{\beta}_R^{SP},\\ \hat{\beta}_{\text{LMOE}}^{NC},\ &\hat{\beta}_{\text{LMOE}}^{C},\ \hat{\beta}_{\text{LMOE}}^{SP}\big)' \\ &= (+3.584,\ +2.571,\ -1.013,\ -5.176,\ -3.130,\ +2.046)' \end{aligned} \tag{31}$$

The parameter vector minimizes an inverse-variance weighted distance between model and data moments. The three fitted parameters are the physical reinstatement rate, differential

[19] Frontier-specific identification argument in Online Appendix D.3.1.

Table 12: Baseline Parameter Values

Parameter	Description	Value	Source
σ	Within-module task substitution	2.0	Acemoglu and Restrepo (2018b)
ε_w	Within-sector skill substitution	1.5	Katz and Murphy (1992)
μ	Physical task share	0.45	ACS occupation shares
κ	Comparative advantage	3.24	Cross-sector skill ratio (ACS)
a_H	College absolute advantage	1.65	Current Population Survey (CPS) wage levels
$\bar{I}_R$	Baseline physical threshold	0.30	30% physical automation
$\bar{I}_A$	Baseline cognitive threshold	0.25	25% cognitive automation
L_S, L_H	Labor supply	0.60, 0.40	CPS education shares
ε_D	Local labor demand elasticity	3.0	Suárez Serrato and Zidar (2016) range

Notes: Parameters held fixed throughout the estimation. Robot and AI shocks: $\Delta I = 0.08$, the frontier advance at which the model's wage responses are matched to the reduced-form moments (Section 4.5, sensitivity in Online Appendix D.3.1). $\varepsilon_D = 3$ sits just above the local labor demand elasticities in Suárez Serrato and Zidar (2016), which run from -1.77 to -2.49 across their specifications against a stated range of -1.7 to -3. $\varepsilon_S^{\text{mig}} = 0.78$ is their aggregate elasticity of labor supply, assigned here to the less mobile group, and $\varepsilon_H^{\text{mig}} = 3.12$ has no external counterpart (Online Appendix D.3.1). Sensitivity to baseline parameters is in Table OD.3, and Online Appendix B.7 shows that the calibrated $\kappa = 3.24$ satisfies the ordering condition $\kappa > 1$ of Proposition 1.

sorting and common frontier concentration. They are calibration values conditional on the reference measurement mapping and not estimates of the latent frontier parameters, and where the text says estimation it means the minimum-distance fit to the four wage targets. The weighting matrix is the inverse of the moment covariance estimated by a commuting-zone pairs bootstrap. The objective, optimization routine, and numerical Jacobian are in Online Appendix D.3.

Two conditions are maintained and not tested, and every structural quantity below is conditional on them. The first is that the residualized Webb Robot and LMOE exposures are orthogonal to the wage innovations, so that the four conditional projection coefficients can be read as the model's responses to a frontier advance. The balance tests of Section 3.3 are diagnostics consistent with that condition, and the paper's own IV evidence does not identify reinstatement. The second is the cardinal mapping ΔI from one standard deviation of exposure to a movement of the task frontier. The wage moments pin down the calibration conditional on that mapping and do not identify the mapping itself. Sensitivity across the

values the external evidence admits is reported in Online Appendix D.3.1.

The numerical Jacobian (Online Appendix D.3.1) is four primitive wage moments against three parameters, and it separates along frontier lines. The physical reinstatement rate loads on the robot moments, +2.17 and +1.28 on the two wage levels, so $\hat{\delta}_R$ reads the positive Webb Robot wage moments in Table 1. It also carries smaller loadings on the cognitive moments, +0.65 and +0.42, which are not an independent cognitive channel, because δ_A is not a free parameter and moves with δ_R along the fixed task-rate ratio. The cognitive moments discipline φ and ν. Both load on them with the same sign and different ratios, -4.28 against -1.06 for sorting and -6.35 against -3.66 for concentration, and it is the difference between those two ratios that tells the parameters apart. The whitened Jacobian has full column rank at the benchmark, with singular values 14.12, 2.95 and 2.17 and a condition number of 6.50, which is the local rank condition on the moment map at the reference mapping. Rank makes $J'WJ$ positive definite, but it does not on its own make the reported vector a strict local minimum, because the criterion's Hessian also carries a residual-weighted curvature term that rank leaves uncontrolled whenever the fit is not exact. A numerical global search over the maintained parameter space finds no competing minimum within a pre-specified tolerance (Online Appendix D.3.1), which is numerical evidence of global uniqueness and not an analytic proof. The reinstatement asymmetry $\delta_R \gg \delta_A$ is imposed by the external ratio, and what the calibration pins down at the reference mapping is the aggregate threshold ω^*_{struct}, the sign pattern, and ν. The full identification discussion, including the loading-versus-content distinction and the leave-one-moment-out diagnostics, is in Online Appendix D.4.

First, $\hat{\delta}_R = 1.233$, the *effective* reinstatement rate, the productive-task-content equivalent of about 1.23 boundary-productivity tasks per unit of physical frontier advance. At the reference mapping it exceeds the displacement threshold, so effective robot reinstatement more than offsets displacement. That comparison is a property of the reference calibration and not one the wage moments deliver on their own, because the level of $\hat{\delta}_R$ moves inversely

Table 13: Benchmark Structural Calibration

Parameter	Value	Source or discipline
Calibrated to the four primitive wage moments at the reference mapping		
δ_R (Physical reinstatement)	1.233	Humlum (2022): [0.6, 2.0]
φ (Differential sorting)	0.411	ALM (2003): > 0
ν (Common frontier concentration)	1.137	—
Externally disciplined, not estimated		
$\chi_\delta^{\text{task}}$ (Reinstatement ratio)	4.8148	Exposure-attributed task rates
γ_R/γ_A (Mapping benchmark)	1	Normalization (Online Appendix A.8)
$\varepsilon_S^{\text{mig}}$ (S mobility)	0.78	Suárez Serrato and Zidar (2016), Table 7
$\varepsilon_H^{\text{mig}}$ (H mobility)	3.12	Maintained, ordering externally documented
Implied		
δ_A (Cognitive reinstatement)	0.256	$= \delta_R/\chi_\delta^{\text{task}}$
Minimum distance, three parameters against four independent moments, $W = \hat{\Sigma}^{-1}$, $J = 0.09$, $p = 0.77$.		

Notes: Units-consistent minimum-distance calibration against the four primitive wage targets of equation 31, in log points $\times$ 100. The parameter values are ratios and carry no wage units. The two premium quantities are implied algebraically, since $\Delta\text{SP} \equiv \Delta w^C - \Delta w^{NC}$, and are not additional information. The bracketed interval for δ_R is the external Humlum comparison range (Online Appendix D.6), and the conditional pairs-bootstrap intervals for the calibration vector are in Online Appendix D.3. ALM refers to Autor et al. (2003).

with the exposure-to-frontier mapping, and over the mapping range the external evidence admits the reinstatement-dominated regime holds across the lower four fifths and turns over in the upper fifth (Online Appendix D.3.1). What the exercise establishes is that structural configurations consistent with the reduced-form pattern and with masking exist, without identifying which configuration generates it. Because the calibration is pinned in the effective rate $\delta_R = \delta_R^{\text{task}}\gamma_R^{\sigma-1}$ (Online Appendix A.8), the underlying task-creation *count* is at least this large, with equality when new tasks enter at full boundary productivity. It lies within the range $[0.6, 2.0]$ implied by Humlum (2022)'s independent firm-level estimates, and his core finding, total labor demand rising while production tasks fall, is qualitatively consistent with strong reinstatement, though the model's formal threshold $\bar{\delta}_R$ is defined on the marginal product of labor and not on employment. That external consistency is what carries the evidentiary weight. Three parameters against four independent moments leave one overidentifying restriction, which is not rejected ($J = 0.09$, $p = 0.77$), conditional on

the externally disciplined relative reinstatement rate and migration elasticities. Second, $\hat{\delta}_A = 0.256$. Cognitive reinstatement is finite, about a fifth of the physical rate, and stays below its own much higher threshold across the sensitivity range, so the asymmetry between frontiers is structural and not a corner. Third, common frontier concentration is strongly supported, $\hat{\nu} = 1.137$ with a conditional interval of [0.379, 1.593] that excludes zero in every re-estimate. The sorting parameter $\hat{\varphi} = 0.411$ is moderate at the benchmark but weakly pinned down on its own, with a conditional interval of [0.000, 1.171], and positive sorting is not sustained when either LMOE wage-level moment is omitted (Online Appendix D.3.2). The independent occupational gradient in Online Appendix A.9 is consistent with a moderate value and is not a calibration of it.

Sensitivity to the calibrated baseline parameters is in Online Appendix D.5. Opposite signs hold in all seven configurations and the structural masking threshold stays interior, ranging from 0.45 to 0.83.

4.6 Moment Fit and Mechanism Decomposition

The reported optimum of the units-consistent (log-point) objective fits the four independent moments to within 0.22 standard errors. Table 14 decomposes the predicted skill premium by adding each mechanism cumulatively, starting from pure displacement. The ordering separates the common channels from the differential ones at each stage: asymmetric reinstatement, then common frontier concentration, then differential sorting, then common local pass-through, then differential mobility. Both empirical signs are matched only in the final row. Each intermediate specification reproduces at most one of the two target signs. This describes the displayed cumulative path. The paper does not run a leave-one-out test that would establish each mechanism as separately necessary, though the rows show each one moving a distinct margin.

Table 14: Mechanism Decomposition

Specification	$\Delta\mathcal{P}_R$	$\Delta\mathcal{P}_{\text{LMOE}}$	Opposite?	Match?
Pure displacement	$+0.90$	-0.83	✓	×
+ Reinstatement	$+0.57$	-0.76	✓	×
+ Common concentration ν	$+0.56$	-0.72	✓	×
+ Differential sorting φ	$+0.43$	$+0.50$	×	×
+ Common pass-through	$+0.29$	$+0.34$	×	×
+ Differential mobility	-1.16	$+2.24$	✓	✓
Data	-1.01	$+2.05$	✓	—

Notes: Each row adds one mechanism to the previous specification, evaluated at the calibrated parameters of Table 13. Units are log points × 100 per standard deviation of exposure. "Opposite?" indicates whether $\Delta\mathcal{P}_R$ and $\Delta\mathcal{P}_{\text{LMOE}}$ have opposite signs. "Match?" indicates whether the signs match the empirical estimates ($\Delta\mathcal{P}_R < 0$, $\Delta\mathcal{P}_{\text{LMOE}} > 0$). Pure displacement uses baseline parameters only ($\delta_R = \delta_A = 0$, $\varphi = 0$, national labor market). The mechanism-by-mechanism reading of the rows is in the text. Data targets from equation 31.

Because $\hat{\delta}_A = 0.256$ against $\hat{\delta}_R = 1.233$, reinstatement operates asymmetrically. It moves the robot premium response from $+0.90$ to $+0.57$, because physical reinstatement is strong enough to blunt the displacement effect, while it barely moves the cognitive one, from -0.83 to -0.76. Without reinstatement exceeding displacement at the physical frontier, robot automation cannot raise wages at all.

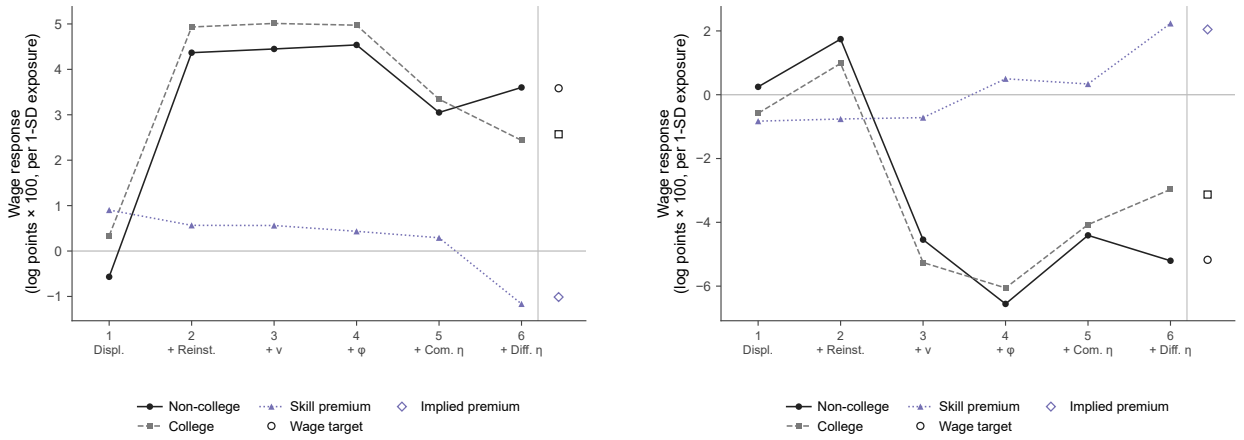
Common frontier concentration is the next rung and changes little on its own, $+0.57$ to $+0.56$ and -0.76 to -0.72, because it raises both cognitive skill weights by the same factor and so cancels from the within-sector premium. Its work is done in levels, not in the gap.

Differential sorting is the first rung that changes a sign. Concentrating cognitive displacement on non-college workers moves the cognitive premium response from -0.72 to $+0.50$, and it also pulls the robot response down to $+0.43$, so at this stage the two responses share a sign and the opposite-sign pattern is temporarily broken. Sorting alone does not deliver

the reversal.

Spatial equilibrium completes it, and in two steps that must be kept apart. Applying a common pass-through to both groups scales both responses toward zero, to +0.29 and +0.34, so it changes magnitudes and no signs. Allowing the pass-through to differ across groups, with $\varepsilon_S^{\text{mig}} = 0.78$ against $\varepsilon_H^{\text{mig}} = 3.12$, is what restores the pattern, moving the robot response to −1.16 and the cognitive one to +2.24 against empirical targets of −1.01 and +2.05. The full model restores opposite signs, in line with Proposition 4(c), so the masking result (Proposition 2) and aggregation bias (Corollary 3) hold under the extended model.

Figure 9 plots the same six rungs for the two wage levels as well as the premium. It shows that the reversal is not produced by any single mechanism, since differential sorting flips the cognitive premium, and differential mobility flips the physical one back, with common concentration and common pass-through moving levels without touching either sign.



(a) Robot effect

(b) LMOE effect

Figure 9: **Mechanism Decomposition.** Predicted local wage and premium responses (log points × 100 per one-standard-deviation exposure) as the six stages are added in turn, from pure displacement to the full calibrated model. Panel A: robot exposure. Panel B: LMOE exposure. Stages: 1 pure displacement, 2 plus asymmetric reinstatement, 3 plus common concentration ν, 4 plus differential sorting φ, 5 plus common pass-through η, 6 plus differential pass-through. Open markers at the right edge are the empirical targets. The two wage-level targets keep their series' marker shape. The premium target carries a distinct open diamond because it is derived from those two and is not an independent moment. Pure displacement, the first rung, already produces opposite signs but in the wrong direction (Table 11). Parameter values from Table 13.

4.7 Masking Quantification

Under the calibrated model, $\hat{\omega}^*_{\text{struct}} = 0.658$ at the reference solution, with a conditional pairs-bootstrap interval of [0.561, 0.764] that is interior and excludes one, and opposite signs hold in 2,000 of 2,000 re-estimates. Under pure displacement, with both reinstatement rates set to zero, the threshold falls to $\omega^*_{\text{pure}} = 0.479$, and the gap between the two is what the three extensions contribute together, reinstatement, sorting and mobility, as the decomposition of Table 14 separates them. The reduced-form counterpart $\omega^*_{\text{RF}} = 0.669$ is computed from the same skill-premium responses the estimation targets, so its proximity is a consistency check and not corroboration (Online Appendix E.14). Full inference is in Table OE.29.

Remark 3 (Quantifying the masking bias). At equal weights ($\tilde{\omega} = 0.5$), the single-index prediction is +0.52, opposite in sign to the robot-driven compression that dominates the U.S. economy. The prediction error is $|+0.52 - (-1.01)| + |+0.52 - (+2.05)| = 3.06$ log points with sign reversal, and at $\tilde{\omega} = \omega^*_{\text{RF}} = 0.67$ the single-index prediction crosses zero. This provides a structural interpretation of the empirical null. A pooled automation index yields $\hat{\beta} = 0.009$ ($t = 1.46$), while separating the frontiers raises the incremental R^2 over controls from 0.004 to 0.030 (Table 7, columns 5 and 8, together with the controls-only $R^2 = 0.360$ in Table OE.11).

4.8 External Comparisons and Directional Benchmarks

A structural model estimated from U.S. commuting-zone regressions must be disciplined by external evidence. The calibrated parameters are not free to take any values that fit the moments. They must also be consistent with independent estimates from different countries, aggregation levels, and identification strategies. Table 15 summarizes the comparison.

The physical reinstatement finding, an effective rate of $\hat{\delta}_R = 1.233$, is consistent with firm-level evidence from Denmark and Spain and with cross-country evidence from 17 countries, which point to the reinstatement asymmetry from outside these data. The Denmark mapping

Table 15: External Comparisons and Directional Benchmarks

Parameter	External Source	Implied Value or Direction	Consistent
$\delta_R = 1.233$	Humlum (2022): Denmark firms	[0.6, 2.0]	✓
Strong physical reinstatement	Graetz & Michaels (2018): 17 countries	Qualitative	Directional
Strong physical reinstatement	Koch et al. (2021): Spain firms	Qualitative	Directional
δ_R dynamics	Chung & Lee (2023): US CZs	neg→pos reversal	Directional
δ_A below threshold	Humlum & Vestergaard (2025): Denmark	null earnings effects	Directional
$\varphi \approx 0.41$	Autor, Levy & Murnane (2003)	within-occ. sorting	✓
Low-skill workers less mobile	Bound & Holzer (2000)	differential mobility	✓

Notes: Each row compares a calibrated parameter (Table 13) or a model direction with evidence from another setting. Humlum (2022): Danish firm estimates, mapped into $[0.6, 2.0]$ in Online Appendix D.6. Graetz and Michaels (2018): robots raise productivity and wages with null employment effects across 17 countries. Koch et al. (2021): output growth of almost 25% and roughly 10% employment expansion in Spanish adopters. Chung and Lee (2023): reversal from negative (2005–2010) to positive (2011–2016), against a 2019 break here. Humlum and Vestergaard (2025): precise null earnings effects two years post-ChatGPT. Autor et al. (2003): within-occupation sorting (Section 4.2). Hornbeck and Moretti (2024): skill-specific mobility ratio of 0.39, against $0.78/3.12 \approx 0.25$ here (Online Appendix D.3.1).

is quantitative and the other two are directional. The much weaker cognitive reinstatement ($\hat{\delta}_A = 0.256$, about a fifth of the physical rate) aligns with Humlum and Vestergaard (2025)'s precise null and with the broader pattern of large individual productivity gains from AI (Noy and Zhang, 2023; Brynjolfsson et al., 2025) coexisting with null aggregate wage effects, which provide contextual evidence and not a direct prediction or estimate of δ_A. Within the model, $\hat{\delta}_A = 0.256$ leaves the cognitive frontier displacement-dominated, since reinstatement there is positive but far below its own boundary, so it offsets only part of the displacement.

Read through the model, the calibrated threshold offers an illustrative interpretation of three eras, with rising inequality under cognitive automation (1980s–1990s, a low robot share $\omega \approx 0$), stagnation as robots introduced an opposing force (2000s–2010s, $\omega \approx \omega^*$), and potentially renewed widening as generative AI shifts the mix back toward the cognitive continuum. The era-specific mix values are an interpretation and not an estimate, because the model is calibrated to the 2005–2019 cross-section alone. Within that estimation window, the fitted $\omega^*_{\text{struct}} = 0.658$ implies that once the robot share exceeded roughly two-thirds, the net effect switched from widening to compression. If cognitive reinstatement stays below its threshold ($\hat{\delta}_A = 0.256$ against a cognitive output boundary of 4.14), the model predicts renewed widening under AI-dominant technological change. The cognitive coefficient is positive over 2005–2024 and indistinguishable from zero over 2019–2024 alone (Section 3.5), so the post-2019 window does not yet show the predicted renewed widening and leaves its future emergence unresolved.

5 Conclusion

This paper asked whether automation widens or compresses inequality, and found the question underspecified. Robot exposure tracks rising wages for both education groups and a compressing skill premium, and AI exposure tracks rising college wages and a widening one. Because the two carry opposite signs, a pooled index of the two explains 4 percent of the

cross-zone variation in the skill premium against 31 percent for the separated frontiers, and under controls it changes sign. The sign of an aggregate automation coefficient therefore depends on the mix of the two frontiers as much as on the strength of either, and it turns negative only above a robot weight of roughly two-thirds.

The calibrated model rationalizes the pattern through asymmetric reinstatement. At the reference calibration physical automation creates new tasks fast enough to offset its own displacement, while cognitive automation creates them at about a fifth of that rate against an output boundary six times higher, because displacement there destroys tasks made valuable by concentration near the frontier while reinstatement returns tasks that carry no such weight. Differential sorting and differential mobility then convert those national movements into the local wage ordering the data show. The calibration is conditional on how the exposure measures are mapped onto the two frontiers.

The general point reaches past automation. Wherever a scalar index aggregates channels with opposite-signed effects on an outcome, the aggregate can return a precisely estimated null while each channel moves the outcome by a large amount. The condition is testable, because the components of a composite index should share the same sign on the outcome, and here they do not. Where the two channels do agree in sign, aggregation amplifies instead of masking, which is what makes the account refutable and what the data show. The immediate implication is dated. Robot density in American manufacturing kept rising through 2023 while the compression channel turned, so that after 2019 the same exposure that had raised non-college wages lowers them, in the Webb AI horse race and not in the LMOE one. Cognitive-frontier adoption accelerated over the same years. If new-task creation at the physical frontier continues to slow relative to AI-dominant automation, the mix shifts toward the cognitive frontier and the skill premium should widen.

References

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of Labor Economics*, 4:1043–1171.

Acemoglu, D., Autor, D., Hazell, J., and Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics*, 40(S1):S293–S340.

Acemoglu, D., Kong, F., and Restrepo, P. (2025). Tasks at work: Comparative advantage, technology and labor demand. In *Handbook of Labor Economics*, volume 6, pages 1–114. Elsevier.

Acemoglu, D. and Restrepo, P. (2018a). Low-skill and high-skill automation. *Journal of Human Capital*, 12(2):204–232.

Acemoglu, D. and Restrepo, P. (2018b). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6):1488–1542.

Acemoglu, D. and Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2):3–30.

Acemoglu, D. and Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy*, 128(6):2188–2244.

Acemoglu, D. and Restrepo, P. (2022). Tasks, automation, and the rise in U.S. wage inequality. *Econometrica*, 90(5):1973–2016.

Adão, R., Koles'ar, M., and Morales, E. (2019). Shift-share designs: Theory and inference. *Quarterly Journal of Economics*, 134(4):1949–2010.

Autor, D. H. 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.

Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review*, 103(6):2121–2168.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118(4):1279–1333.

Babina, T., Fedyk, A., He, A., and Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, 151:103745.

Babina, T., Fedyk, A., He, A., and Hodson, J. (2025). Artificial intelligence makes firm operating performance less volatile. *AEA Papers and Proceedings*, 115:35–39.

Bloom, D. E., Prettner, K., Saadaoui, J., and Veruete, M. (2025). Artificial intelligence and the skill premium. *Finance Research Letters*, 81:107401. NBER Working Paper No. 32430, 2024.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *Review of Economic Studies*, 89(1):181–213.

Bound, J. and Holzer, H. J. (2000). Demand shifts, population adjustments, and labor market outcomes during the 1980s. *Journal of Labor Economics*, 18(1):20–54.

Brynjolfsson, E., Li, D., and Raymond, L. R. (2025). Generative AI at work. *Quarterly Journal of Economics*, 140(2):889–942.

Chung, J. and Lee, Y. S. (2023). The evolving impact of robots on jobs. *ILR Review*, 76(2):290–319.

Combemale, C., Whitefoot, K. S., Ales, L., and Fuchs, E. R. H. (2021). Not all technological change is equal: How the separability of tasks mediates the effect of technology change on skill demand. *Industrial and Corporate Change*, 30(6):1361–1387.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, 384(6702):1306–1308.

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

Fossen, F. M., Samaan, D., and Sorgner, A. (2022). How are patented AI, software and robot technologies related to wage changes in the united states? *Frontiers in Artificial Intelligence*, 5:869282.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Graetz, G. and Michaels, G. (2018). Robots at work. *Review of Economics and Statistics*, 100(5):753–768.

Hornbeck, R. and Moretti, E. (2024). Estimating who benefits from productivity growth: Local and distant effects of city productivity growth on wages, rents, and inequality. *The Review of Economics and Statistics*, 106(3):587–607.

Humlum, A. (2022). Robot adoption and labor market dynamics. ROCKWOOL Foundation Study Paper 175, May 2022.

Humlum, A. and Vestergaard, E. (2025). Still waters, rapid currents: Early labor market transformation under generative AI. NBER Working Paper No. 33777, revised March 2026.

International Federation of Robotics (2024). World robotics 2024: Industrial robots. Technical report, International Federation of Robotics, Frankfurt am Main, Germany.

Katz, L. F. and Murphy, K. M. (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics*, 107(1):35–78.

Koch, M., Manuylov, I., and Smolka, M. (2021). Robots and firms. *The Economic Journal*, 131(638):2553–2584.

Lankisch, C., Prettner, K., and Prskawetz, A. (2019). How can robots affect wage inequality? *Economic Modelling*, 81:161–169.

Montiel Olea, J. L. and Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3):358–369.

Moretti, E. (2011). Local labor markets. *Handbook of Labor Economics*, 4:1237–1313.

Notowidigdo, M. J. (2020). The incidence of local labor demand shocks. *Journal of Labor Economics*, 38(3):687–725.

Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192.

Ribeiro, M. J. and Prettner, K. (2025). The skill premium across countries in the era of industrial robots and generative AI. Working paper.

Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear IV regression. In Andrews, D. W. K. and Stock, J. H., editors, *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, pages 80–108. Cambridge University Press, New York.

Suárez Serrato, J. C. and Zidar, O. (2016). Who benefits from state corporate tax cuts? a local labor markets approach with heterogeneous firms. *American Economic Review*, 106(9):2582–2624.

Tolbert, C. M. and Sizer, M. (1996). U.S. commuting zones and labor market areas: A 1990 update. Staff Report AGES-9614, U.S. Department of Agriculture, Economic Research Service.

Webb, M. (2020). The impact of artificial intelligence on the labor market. Stanford University Working Paper.

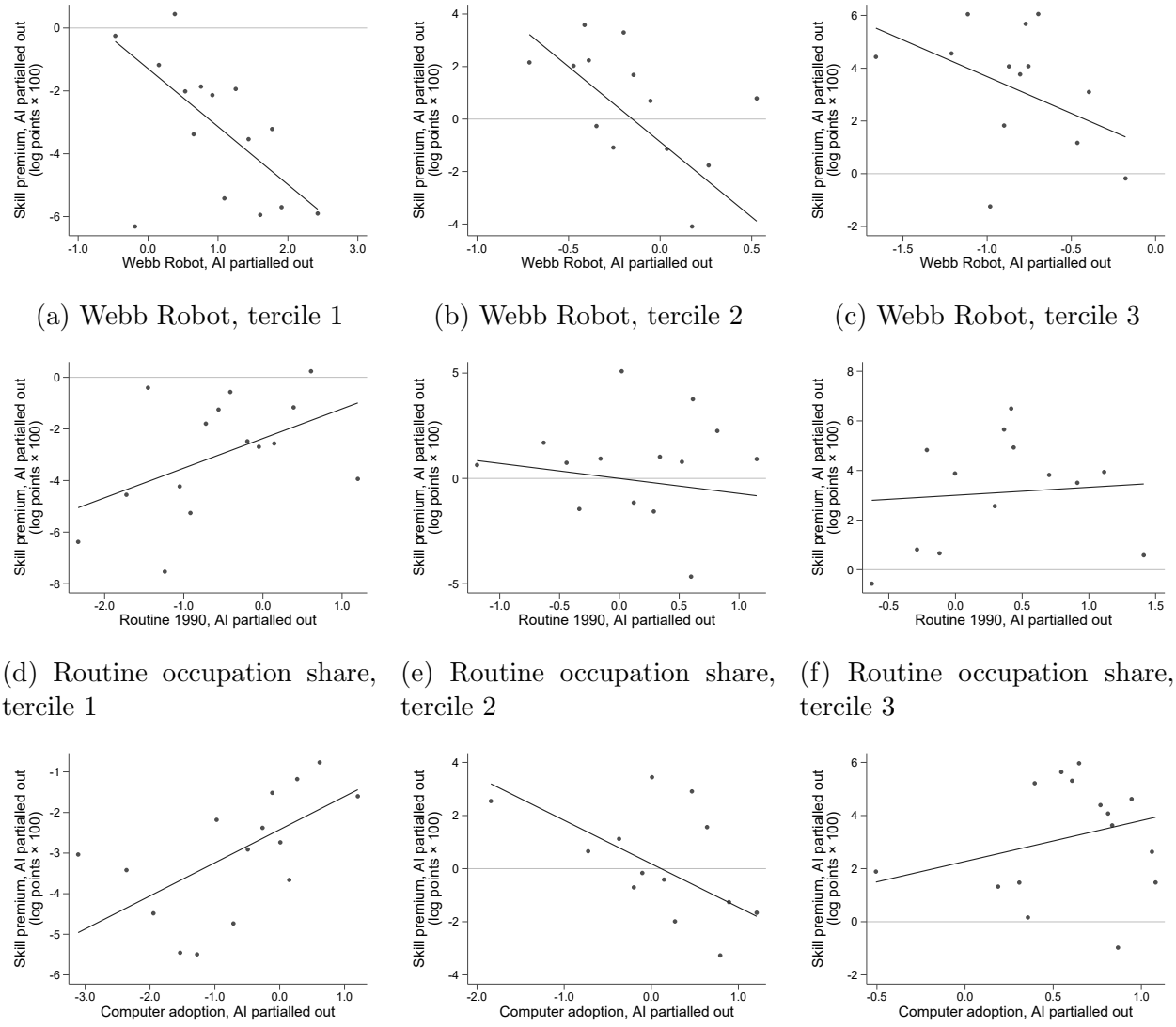
Zeira, J. (1998). Workers, machines, and economic growth. *Quarterly Journal of Economics*, 113(4):1091–1117.

A Additional Tables and Figures

Table A.1: Cross-Frontier Interaction: $\hat{\xi}$ on $Z_c^P \times A_c$

	Webb Robot (1)	Routine occ. share (1990) (2)	Computer adoption (3)
Panel A: Raw interaction			
No controls	-0.007^{**}	$+0.009^{**}$	-0.006^{***}
	(0.003)	(0.004)	(0.002)
Baseline controls	-0.006^{*}	$+0.010^{**}$	-0.004^{**}
	(0.003)	(0.004)	(0.002)
Controls + SP_0	-0.006^{**}	$+0.009^{**}$	-0.003^{*}
	(0.003)	(0.004)	(0.001)
Kitchen sink	-0.007^{*}	$+0.011^{***}$	-0.005^{***}
	(0.004)	(0.004)	(0.002)
Panel B: Residualized interaction ($Z_c^{P\perp} \times A_c^{\perp}$)			
Resid. on X	-0.009	$+0.010$	-0.013^{***}
	(0.007)	(0.006)	(0.005)
Resid. on X + tercile FE	-0.011	$+0.009$	-0.013^{**}
	(0.007)	(0.007)	(0.006)
Panel C: Within college-share terciles (raw + X)			
Low college	-0.004	-0.007	-0.011
	(0.005)	(0.005)	(0.010)
Mid college	-0.026	$+0.028^{**}$	-0.018^{*}
	(0.018)	(0.013)	(0.011)
High college	-0.038^{***}	$+0.023^{***}$	-0.008^{*}
	(0.013)	(0.008)	(0.004)
N (full sample)	722	722	677

Notes: Each cell reports $\hat{\xi}$ from equation (24). A_c is Webb AI throughout. Panel A: raw interaction across four control sets (no controls, baseline X, X plus initial skill premium, and a kitchen sink with quadratics, spline, and tercile FEs). Panel B: P and A residualized on baseline controls (and tercile FEs in second row). The interaction is further orthogonalized to $P^{\perp}$ and $A^{\perp}$ levels. Panel C: estimated within college-share terciles with baseline controls. Population weights, robust SEs. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.



(a) Webb Robot, tercile 1

(b) Webb Robot, tercile 2

(c) Webb Robot, tercile 3

(d) Routine occupation share, tercile 1

(e) Routine occupation share, tercile 2

(f) Routine occupation share, tercile 3

(g) Computer adoption, tercile 1

(h) Computer adoption, tercile 2

(i) Computer adoption, tercile 3

Figure A.1: **Physical-frontier exposure and skill premium changes, within college-share terciles, after partialling out AI.** Each row shows one physical-frontier measure and each panel within a row corresponds to a college-share tercile. Both axes are residualized on AI exposure, so the vertical axis is $\Delta\mathcal{P}$ net of AI and the horizontal axis is the physical measure net of AI, and the plotted slopes are within-tercile physical-exposure gradients, not the interaction coefficients $\hat{\xi}$ of equation (24) in Table A.1. Top row: Webb Robot. Middle row: the routine occupation share measured in 1990. Bottom row: computer adoption, the Babina et al. (2024) share of workers in computer occupations. 15 quantile bins, population-weighted.

Table A.2: Stacked Long Differences: Technology Effects on $\Delta\mathcal{P}$ Across Sub-Periods

	(1) 2005–2010	(2) 2010–2019	(3) Pooled	(4) Exposures × Late	(5) Fully interacted
Robot (Webb)	-0.0111^{***}	-0.0091^{***}	-0.0096^{***}	-0.0042^{*}	-0.0111^{***}
	(0.0027)	(0.0035)	(0.0020)	(0.0023)	(0.0027)
AI (Webb)	$+0.0046^{***}$	$+0.0090^{***}$	$+0.0065^{***}$	$+0.0037^{**}$	$+0.0046^{***}$
	(0.0017)	(0.0025)	(0.0015)	(0.0017)	(0.0017)
Robot × Late				-0.0112^{***}	$+0.0020$
				(0.0025)	(0.0047)
AI × Late				$+0.0059^{**}$	$+0.0044$
				(0.0030)	(0.0031)
Baseline SP	-0.232^{***}	-0.375^{***}	-0.288^{***}	-0.304^{***}	-0.232^{***}
	(0.031)	(0.040)	(0.025)	(0.025)	(0.031)
Baseline SP × Late					-0.144^{***}
					(0.051)
Period FE	No	No	Yes	Yes	Yes
Controls × Late	—	—	No	No	Yes
N	722	722	1,444	1,444	1,444
R^2	0.266	0.441	0.329	0.356	0.378

Notes: Dependent variable: $\Delta\mathcal{P}$ for the indicated sub-period. All specifications include baseline (2005) demographic controls and period-start skill premium. All columns use CZ-clustered SEs, and columns (3)–(5) stack the two windows with period fixed effects. "Late" = $\mathbf{1}$[period = 2010–2019]. WLS with 2005 population weights. The interactions in column (4) hold the demographic controls and the period-start skill premium to one coefficient across periods. Column (5) lets every regressor move with the period, so its exposure interactions are the differences between the two sub-period coefficients of columns (1)–(2). The restriction separating the two columns is testable and it fails, since the period-start skill premium alone moves by -0.144 (SE 0.051) in column (5). The parallel *LMOE specification* (*Webb Robot* + LMOE, pooled controls) gives, in the interacted column, an LMOE base of $+0.016$ (SE 0.006) and an LMOE × Late interaction of $+0.017$ (SE 0.006), implying a late-period LMOE effect of $+0.033$ (SE 0.007), with a Robot base of $+0.004$ and a Robot × Late interaction of $+0.008$, $N = 1{,}444$. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A.3: Horse-Race Specification by College-Share Tercile

	College-Share Tercile			
	Low	Mid	High	Full
Panel A: No controls				
Robot Exposure	-0.016^{***}	-0.058^{***}	-0.026^{**}	-0.031^{***}
	(0.005)	(0.021)	(0.011)	(0.003)
AI Exposure	$+0.012^{*}$	$+0.024^{**}$	$+0.025^{***}$	$+0.023^{***}$
	(0.007)	(0.010)	(0.008)	(0.005)
R^2	0.044	0.262	0.363	0.308
Panel B: Baseline controls				
Robot Exposure	-0.002	-0.033^{**}	-0.024	-0.014^{**}
	(0.008)	(0.015)	(0.017)	(0.007)
AI Exposure	$+0.005$	$+0.019^{*}$	$+0.024^{**}$	$+0.015^{***}$
	(0.007)	(0.011)	(0.011)	(0.005)
R^2	0.119	0.387	0.531	0.390
Panel C: Baseline controls + baseline SP				
Robot Exposure	-0.008	-0.022^{*}	-0.029^{**}	-0.014^{**}
	(0.007)	(0.013)	(0.014)	(0.006)
AI Exposure	$+0.003$	$+0.022^{**}$	$+0.023^{**}$	$+0.013^{***}$
	(0.005)	(0.011)	(0.009)	(0.005)
R^2	0.250	0.439	0.696	0.471
N	579	100	43	722

Notes: Each column restricts to one population-weighted tercile of baseline college share (low: $\leq 27.1\%$, mid: 27.1–33.1%, high: $\geq 33.1\%$). Dependent variable: $\Delta\mathcal{P}$ (2005–2019), computed from single-year endpoints (2005 to 2019), so the full-sample coefficients differ slightly from the endpoint-window construction of Table 9 (Robot is -0.014 here and -0.015 there). Population-weighted, robust standard errors. Baseline controls: college share, manufacturing share, female share, nonwhite share, mean age. Panel C adds baseline skill premium, the 2005 value. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.4: Seven Measures of Automation Exposure

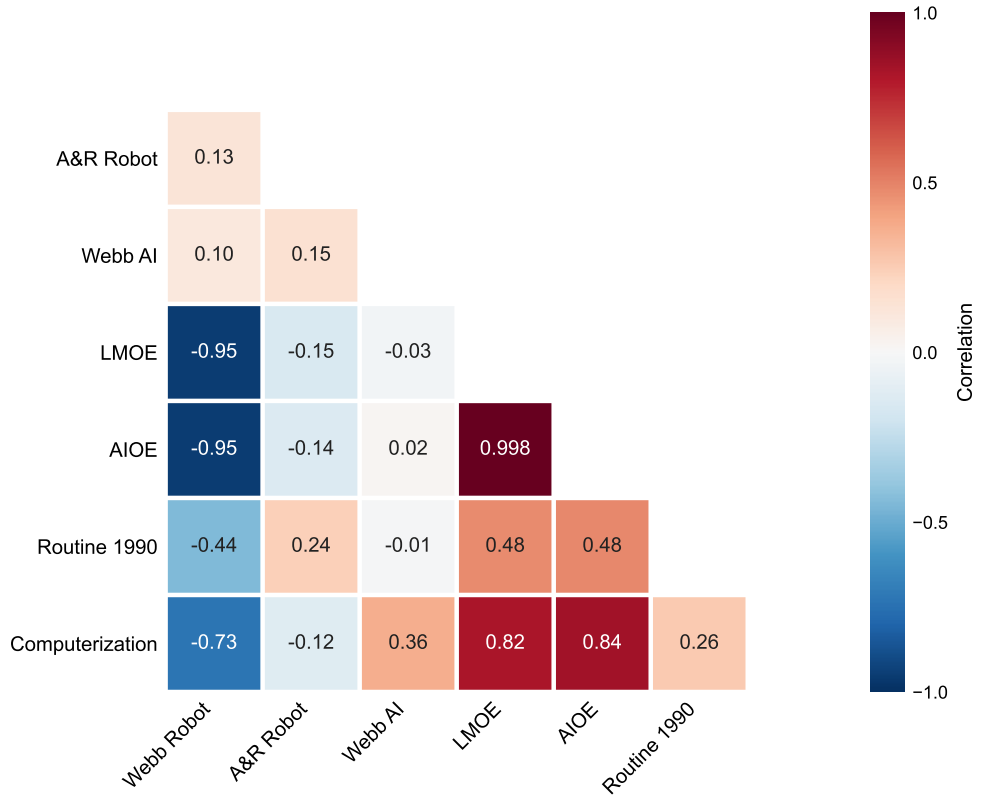
Measure	Source	Construction	Type	Grouping
Panel A: Robot/physical automation				
Webb Robot	Webb (2020)	Patent–occupation overlap using NLP of patent texts and O*NET task descriptions	Feasibility	Physical
A&R Robot	Acemoglu and Restrepo (2020)	Industry-level robot adoption (IFR) allocated to CZs via employment shares	Deployment	Physical
Panel B: AI/cognitive automation				
Webb AI	Webb (2020)	Same NLP method applied to AI patents	Feasibility	Cognitive
LMOE (GenAI)	Eloundou et al. (2024)	Language modeling occupational exposure, based on GPT-4 capabilities	Feasibility	Physical†
AIOE (General AI)	Felten et al. (2021)	AI occupational exposure, mapping AI benchmarks to O*NET abilities	Feasibility	Physical†
Panel C: Other technology measures				
Routine task intensity (AD)	Autor and Dorn (2013)	Routine task intensity of CZ employment	Feasibility	Physical
Computerization	Acemoglu and Restrepo (2020)	Task-based computerization intensity of CZ employment	Feasibility	—

Notes: All measures are standardized to mean zero and unit standard deviation. "Type" distinguishes *feasibility* measures, which record whether the technology to automate a task exists, from *deployment* measures, which record observed firm- or industry-level adoption. "Grouping" reports the automation frontier a measure groups with empirically, from its factor loading. A&R Robot is grouped from its construction as a deployment measure, because the factor structure leaves it unclassified, and Computerization is excluded from the factor analysis and has no grouping. † LMOE and AIOE are built as AI measures but load on the physical frontier ($r = -0.95$ with Webb Robot), which Section 3.2 discusses.

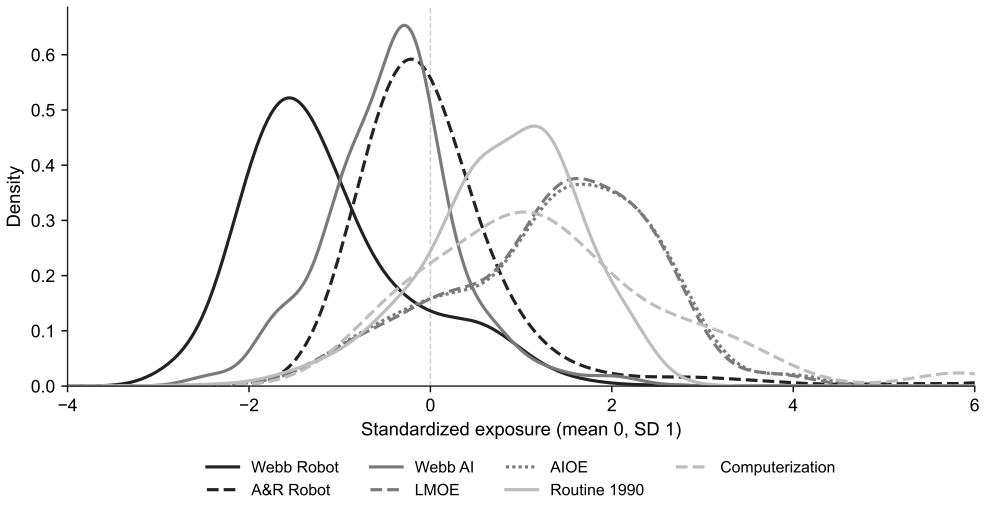
Table A.5: Descriptive Statistics

	Mean	S.D.	p10	Median	p90	N
Panel A: Outcome variables (2005–2019)						
Δ Skill premium	−0.012	0.047	−0.123	−0.034	0.033	722
$\Delta \ln w^C$ (college)	0.173	0.046	0.107	0.190	0.289	722
$\Delta \ln w^{NC}$ (non-coll)	0.185	0.044	0.162	0.226	0.321	722
Panel B: Outcome variables (2019–2024)						
Δ Skill premium	−0.021	0.034	−0.095	−0.024	0.058	722
$\Delta \ln w^C$ (college)	0.175	0.035	0.105	0.169	0.239	722
$\Delta \ln w^{NC}$ (non-coll)	0.196	0.026	0.144	0.196	0.241	722
Panel C: Outcome variables (2005–2024)						
Δ Skill premium	−0.033	0.055	−0.162	−0.058	0.031	722
$\Delta \ln w^C$ (college)	0.349	0.058	0.277	0.364	0.470	722
$\Delta \ln w^{NC}$ (non-coll)	0.381	0.048	0.357	0.422	0.514	722
Panel D: Technology exposure measures (standardized)						
Robot exposure (Webb)	0.000	1.000	−1.419	0.153	1.135	722
AI exposure (Webb)	0.000	1.000	−1.134	−0.174	1.462	722
Robot exposure (A&R)	0.000	1.000	−0.890	−0.240	0.887	722
LMOE (GenAI)	0.000	1.000	−1.128	−0.139	1.459	722
AIOE (General AI)	0.000	1.000	−1.092	−0.141	1.448	722
Panel E: Baseline CZ characteristics (2005)						
College share	0.301	0.076	0.161	0.210	0.309	722
Manufacturing share	0.125	0.055	0.052	0.130	0.227	722
Female share	0.468	0.017	0.451	0.476	0.497	722
Nonwhite share	0.238	0.128	0.035	0.104	0.321	722
Mean worker age	39.660	0.868	38.485	39.875	40.963	722
Skill premium level	0.565	0.073	0.398	0.498	0.591	722

Notes: Panels A–C and E report population-weighted moments. Panel D reports the unweighted standardization moments, so each exposure has mean zero and unit standard deviation by construction. The population-weighted standard deviations are not one and differ across measures. Percentiles are unweighted throughout. The 1990 commuting-zone partition yields 722 mainland zones, and all three outcome panels hold the same 722. The panels are therefore additive, and the 2005–2019 and 2019–2024 premium changes sum to the 2005–2024 figure up to rounding.



(a) Pairwise correlations (population-weighted)



(b) Kernel density estimates

Figure A.2: **Technology Exposure Measures: Correlations and Distributions.** Panel (a): population-weighted pairwise correlations across 722 commuting zones. Panel (b): kernel densities of the standardized measures. The robot measures (navy) differ: Webb Robot is mildly left-skewed while A&R Robot is strongly right-skewed. The AI measures (brick) are right-skewed. The panels document the measures' covariance structure and marginal distributions, Table 7 reports the pooled-versus-frontier-specific regression comparison.