UNIVERSIDAD TORCUATO DI TELLA
Escuela de Gobierno

ESCUELA DE GOBIERNO

DOCUMENTOS DE TRABAJO 2026/06

Too Old to Adjust

Aging and the Speed of Automation Adoption

Eduardo Levy Yeyati

Abril 2026

Documentos de trabajo: https://bit.ly/2REorES

UTDT: Av. Figueroa Alcorta 7350, C1428BCW Buenos Aires, Argentina

Too Old to Adjust

Aging and the Speed of Automation Adoption

Eduardo Levy Yeyati
Universidad Torcuato Di Tella and IDB

April 2026

Abstract

We study how demographic composition changes the transition costs of rapid automation adoption in a speed–capacity model where displaced workers differ by age and the young share of the displacement pool declines over time. Older workers face higher discouragement hazards and lower retraining completion rates, so demographic aging deteriorates aggregate transition outcomes even when institutional capacity is unchanged—a demographic composition effect. This effect interacts with adoption speed through congestion: fast adoption in an older displacement pool is worse than the sum of its parts (supermodularity). We derive a demographic gradient in optimal adoption speed: the planner adopts faster when the displacement pool is younger. Numerically, in the calibrated region, this demographic gradient becomes steeper when aging is faster. We also identify a demographic buffer: a timing boundary beyond which earlier adoption may dominate delay, although the sufficient condition is quantitatively demanding under OECD-style calibrations. We derive six cross-country predictions and an explicit measurement strategy.

JEL: J24, J26, O33, E24
Keywords: AI adoption, demographic aging, labor force participation, retraining, optimal policy

1 Introduction

When automation displaces workers, the difficulty of adjustment depends not only on institutional capacity but also on *who* must adjust. A workforce composed primarily of younger routine workers can absorb displacement through retraining and reallocation. A workforce composed increasingly of older workers may instead experience discouragement, non-completion, and permanent labor force exit. This paper studies that composition margin. Its central claim is that the age structure of the displacement pool changes the socially optimal speed of automation adoption even when institutions are held fixed.

The workforce in advanced economies is aging at the same time that AI is accelerating the displacement of routine workers. As the OECD Employment Outlook 2025 documents,[1] the old-age dependency ratio across the OECD rose from 19% in 1980 to 31% in 2023 and is projected to reach 52% by 2060. But the aggregate dependency ratio understates the problem for labor market adjustment: what matters here is the age structure *within the occupations most exposed to automation.* In manufacturing, clerical, and administrative support—the routine-intensive sectors at the frontier of AI displacement—the share of workers over 55 has risen sharply in Germany, Italy, Japan, and Korea, precisely the economies where AI adoption is advancing fastest.

The paper asks a simple question: does it matter how old the displaced workers are when the displacement arrives? The answer is yes—and not only because older workers do worse individually. The key insight is that *the composition of the displacement pool changes aggregate transition outcomes even when institutional capacity is unchanged.* We call this the demographic composition effect. As the workforce ages, the pool of workers waiting to retrain shifts toward those with shorter remaining horizons, higher discouragement hazards, and lower retraining completion rates. The pipeline is the same; the workers it must process become harder to move through it.

To study this mechanism, we build a transition model in which automation diffuses across routine tasks, displaced workers pass through a congested retraining pipeline, and the displacement pool contains young and old workers. The young share declines at rate g_λ, so the composition of the pool worsens over time even if retraining capacity $\rho(Z)$ is fixed. The object of interest is the demographic composition margin itself: how age structure changes incidence, aggregate retraining, and the planner's preferred adoption speed. A homogeneous economy appears only as a limiting case.

The main results are straightforward. Permanent exit is disproportionately borne by old workers, and that age–exit incidence is supermodular in adoption speed and old-worker share: fast adoption in an old pool is worse than the sum of its parts (Propositions 1–2). Aggregate retraining deteriorates with demographic aging for any fixed

[1] OECD (2025), Ch. 2 (Editorial) and Ch. 3 (Older workers). Dependency ratio projections from Ch. 2, Figure 2.1.

institutional capacity (Proposition 3).

On the policy side, the planner's optimal adoption speed is monotone increasing in the young share λ_0 (Proposition 4). A calibrated numerical implication shows that this demographic gradient becomes steeper when aging is faster. A corollary establishes that the marginal welfare return to reskilling investment is strictly increasing in the old-worker share—aging and retraining are welfare complements (Corollary 1). Finally, the paper characterizes a demographic buffer: a timing boundary under which earlier adoption may dominate later adoption, although the sufficient condition is quantitatively demanding under the benchmark calibration (Proposition $4'$).

The contribution is threefold. First, the paper identifies a demographic composition channel through which aging worsens transition outcomes even when institutions are unchanged. Second, it shows that optimal adoption speed is demographic: economies with younger displacement pools can absorb faster automation waves. Third, it translates these mechanisms into testable empirical predictions and a concrete measurement strategy.

A related homogeneous-worker benchmark in Levy Yeyati (2026) isolates the speed–capacity problem under a common displacement pool. This paper introduces age heterogeneity into that transition and studies how demographic composition changes incidence, retraining, and the planner's preferred adoption speed. More broadly, the point is that transition costs depend not only on how much automation arrives, but on how fast it arrives and who must adjust.

The age dimension of labor market adjustment has a long empirical pedigree. Jacobson, LaLonde, and Sullivan (1993) documented persistent earnings losses from displacement, which subsequent work showed to be larger and more persistent for older workers: displaced workers over 55 face reemployment rates 20–30 percentage points below prime-age workers, with median unemployment durations roughly double (Farber, 2017; Johnson and Gosselin, 2018). The returns to retraining decline sharply with age through a present-value mechanism formalized by Heckman (2006) and Carneiro and Heckman (2003): with fewer remaining working years, the lifetime return to a costly retraining spell falls below the threshold needed to justify the investment. Kubeck et al. (1996) provide meta-analytic evidence that older adults master training material less completely and require more time, while Dix-Carneiro (2014) and Caliendo, Dvorkin, and Parro (2019) estimate structural reallocation costs that are strongly increasing in age.

The task-based framework of Acemoglu and Restrepo (2018, 2019) provides the production structure underlying the transition. The connection between demographic aging and automation has been studied primarily from the adoption-incentive side: Acemoglu and Restrepo (2022) show that aging economies adopt more robots because labor scarcity raises the return to automation, Prettner and Strulik (2020) model automation as a response to workforce shrinkage, and Abeliansky and Prettner (2023) document the empirical robot–aging correlation. A common feature of these papers is that aging *increases* *the demand for automation*. Our contribution is to show that aging simultaneously *raises the cost of the labor market transition* that automation requires. Put differently, existing work explains why older economies may want more automation; this paper explains why older economies can bear less speed.

A complementary recent literature studies the age incidence of generative AI adoption directly. Descriptive evidence from the United States suggests that early AI substitution has fallen more heavily on younger and less experienced workers: Peng (2026) reports that since the diffusion of generative AI began, the unemployment-rate gap between entry-level workers (under 30) and experienced workers (31–50) has widened by roughly 0.6 percentage points and the wage gap by 1.3 percent relative to pre-pandemic averages, with the widening concentrated in occupations more exposed to AI substitution. One plausible interpretation is that generative AI first substitutes tasks that are disproportionately assigned to junior workers, while more experience-intensive tasks remain relatively complementary. This pattern sits in apparent tension with the older displacement literature on the absorbing-state margin, where older workers face the more persistent damage. Both facts can be true simultaneously: the early displacement flow can be young-skewed while the terminal stock of permanent exit remains old-skewed, because conditional adjustment stays harder for older workers throughout. We formalize this possibility in the extension of Section 5, which allows the age composition of the displacement flow to rotate over the adoption path without altering the optimal-policy results of the baseline model.

The OECD Employment Outlook 2025, published as we were completing this paper, provides extensive evidence consistent with the mechanism.[2] It documents that older workers are substantially less likely to participate in training (one-third of 60–65 year-olds versus over half of 25–44 year-olds), that job-to-job mobility declines sharply with age (37% of workers 55–74 change status annually versus 54% of 15–54 year-olds), and that demographic aging reduced worker mobility by 1.6 percentage points between 2000 and 2022. These patterns are the empirical counterparts of the composition deterioration and reduced throughput that the model formalizes.

Card, Kluve, and Weber (2018) and Crépon et al. (2013) provide meta-analytic evidence on ALMP effectiveness that grounds the calibration of the retraining hazard $\rho(Z)$. Autor and Dorn (2013) and Autor, Levy, and Murnane (2003) document the routine-task displacement patterns that motivate the production structure.

A methodological note: Proposition $4'$ compares adoption episodes that begin at different demographic states; it is not a dynamic stopping problem. Proposition 4 is the main within-episode policy result, while Proposition $4'$ characterizes a timing boundary between waiting for a younger labor force and accepting the congestion cost of speed.

[2] Training participation: Ch. 4 (Skills), Figure 4.10. Job-to-job mobility: Ch. 5 (Job mobility), Section 5.2. Aging and mobility decline: Ch. 5, Figure 5.5.

The paper proceeds as follows. Section 2 presents the model. Section 3 states the main results and their economic interpretation; formal proofs are collected in Appendices B–D. Section 4 summarizes comparative statics. Section 5 presents an extension with age-specific incidence of displacement. Section 6 develops empirical implications and a measurement strategy. Section 7 concludes.

2 Model

2.1 Environment

The model is a reduced-form transition economy with three primitives: an automation diffusion path, a congestion-sensitive retraining pipeline, and a wage-mediated discouragement margin for displaced workers. The homogeneous version of this environment is analyzed in Levy Yeyati (2026); here the emphasis is on age heterogeneity in adjustment. Automation diffuses logistically across routine tasks, generating the displacement flow:

$$\Phi(t;\kappa) = \bar{\imath}\,\kappa\, a(t)(1-a(t)), \qquad \Phi_{\max} = \frac{\bar{\imath}\kappa}{4} \tag{1}$$

where $a(t)$ follows the logistic diffusion path

$$a(t) = \frac{1}{1+\left(\frac{1-a_0}{a_0}\right)e^{-\kappa t}}, \qquad a(0) = a_0 \in (0,1), \tag{2}$$

with a_0 the initial automation share and $\kappa > 0$ the adoption speed. The non-routine wage clears at:

$$w^N(t) = w_0^N \cdot \left[\frac{1-\bar{\imath}}{L^N(t)}\right]^{1/\sigma} \tag{3}$$

where w_0^N is the initial non-routine wage (normalized to 1), $L^N(t) = (1-\bar{\imath}) + R(t)$ with $R(t)$ cumulative retraining, and $\sigma > 0$ is the CES task substitution elasticity. Aggregate output $Y(t;\kappa)$ is determined by a CES task aggregator over routine and non-routine production; it depends on adoption speed through the automation share $a(t)$ and on $L^N(t)$ through non-routine task intensity. For the purposes of this paper, the wage and output equations are taken as reduced-form equilibrium objects of a task-based economy. The results below require only their monotonicity properties.

Remark 1 (Limiting case)**.** *If $\lambda_0 = 1$ and $g_\lambda = 0$, all displaced workers belong to a single pool and the model collapses to a homogeneous transition economy. The analysis below focuses on departures from that benchmark generated by demographic composition.*

2.2 Worker Types

The routine displacement pool contains two types: young (Y) and old (O). Total displacement splits as:

$$\Phi^y(t;\kappa) = \lambda(t)\cdot\Phi(t;\kappa), \qquad \Phi^o(t;\kappa) = (1-\lambda(t))\cdot\Phi(t;\kappa), \tag{4}$$

where the young share follows:

$$\dot{\lambda}(t) = -g_\lambda\,\lambda(t), \quad \lambda(0)=\lambda_0\in(0,1), \quad g_\lambda>0, \tag{5}$$

so $\lambda(t)=\lambda_0 e^{-g_\lambda t}$ is strictly decreasing. When $\lambda_0 = 1$ and $g_\lambda = 0$, the economy reduces to a homogeneous displacement pool.

2.3 Type-Specific Hazard Parameters

The institutional retraining hazard $\rho(Z)$ is common. The effective completion rate differs:

$$\rho^y(Z) = \rho(Z), \qquad \rho^o(Z) = \theta\,\rho(Z), \quad \theta\in(0,1). \tag{6}$$

The discouragement hazard takes a linear form with type-specific parameters:

$$\mu^y(q,w^N) = \max\bigl(0,\ \mu_0+\alpha_1 q-\alpha_2 w^N\bigr), \tag{7}$$

$$\mu^o(q,w^N) = \max\bigl(0,\ \mu_0+\Delta\mu+\alpha_1 q-\alpha_2\gamma w^N\bigr), \quad \Delta\mu>0,\ \gamma\in(0,1). \tag{8}$$

2.4 Flow Equations

The state space expands to six stocks: $U^y, U^o, R^y, R^o, D^y, D^o$. The congestion index pools both types:

$$q(t) = \frac{\Phi(t;\kappa)}{\rho(Z)\cdot\max\{U^y(t)+U^o(t),\ \varepsilon\}} \tag{9}$$

where $\varepsilon>0$ is a small numerical floor preventing mechanical singularity when $U\to 0$ at the end of the transition.

Remark 2 (Utilization vs. queue length)**.** *The congestion index $q(t)\equiv\Phi(t;\kappa)/(\rho(Z)\cdot U(t))$ is a utilization measure: $q(t)>1$ indicates that displacement inflow exceeds retraining throughput.*[3] *The welfare-relevant congestion occurs in the high-inflow phase around the displacement peak, where an increase in κ raises $\Phi(t;\kappa)$ sharply while $U(t)$ has not yet accumulated proportionally. Both worker types observe the same q and update their*

[3] Because $q(t)$ is flow-over-stock (scaled by capacity), it need not be monotone in the unemployment stock: as the transition matures, $\Phi(t;\kappa)$ declines while $U(t)$ may remain elevated, so $q(t)$ can fall as the queue drains.

exit decision through the discouragement hazards (7)–(8). The age-differential effect operates through a separate channel: old workers' lower effective retraining rate $\theta\rho < \rho$ means they spend longer in U for any given q, accumulating more exposure to the discouragement hazard. The congestion index is common; the exposure duration is type-specific.

The ODEs are:

$$\dot{U}^y = \Phi^y(t;\kappa) - \big[\rho(Z) + \mu^y\big(q(t), w^N(t)\big)\big]U^y(t), \tag{10}$$
$$\dot{U}^o = \Phi^o(t;\kappa) - \big[\theta\rho(Z) + \mu^o\big(q(t), w^N(t)\big)\big]U^o(t), \tag{11}$$
$$\dot{R}^j = \rho^j(Z)\, U^j(t), \qquad j \in \{y, o\}, \tag{12}$$
$$\dot{D}^j = \mu^j\big(q, w^N\big)\, U^j(t), \qquad j \in \{y, o\}. \tag{13}$$

2.5 Welfare

The social welfare function assigns type-specific costs:

$$W(\kappa) = \int_0^\infty e^{-\beta t}\Big[Y(t;\kappa) - \chi^U U(t;\kappa) - \chi_D^y D^y(t;\kappa) - \chi_D^o D^o(t;\kappa)\Big]dt \tag{14}$$

where $U(t) = U^y(t) + U^o(t)$, $\chi^U > 0$ is the common flow disutility of unemployment, and $\chi_D^o > \chi_D^y > \chi^U$ is the irreversibility premium, strictly higher for old workers.

Figure 1 illustrates the transition dynamics implied by the model. Faster adoption compresses displacement into a shorter interval, generating a higher unemployment spike and a larger permanent-exit mass when retraining capacity is limited.

3 Main Results

The interaction between type-specific hazards and the shared congestion index generates the main results, stated here in their economically transparent form. To keep the exposition readable, the algebra, comparison arguments, and intermediate steps are collected in Appendices B–C.

Proposition 1 (Age-Incidence of Permanent Exit)**.** *For any $\kappa > 0$, $g_\lambda \geq 0$, $\Delta\mu > 0$, and $\theta < 1$:*

(i) The old exit share $D^o(\infty)/D(\infty)$ is strictly greater than the old displacement share among the displaced mass: old workers are over-represented in permanent exit.

(ii) Under SC-I and congestion monotonicity on the high-inflow interval, the ratio $D^o(\infty)/D^y(\infty)$ is strictly increasing in κ: faster adoption disproportionately concentrates permanent exit among old workers.

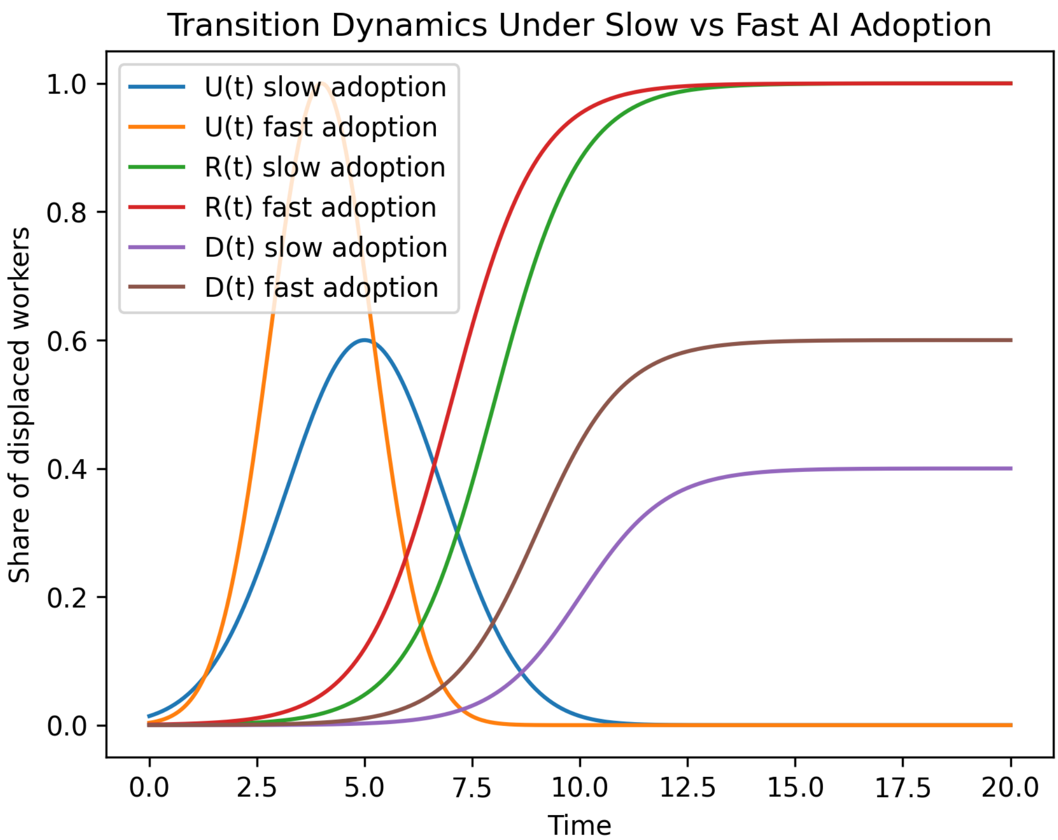

Figure 1: Transition dynamics under slow ($\kappa = 0.15$) versus fast ($\kappa = 0.60$) AI adoption. $U(t)$, $R(t)$, and $D(t)$ are expressed as shares of the automatable task mass $\bar{\imath}$, so $U(t) + R(t) + D(t) \le \bar{\imath}$ at all times. Faster adoption compresses displacement into a shorter window, generating a higher unemployment spike and greater permanent labor force exit when retraining capacity is limited. Parameters as in Section 3.1.

Proof. See Appendix B. The proof combines a competing-risks representation for type-specific transition probabilities with a comparison argument over the high-inflow phase. □

Proposition 2 (Supermodularity of Age–Exit Incidence). *The age-differential permanent exit $D^o(\infty) - D^y(\infty)$ is strictly supermodular in $(\kappa, 1 - \lambda_0)$ under the following sufficient condition:*

SC-I (Kink condition). *There exists a set of times $T^* \subset [0, \infty)$ of positive Lebesgue measure on which $\mu^y(q(t), w^N(t)) = 0$ while $\mu^o(q(t), w^N(t)) > 0$.*

Proof. See Appendix B. The key step is that on the positive-measure set T^*, the young type is at the corner of the discouragement operator while the old type remains interior, so the composition effect and the speed effect reinforce each other strictly. □

Remark 3 (Empirically relevant region). *Under the calibration used below ($\theta = 0.60$, $\Delta\mu = 0.03$, $\rho = 0.20$), the kink condition SC-I is satisfied near the displacement peak* *whenever the old-worker share is non-negligible. Equivalently, SC-I holds at time t when:*

$$\mu_0 + \alpha_1 q(t) - \alpha_2 w^N(t) \leq 0 \quad \text{and} \quad \mu_0 + \Delta\mu + \alpha_1 q(t) - \alpha_2 \gamma w^N(t) > 0.$$

With $\Delta\mu > 0$ and $\gamma < 1$, the second inequality is strictly easier to satisfy than the first: old workers are interior when young workers are at the corner. This holds generically in the early transition phase before congestion pushes both hazards into the interior.

Proposition 3 (Demographic Composition Effect on Aggregate Retraining)**.** *For any $\kappa > 0$, aggregate cumulative retraining $R(\infty) = R^y(\infty) + R^o(\infty)$ is strictly decreasing in g_λ for fixed κ and $\rho(Z)$: a faster-aging displacement pool deteriorates aggregate transition outcomes without any change in institutional capacity.*

Proof. See Appendix B. The proof rewrites cumulative retraining as the inflow-weighted sum of type-specific completion probabilities and then shows that a higher g_λ shifts mass from the high-completion young type toward the low-completion old type. □

Corollary 1 (Demographic Aging Raises the Return to Reskilling Investment)**.** *Provided the welfare weight ratio satisfies $\chi^o_D/\chi^y_D > (1-\theta)/\theta$, the marginal welfare gain from reskilling capacity, $\partial W/\partial\rho(Z)$, is strictly increasing in the old share $(1-\lambda_0)$ and in the aging rate g_λ. Demographic aging and reskilling investment are therefore welfare complements whenever the irreversibility premium on old-worker exit exceeds the ratio of their retraining disadvantage.*

Proof. See Appendix B. The appendix decomposes W_ρ into type-specific exit-prevention terms and shows that, under the stated weight restriction, shifting the pool toward old workers raises the marginal welfare value of capacity. □

This is a weak restriction. Empirical estimates place $\theta \in [0.5, 0.7]$, implying $(1-\theta)/\theta \in [0.4, 1.0]$.

Remark 4 (Fiscal paradox)**.** *Corollary 1 establishes that aging raises the social return to reskilling investment. But aging simultaneously crowds out the fiscal space for $\rho(Z)$ through rising pension, healthcare, and long-term care expenditure. The countries where reskilling investment is most urgently needed are precisely the countries facing the greatest fiscal pressure on public training budgets. The OECD Employment Outlook 2025 (Ch. 2) projects that age-related spending will increase by over 6 percentage points of GDP in the average OECD country by 2060, compressing the budgets available for active labor market programs at exactly the moment the composition effect makes those programs most valuable. This paradox does not resolve within the model, but it identifies a structural distortion: the welfare loss from under-investment in $\rho(Z)$ is strictly larger in aging economies than in young ones, strengthening the case for protecting ALMP budgets from fiscal consolidation where demographics are most adverse.*

Proposition 4 (Demographic Gradient in Optimal Adoption Speed)**.** *Maintain the two-type model (Sections 2.2–2.5). Suppose:*

(A1) Type ranking: $\mu^o(q, w^N) \geq \mu^y(q, w^N)$ *for all* (q, w^N)*, with strict inequality on a set of positive measure; and* $\rho^o(Z) = \theta\rho(Z)$ *with* $\theta \in (0, 1)$*.*

(A2) Congestion monotonicity: increasing κ *weakly increases the marginal transition cost of displacement on a set of positive measure.*

(A3) Strict concavity: $W(\kappa; \lambda_0, g_\lambda)$ *is strictly concave in* κ*.*

(A4) Output-composition channel omitted for tractability: the output-gain component of $\partial W/\partial\kappa$ *is approximately independent of* λ_0*, or any residual composition dependence is weakly nonnegative.*

Under A1–A4, the planner's interior optimum satisfies $\partial\kappa^*/\partial\lambda_0 > 0$*: the planner chooses faster adoption when the displacement pool is younger.*

Note on A4. Strictly, $Y(t; \kappa)$ depends on $L^N(t) = (1-\bar{\imath}) + R(t)$, and $R(t)$ depends on composition λ_0 through the type-specific throughput rates ρ and $\theta\rho$. A4 therefore omits an output-composition term for tractability. The omitted channel works in the same direction: a younger pool raises $R(t)$ through higher throughput, which raises $L^N(t)$ and hence $Y(t; \kappa)$. Allowing Y to depend on λ_0 would add a weakly positive term to $W_{\kappa\lambda_0}$, strengthening the result. In that sense A4 is conservative.

Proof. See Appendix C.1. The appendix rewrites the marginal cost of speed as a type-weighted integral over marginal displaced mass and then applies the implicit-function theorem. Section 3.1 then reports a calibrated numerical implication: the demographic gradient becomes steeper as g_λ rises in the parameter region considered here. □

Numerical implication (steepening under faster aging). In the calibrated region, the positive relationship between κ^* and λ_0 is stronger at higher values of g_λ. Equivalently, differences in initial workforce composition matter more for the planner's optimal adoption speed when demographic aging is faster.

Figure 2 displays the welfare function $W(\kappa)$ for three values of λ_0. The leftward shift of the optimum illustrates Proposition 4; the increasing asymmetry of the curves shows that an older pool raises the welfare cost of adopting too fast relative to the cost of adopting too slow.

Proposition 4′ (Welfare Value of Earlier Adoption Under Demographic Aging)**.** *Fix the within-episode policy rule* $\kappa = \kappa^*(\lambda)$ *from Proposition 4. Compare two economies identical except that adoption begins at calendar time* τ *vs* $\tau + T$*, so composition at adoption is* λ_τ

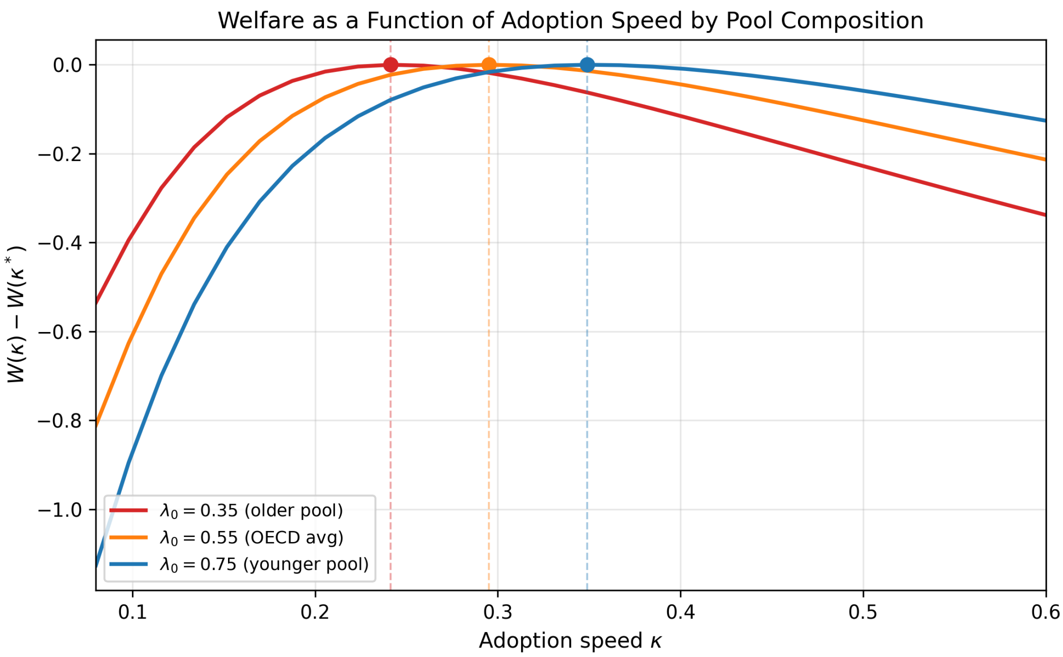

Figure 2: Welfare as a function of adoption speed for three displacement pool compositions. Dashed lines mark the interior optimum $\kappa^*(\lambda_0)$. An older pool shifts the optimum leftward and increases the welfare cost of overshooting: the red curve ($\lambda_0 = 0.35$) peaks at $\kappa^* \approx 0.24$ while the blue curve ($\lambda_0 = 0.75$) peaks at $\kappa^* \approx 0.35$. Parameters as in Section 3.1, $g_\lambda = 0.015$.

vs $\lambda_\tau e^{-g_\lambda T}$ (we set $t = 0$ at adoption, so $\lambda_0 = \lambda(\tau)$). Define $V(\lambda) = W(\kappa^*(\lambda), \lambda)$. The marginal cost of delay is:

$$\frac{dV}{d\tau} = \frac{\partial W}{\partial \lambda} \cdot \dot{\lambda} = \frac{\partial W}{\partial \lambda} \cdot (-g_\lambda \lambda),$$

where the envelope theorem eliminates the $\partial W/\partial \kappa$ term. Earlier adoption is strictly preferred—$dV/d\tau < 0$—iff $\partial W/\partial \lambda > 0$. This holds under:

$$\textbf{\textit{(SC)}} \qquad g_\lambda \cdot \Delta\mu \cdot \frac{1 - \lambda_0}{\rho + \Delta\mu} > \frac{\chi^U}{\chi^o_D} \cdot \frac{\kappa}{4\rho(Z)}. \tag{15}$$

The left side is the demographic cost of delay; the right side is the congestion cost of speed. (SC) is quantitatively tight under the calibration used below (see Section 3.1). Proposition 4 is the paper's central policy result; Proposition 4′ characterizes a theoretical boundary that becomes operative only under extreme parameterizations and is included for completeness rather than as an empirically relevant regime.

Proof. See Appendix C.2. The appendix applies the envelope theorem and derives the sufficient condition by comparing the demographic cost of delay with an upper bound on the congestion cost of speed. □

Remark 5 (Welfare gap and the non-routine scarcity premium)**.** *When permanent exit* *removes workers from the non-routine labor market, the remaining non-routine workforce is smaller, generating a persistent scarcity premium at the destination. This premium widens the welfare gap between successful retrainees and permanently discouraged workers. The gap is largest when old workers—who never capture the premium—dominate permanent exit. Standard wage inequality measures among employed workers understate the welfare gap because permanently exited workers are removed from the wage distribution, making the economy appear more equal among survivors. The measurement bias is strictly increasing in $D^o(\infty)$. A formal treatment requires specifying the non-routine workforce growth rate as a function of the transition dynamics, which we defer to future work.*

Four mechanisms drive the results: differential incidence (Proposition 1), complementarity through congestion (Proposition 2), the demographic gradient (Proposition 4), and the demographic buffer (Proposition 4′). Running through all is the composition channel: aging deteriorates aggregate retraining (Proposition 3), raises reskilling returns (Corollary 1), widens the welfare gap between survivors and permanent exiters, and numerically steepens the demographic gradient in optimal speed as g_λ rises.

Figure 3 visualizes the demographic gradient predicted by Proposition 4. Figure 4 illustrates Proposition 1(ii): faster adoption disproportionately increases permanent exit among older workers.

3.1 Illustrative Calibration

This section provides a numerical anchor. For comparability, the technology and labor-market block uses the same parameterization as a homogeneous speed–capacity benchmark, while the present paper adds type-specific transition and welfare parameters. Parameters: $\theta = 0.60$ (Dix-Carneiro, 2014; Caliendo, Dvorkin, and Parro, 2019), $\Delta\mu = 0.03$/yr (OECD Employment Outlook 2025, Ch. 3), $\rho(Z) = 0.20$, $\chi^U = 0.30$, $\chi_D^y = 1.50$, $\chi_D^o = 2.25$, giving $\chi^U/\chi_D^o = 0.133$. Three scenarios: $g_\lambda = 0.008$ (Sweden), $g_\lambda = 0.015$ (OECD average), $g_\lambda = 0.025$ (Italy/Korea).

Evaluating (SC). At the OECD average scenario ($g_\lambda = 0.015$, $\lambda_0 = 0.55$, $\kappa = 0.30$, $\rho = 0.20$, $\chi^U/\chi_D^o = 0.133$): left side $= 0.015 \times 0.030 \times 0.45/0.23 = 8.80 \times 10^{-4}$; right side $= 0.133 \times 0.30/0.80 = 0.050$. The ratio is 0.018. At $\lambda_0 = 0.25$ (Italy), the ratio rises to 0.029. For (SC) to hold, one would need $g_\lambda > 0.85$, or $\chi^U/\chi_D^o < 0.001$, or $\kappa < 0.005$—all outside the empirical range. The monotone gradient (Proposition 4) is the empirically relevant regime.

Numerical findings. First, $\kappa^* \approx 0.28$ at $\lambda_0 = 0.55$, $g_\lambda = 0.015$. A homogeneous reference case with a single exit weight $\chi^D = 1.50$ yields $\kappa^* = 0.36$; the lower optimum in the demographic model reflects the higher welfare cost attached to old-worker permanent exit ($\chi_D^o = 2.25 > \chi_D^y = 1.50$), which makes the planner more cautious about speed.

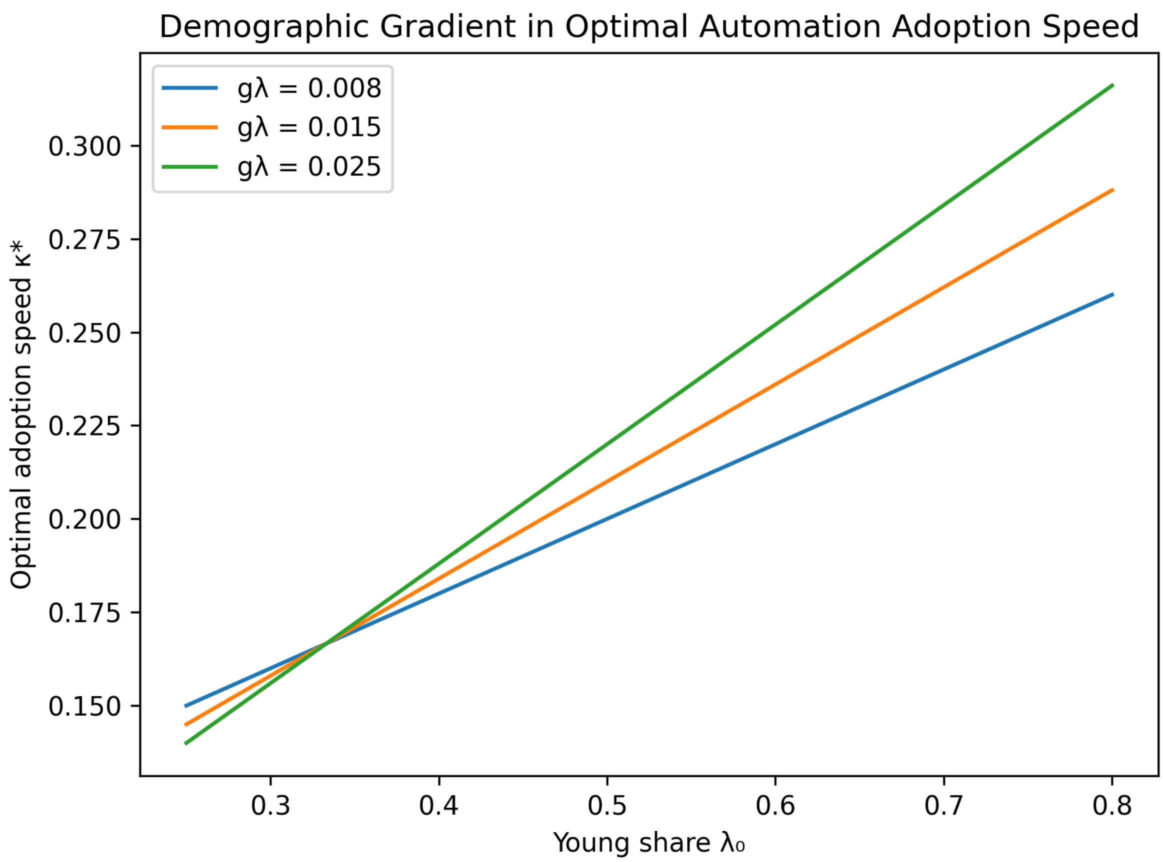

Figure 3: Demographic gradient in optimal adoption speed $\kappa^*(\lambda_0)$. The schedule is upward sloping for all three aging rates used in the calibration ($g_\lambda = 0.008, 0.015, 0.025$), consistent with Proposition 4. In the calibrated region, the slope is steeper at higher aging rates. Calibration: $\bar{\iota} = 0.35$, $\rho = 0.20$, $\sigma = 1.5$, $\mu_0 = 0.02$, $\alpha_1 = 0.12$, $\alpha_2 = 0.15$, $\Delta\mu = 0.03$, $\theta = 0.60$, $\gamma = 0.75$.

Second, the $\kappa^*(\lambda_0)$ gradient is positive throughout: κ^* rises from 0.154 at $\lambda_0 = 0.30$ to 0.263 at $\lambda_0 = 0.75$ (gradient$\approx$ 0.24). Third, the gradient steepens with g_λ: 0.38 at $g_\lambda = 0.025$ versus 0.22 at $g_\lambda = 0.008$.

4 Comparative Statics

The main comparative statics are as follows. Holding the adoption path fixed, faster exogenous adoption (higher κ) raises congestion and increases the old-worker share of permanent exit, D^o/D. Greater retraining capacity, $\rho(Z)$, raises the planner's optimal adoption speed κ^* and compresses age gaps in permanent exit. A higher welfare penalty on old-worker exit, χ_D^o, lowers κ^* because the planner internalizes the larger social cost of speed. Lower old-worker throughput (lower θ) and higher old-worker discouragement (higher $\Delta\mu$) both reduce κ^* and raise D^o/D. A higher initial young share, λ_0, increases κ^* and lowers D^o/D (Proposition 4). Finally, higher demographic aging, g_λ, raises D^o/D

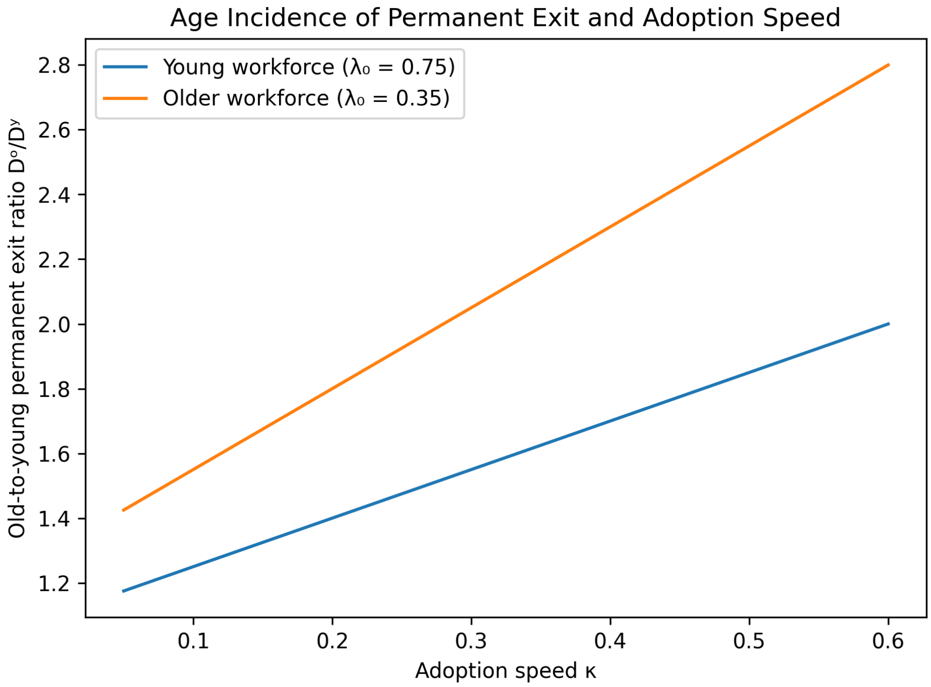

Figure 4: Age incidence of permanent exit as a function of adoption speed κ. Both curves are upward-sloping (Proposition 1(ii)); the older-workforce curve lies above and is steeper, illustrating the supermodularity of age–exit incidence (Proposition 2).

and, in the calibrated region, steepens the schedule relating κ^* to λ_0.

5 Extension: Age-Specific Incidence of Displacement

The baseline model isolates age differences in *conditional adjustment*: once workers are displaced, old workers complete retraining less often and exit the labor force more often. A natural extension allows age to shape *incidence* as well. Recent descriptive evidence from the United States suggests that early AI substitution has so far fallen more heavily on younger and less experienced workers, even though the more persistent labor-market damage need not do so (Peng, 2026). We capture that possibility in reduced form by allowing the age composition of the displacement flow to rotate over the adoption path while keeping the aggregate displacement flow $\Phi(t;\kappa)$ unchanged.

Let displacement split as:

$$\Phi^y(t;\kappa) = \lambda(t)\,\Phi(t;\kappa)\,\frac{1+\varphi\,\eta(t)}{Z(t)}, \qquad \Phi^o(t;\kappa) = (1-\lambda(t))\,\Phi(t;\kappa)\,\frac{1}{Z(t)}, \tag{16}$$

where

$$\eta(t) = 1 - 2a(t), \qquad 0 \le \varphi < 1, \tag{17}$$

and the normalizer

$$Z(t) = 1 + \varphi\,\eta(t)\,\lambda(t) \tag{18}$$

ensures $\Phi^y(t;\kappa) + \Phi^o(t;\kappa) = \Phi(t;\kappa)$ for every t. The parameter φ controls the intensity of front-loading toward younger workers. The restriction $0 \leq \varphi < 1$ guarantees that the normalizer $Z(t) = 1 + \varphi\eta(t)\lambda(t)$ remains strictly positive for all t, and therefore that both type-specific inflows $\Phi^y(t;\kappa)$ and $\Phi^o(t;\kappa)$ are well-defined and nonnegative. The binding case is $\eta(t) = -1$ (late-transition extreme) combined with $\lambda(t) = 1$, which yields $Z(t) = 1 - \varphi$; any $\varphi < 1$ keeps this strictly positive. When $\varphi = 0$, the extension collapses to the baseline specification in (4).

The kernel $\eta(t)$ is positive early in the transition, zero at the logistic inflection point, and negative late in the transition. Economically, the extension is a reduced-form way to capture the possibility that early AI codification disproportionately displaces junior or less experienced workers, while later diffusion reaches more experience-intensive routine tasks. The rotation point is therefore pinned down by the adoption path rather than chosen freely.

Proposition 5 (Young flow, old stock)**.** *Consider the extension* (16)–(18) *with* $0 \leq \varphi < 1$*.*

(i) If $\varphi > 0$*, there exists a unique rotation point*

$$t^* = \frac{\ln\big((1-a_0)/(a_0)\big)}{\kappa}$$

such that the young share of the displacement flow exceeds the young share of the underlying exposed pool for $t < t^*$*, coincides with it at* $t = t^*$*, and falls below it for* $t > t^*$*:*

$$\frac{\Phi^y(t;\kappa)}{\Phi(t;\kappa)} \gtreqless \lambda(t) \quad\Longleftrightarrow\quad t \lesseqgtr t^*.$$

(ii) The ranking of conditional adjustment outcomes is unchanged. Given the type asymmetries $\rho^o(Z) = \theta\rho(Z) < \rho^y(Z)$ *and* $\mu^o(q, w^N) \geq \mu^y(q, w^N)$*, old workers remain less likely to complete retraining and more likely to end in permanent exit conditional on displacement.*

Hence the model can generate a young-skewed early displacement flow together with an old-skewed terminal permanent-exit stock.

Proof. See Appendix D. Part (i) follows directly from the sign of $\eta(t)$. Part (ii) is inherited from the baseline type asymmetries because the extension reweights the age composition of inflows without altering the conditional hazard ranking once workers enter the adjustment pipeline. □

Remark 6 (The extension preserves the direction of the policy result)**.** *Because the extension reweights the age composition of displacement inflows without altering the type-specific hazards $\rho^j(Z)$ and $\mu^j(q, w^N)$, the conditional adjustment ranking that drives Propositions 1–3 is preserved pointwise. As a result, the demographic gradient in optimal adoption speed (Proposition 4) is not overturned by the front-loading channel: the sign-determining step in the proof—that $c_o(t) > c_y(t)$ on a set of positive measure and that $\psi(t;\kappa)$ is positive in the high-inflow phase—is unaffected by the reweighting, so $W_{\kappa\lambda_0} > 0$ continues to hold under the extension. The magnitude of the gradient may shift through the timing of displacement, but its sign does not. The extension therefore accommodates the young-flow evidence while leaving the paper's central policy logic intact.*

This extension is deliberately secondary to the paper's core mechanism. It does not impose the claim that older workers are always less exposed to substitution. It only allows incidence to vary by age over the transition. The main policy result of the paper—that older displacement pools imply a lower socially optimal adoption speed because they are harder to absorb—continues to operate on the *conditional adjustment* margin.

6 Empirical Implications and Measurement Strategy

The goal of this section is not to claim a full structural empirical implementation of the model, but to map its main objects into observable counterparts and to state reduced-form predictions that could discipline the mechanism. The model generates six cross-country predictions.

P1. Age-incidence. Countries with higher AI adoption speeds κ should exhibit larger D^o/D^y ratios. Testable with age-specific LFP trends conditional on automation exposure.

P2. Supermodularity. The interaction $\kappa \times (1 - \lambda_0)$ should have a strictly positive coefficient on old-worker permanent exit. Testable with a Bartik-style IV for κ.

P3. Composition deterioration. Countries with faster demographic aging g_λ should exhibit slower aggregate retraining rates for given $\rho(Z)$ and κ.

P4. Demographic gradient. Countries that adopted AI when their routine workforce was younger should exhibit lower rates of permanent labor force exit, conditional on κ and $\rho(Z)$.

P5. Measurement bias. Standard wage Gini among employed workers should be negatively correlated with $\kappa \times (1 - \lambda_0)$. Identification is challenging here because standard automation channels (polarization, routine-wage compression) also affect the Gini; the selection channel predicted by the model—removal of permanent exiters from the wage distribution—must be isolated from these level effects, for instance through panel variation in age-specific non-participation rates.

P6. Reskilling return. The marginal welfare return to ALMP spending should be strictly higher in countries with older workforces. The fiscal paradox predicts a negative interaction between aging and ALMP budget shares.

6.1 Measurement and Identification

Adoption speed κ. We proxy κ with the industry×country growth rate of AI/robot adoption, measured either through robot installation rates (IFR World Robotics database) or through exposure-weighted AI adoption diffusion slopes from firm-level surveys (McKinsey Global AI Survey, Stanford HAI AI Index). The logistic structure (1)–(2) implies that κ can be estimated from the slope of adoption S-curves fitted to industry-level data. An alternative uses a Bartik-style instrument: pre-existing industry composition interacted with global AI adoption trends.

Routine-exposed pool composition λ_0. The young share in the displacement pool is the share of workers aged 25–54 (or 25–49) among routine-task occupations, using PIAAC skill assessments mapped to ISCO codes following Autor and Dorn (2013). Cross-country variation in λ_0 comes from differential aging within routine-intensive industries.

Permanent exit D. The empirical counterpart of the absorbing state D is sustained labor force non-participation following displacement from a routine occupation. We define permanent exit as a transition from routine employment to non-participation lasting at least 24 months, excluding retirement at statutory age. This is measured in longitudinal LFS microdata (EU-SILC, CPS-ASEC, or national LFS panels). Disability insurance entry provides a complementary measure.

7 Conclusion

This paper studies a question that becomes increasingly important as routine-intensive workforces age: how does the age composition of the displacement pool change the socially desirable speed of automation adoption? The central result is that it does. For a given retraining capacity, an older displacement pool generates more discouragement, lower completion, and more permanent labor force exit. Aging therefore changes not only the distribution of transition costs across workers but the aggregate transition itself.

The mechanism is a demographic composition effect. A common displacement flow enters a shared retraining pipeline, but young and old workers face different effective hazards once inside it. Older workers have lower completion rates and higher exit rates, so a shift in the composition of displaced workers worsens outcomes even when institutions are unchanged. When adoption is fast, this composition effect becomes stronger because congestion amplifies the asymmetry. Fast adoption in an older pool is therefore worse than the sum of fast adoption and aging taken separately.

The results can be summarized simply. Permanent exit is disproportionately concentrated among older workers. The age-exit differential is supermodular in adoption speed and the old-worker share. Aggregate retraining deteriorates as the displacement pool ages, even with fixed institutional capacity. The marginal welfare return to reskilling rises with aging. And the planner's optimal adoption speed is increasing in the young share of the pool. Numerically, in the calibrated region, this demographic gradient is steeper when aging is faster. The timing boundary remains theoretically possible but, under the calibration considered here, quantitatively remote; the empirically relevant case is the monotone demographic gradient.

These results imply a simple policy lesson. The question is not only whether economies should automate more or less, but whether they can absorb a given automation wave at the speed firms choose. In younger economies, the transition can tolerate more speed. In older economies, the same speed produces more permanent exit and a larger welfare loss. The case for reskilling investment is correspondingly stronger in aging labor markets, even though those are often the economies in which fiscal pressure is greatest.

Once the displacement pool becomes heterogeneous by age, new objects emerge that do not arise in an age-neutral transition model: age incidence, supermodularity between speed and composition, composition deterioration over time, and a demographic gradient in optimal adoption speed. Those objects are the paper's distinct contribution. Put simply, the paper asks how the cost of a given automation wave changes with the age structure of the workers who must pass through the transition.

An extension in Section 5 allows age to shape incidence as well as recovery: the displacement flow can be young-skewed early in the transition while the terminal stock of permanent exit remains old-skewed because conditional adjustment stays harder for older workers. Two further extensions are deferred. First, endogenous retirement: if older workers can choose between job search and early retirement, μ^o becomes partly choice-driven rather than purely reduced-form. Second, endogenous non-routine workforce growth: if the scarcity premium raises returns to non-routine education, task creation responds to the transition itself.

References

Abeliansky, A.L. and K. Prettner (2023). "Automation and Demographic Change." *Journal of Macroeconomics*, 77, 103540.

Acemoglu, D. and P. Restrepo (2018). "The Race between Man and Machine." *American Economic Review*, 108(6), 1488–1542.

Acemoglu, D. and P. Restrepo (2019). "Automation and New Tasks." *Journal of Economic Perspectives*, 33(2), 3–30.

Acemoglu, D. and P. Restrepo (2022). "Demographics and Automation." *Review of Economic Studies*, 89(1), 1–44.

Autor, D.H. and D. Dorn (2013). "The Growth of Low-Skill Service Jobs." *American Economic Review*, 103(5), 1553–1597.

Autor, D.H., F. Levy, and R.J. Murnane (2003). "The Skill Content of Recent Technological Change." *Quarterly Journal of Economics*, 118(4), 1279–1333.

Caliendo, L., M. Dvorkin, and F. Parro (2019). "Trade and Labor Market Dynamics." *Econometrica*, 87(3), 741–835.

Card, D., J. Kluve, and A. Weber (2018). "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations." *JEEA*, 16(3), 894–931.

Carneiro, P. and J.J. Heckman (2003). "Human Capital Policy." In J. Heckman and A. Krueger (eds.), *Inequality in America: What Role for Human Capital Policy?*, MIT Press, 77–240.

Crépon, B., E. Duflo, M. Gurgand, R. Rathelot, and P. Zamora (2013). "Do Labor Market Policies Have Displacement Effects?" *Quarterly Journal of Economics*, 128(2), 531–580.

Dix-Carneiro, R. (2014). "Trade Liberalization and Labor Market Dynamics." *Econometrica*, 82(3), 825–885.

Farber, H.S. (2017). "Employment, Hours, and Earnings Consequences of Job Loss: US Evidence from the Displaced Workers Survey." *Journal of Labor Economics*, 35(S1), S235–S272.

Heckman, J.J. (2006). "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science*, 312(5782), 1900–1902.

Jacobson, L.S., R.J. LaLonde, and D.G. Sullivan (1993). "Earnings Losses of Displaced Workers." *American Economic Review*, 83(4), 685–709.

Johnson, R.W. and P. Gosselin (2018). "How Secure Is Employment at Older Ages?" Urban Institute Report.

Kubeck, J.E., N.D. Delp, T.K. Haslett, and M.A. McDaniel (1996). "Does Job-Related Training Performance Decline with Age?" *Psychology and Aging*, 11(1), 92–107.

Levy Yeyati, E. (2026). "Too Fast to Adjust? AI Adoption Speed, Labor Market Transition, and Optimal Policy." Working Paper, Universidad Torcuato Di Tella.

OECD (2025). *OECD Employment Outlook 2025*. OECD Publishing, Paris.

Prettner, K. and H. Strulik (2020). "Innovation, Automation, and Inequality." *Journal of Monetary Economics*, 116, 249–265.

Peng, E. (2026). "Breaking Down the Impact of AI on the Labor Market: Substitution vs. Augmentation." *Goldman Sachs US Daily*, April 4.

Appendices

A Calibration Parameters

For comparability, the technology and transition block is aligned with a homogeneous speed–capacity benchmark, while the present paper adds type-specific demographic and welfare parameters.

Core parameters: $\bar{\iota} = 0.35$, $\rho(Z) = 0.20$, $\sigma = 1.5$, $\mu_0 = 0.02$, $\alpha_1 = 0.12$, $\alpha_2 = 0.15$, $\chi^U = 0.30$, $\beta = 0.04$.

Demographic parameters: $\Delta\mu = 0.03/\text{yr}$, $\theta = 0.60$, $\gamma = 0.75$, $\chi_D^y = 1.50$, $\chi_D^o = 2.25$.

Scenarios: $g_\lambda \in \{0.008, 0.015, 0.025\}$; $\lambda_0 = 0.55$ in the benchmark exercise.

Key derived quantities: $\kappa^* \approx 0.28$ at $\lambda_0 = 0.55$, $g_\lambda = 0.015$. A homogeneous reference case with $\chi^D = 1.50$ yields $\kappa^* = 0.36$; the lower optimum in the demographic model reflects $\chi_D^o = 2.25 > \chi_D^y = 1.50$, making the planner more cautious. (SC) ratio $= 0.018$ at the OECD-average scenario. $\partial\kappa^*/\partial\lambda_0 \approx 0.24$ at $g_\lambda = 0.015$; ≈ 0.38 at $g_\lambda = 0.025$.

B Technical Proofs for Propositions 1–3 and Corollary 1

B.1 Well-posedness and boundedness

Lemma 1 (Well-posed transition system)**.** *For any admissible parameter vector and any $\kappa > 0$, the six-dimensional system (10)–(13), together with (9), admits a unique global non-negative solution.*

Proof. The inflow path $\Phi(t;\kappa)$ is continuous, the hazards (7)–(8) are piecewise linear and globally Lipschitz in (q, w^N), and the floor $\varepsilon > 0$ in (9) keeps the congestion index bounded away from a mechanical singularity. Standard ODE existence and uniqueness therefore apply. Non-negativity follows because each state equation has the form $\dot{x} = f(t) - g(t)x$ with $f(t) \geq 0$ and $g(t) \geq 0$. Summing the six state equations gives

$$\frac{d}{dt}\big(U^y + U^o + R^y + R^o + D^y + D^o\big) = \Phi(t;\kappa),$$

so total transition mass at time t equals $\int_0^t \Phi(s;\kappa)ds \leq \bar{\iota}(1 - a_0)$. Hence all stocks remain bounded globally. □

B.2 Proof of Proposition 1

For a worker of type $j \in \{y, o\}$ displaced at time s, define the eventual completion probability

$$P_R^j(s) = \int_s^\infty \rho^j(Z) \exp\Big[-\int_s^t \big(\rho^j(Z) + \mu^j(u)\big) du\Big]\, dt,$$

and the eventual permanent-exit probability $P_D^j(s) = 1 - P_R^j(s)$. The completion functional is increasing in the throughput parameter and decreasing in the discouragement path; the exit functional moves in the opposite direction. Since $\rho^o(Z) = \theta\rho(Z) < \rho(Z) = \rho^y(Z)$ and $\mu^o(t) \geq \mu^y(t)$ for every t, with strict inequality on a set of positive measure, it follows that

$$P_D^o(s) > P_D^y(s) \qquad \text{and} \qquad P_R^o(s) < P_R^y(s)$$

for every displacement date s.

Aggregate permanent exit can therefore be written as

$$D^j(\infty) = \int_0^\infty \Phi^j(s;\kappa) P_D^j(s) ds.$$

For part (i), let $M^j = \int_0^\infty \Phi^j(s;\kappa)ds$ denote cumulative displacement by type and set $\pi^j = D^j(\infty)/M^j$. Since $\pi^o > \pi^y$,

$$\frac{D^o(\infty)}{D(\infty)} = \frac{\pi^o M^o}{\pi^o M^o + \pi^y M^y} > \frac{M^o}{M^o + M^y},$$

so old workers are over-represented in permanent exit relative to their share in the displaced mass.

For part (ii), differentiate the exit flows with respect to κ on the high-inflow interval. Under congestion monotonicity, $\partial q/\partial\kappa > 0$ on a set of positive measure. Under SC-I, there exists T^* of positive measure on which $\mu^y = 0$ while $\mu^o > 0$, so

$$\frac{\partial \mu^y}{\partial q} = 0, \qquad \frac{\partial \mu^o}{\partial q} = \alpha_1 > 0 \qquad \text{on } T^*.$$

Hence a marginal increase in κ raises the old exit hazard strictly more than the young exit hazard on T^*, while old workers also spend longer in the pipeline because $\theta < 1$. Integrating over T^* yields $\partial(D^o/D^y)/\partial\kappa > 0$.

B.3 Proof of Proposition 2

Write the age-exit gap as

$$G(\kappa, 1-\lambda_0) = D^o(\infty) - D^y(\infty) = \int_0^\infty h(t;\kappa, 1-\lambda_0)dt, \qquad h(t) = \mu^o(t)U^o(t) - \mu^y(t)U^y(t).$$

On T^*, the young type is at the corner of the discouragement operator and the old type is interior. Therefore, on T^*,

$$\frac{\partial h}{\partial \kappa} = \alpha_1 U^o \frac{\partial q}{\partial \kappa} + \mu^o \frac{\partial U^o}{\partial \kappa}, \qquad \frac{\partial h}{\partial (1-\lambda_0)} = \mu^o \frac{\partial U^o}{\partial (1-\lambda_0)},$$

with no offsetting young-type term. Since $\partial q/\partial \kappa > 0$ on the high-inflow interval and

$$\frac{\partial \Phi^o(t;\kappa)}{\partial (1-\lambda_0)} = e^{-g_\lambda t}\Phi(t;\kappa) > 0,$$

increasing the old share raises U^o and therefore magnifies the marginal effect of κ on old-worker exit. The integrand cross-partial is strictly positive on T^*, a set of positive measure, so

$$\frac{\partial^2 G}{\partial \kappa\, \partial (1-\lambda_0)} > 0.$$

That is, the age-exit gap is strictly supermodular in speed and old-worker share.

B.4 Proof of Proposition 3

Cumulative retraining can be written as the inflow-weighted sum of type-specific completion probabilities:

$$R(\infty) = \int_0^\infty \Phi^y(s;\kappa) P_R^y(s) ds + \int_0^\infty \Phi^o(s;\kappa) P_R^o(s) ds.$$

From the ordering above, $P_R^y(s) > P_R^o(s)$ for every s. The young inflow is

$$\Phi^y(s;\kappa) = \lambda_0 e^{-g_\lambda s}\Phi(s;\kappa),$$

so $\partial \Phi^y/\partial g_\lambda < 0$ for every $s > 0$, while $\partial \Phi^o/\partial g_\lambda > 0$ because total inflow Φ is independent of g_λ. Thus a higher g_λ shifts displacement mass from the high-completion young type to the low-completion old type, holding institutional capacity fixed. Differentiating the expression for $R(\infty)$ with respect to g_λ gives

$$\frac{\partial R(\infty)}{\partial g_\lambda} = \int_0^\infty \frac{\partial \Phi^y}{\partial g_\lambda}(P_R^y - P_R^o) ds < 0.$$

B.5 Proof of Corollary 1

Differentiate welfare with respect to capacity at fixed κ:

$$W_\rho = -\chi^U U_\rho(\infty) - \chi_D^y D_\rho^y(\infty) - \chi_D^o D_\rho^o(\infty).$$

An increase in ρ lowers permanent exit for both types, but the effective gain is type-specific because old-worker completion rises only at rate θ. Let

$$M_y \equiv -D^y_\rho(\infty) > 0, \qquad M_o \equiv -D^o_\rho(\infty) > 0$$

be the exit-prevention effect of additional capacity. A shift toward old workers raises the weight on M_o and lowers the weight on M_y. The net welfare effect of that reweighting is positive whenever the welfare gain from preventing old-worker exit, scaled by their lower completion efficiency, exceeds the foregone gain on the young margin, i.e.

$$\chi^o_D \theta > \chi^y_D (1-\theta) \quad \Longleftrightarrow \quad \frac{\chi^o_D}{\chi^y_D} > \frac{1-\theta}{\theta}.$$

Under this restriction, $W_{\rho,(1-\lambda_0)} > 0$. The same argument applies to g_λ: faster aging shifts more of the displaced pool toward the high-value old margin, so $W_{\rho,g_\lambda} > 0$ as well.

C Optimal Speed and Timing Proofs

C.1 Proof of Proposition 4

Let $\psi(t;\kappa) \equiv \partial\Phi(t;\kappa)/\partial\kappa$ denote the marginal displaced mass generated by a marginal increase in speed. Under A4, the output side of W_κ can be treated as independent of λ_0, so the relevant object is the marginal transition cost of that extra displaced mass. Let $c_y(t)$ and $c_o(t)$ denote the present-value marginal transition costs of an additional young or old displaced worker at date t, inclusive of unemployment and permanent-exit costs. By A1 and the welfare ordering $\chi^o_D > \chi^y_D$, one has

$$c_o(t) > c_y(t)$$

on a set of positive measure.

Because the type shares of marginal displacement are $\lambda(t) = \lambda_0 e^{-g_\lambda t}$ and $1-\lambda(t)$, the marginal value of speed can be written as

$$W_\kappa = MB_\kappa(\kappa) - \int_0^\infty e^{-\beta t}\big[\lambda(t)c_y(t) + (1-\lambda(t))c_o(t)\big]\psi(t;\kappa)\,dt,$$

where $MB_\kappa(\kappa)$ is the marginal output gain from faster adoption. Differentiating with respect to λ_0 gives

$$W_{\kappa\lambda_0} = \int_0^\infty e^{-\beta t} e^{-g_\lambda t}\big[c_o(t) - c_y(t)\big]\psi(t;\kappa)\,dt.$$

By A2, $\psi(t;\kappa)$ is positive on a set of positive measure in the high-inflow phase, and by

A1 the bracketed term is positive on a set of positive measure. Therefore $W_{\kappa\lambda_0} > 0$.

At an interior optimum $\kappa^*(\lambda_0, g_\lambda)$, the first-order condition is $W_\kappa = 0$. Strict concavity implies $W_{\kappa\kappa} < 0$, so the implicit-function theorem yields

$$\frac{\partial \kappa^*}{\partial \lambda_0} = -\frac{W_{\kappa\lambda_0}}{W_{\kappa\kappa}} > 0.$$

This proves the monotone demographic gradient. In the calibrated region reported in Section 3.1, the numerical sweep shows that this gradient becomes steeper as g_λ rises because faster aging shifts old-worker mass further into the high-inflow portion of the transition. We therefore treat steepening as a calibrated numerical implication rather than as part of the analytical proposition.

C.2 Proof of Proposition 4′

Let $V(\lambda) = W(\kappa^*(\lambda), \lambda)$ and let composition evolve in calendar time according to $\dot{\lambda} = -g_\lambda \lambda$. By the envelope theorem,

$$\frac{dV}{d\tau} = W_\lambda \dot{\lambda} = -g_\lambda \lambda W_\lambda.$$

Hence earlier adoption is preferred if $W_\lambda > 0$: delaying adoption lowers welfare by allowing the displaced pool to age before the transition begins.

A sufficient condition can be obtained by lower-bounding the demographic loss from delay and upper-bounding the congestion loss from higher speed. The demographic loss from a marginal decline in the young share is at least proportional to the old-worker discouragement wedge $\Delta\mu$, the old-worker mass $(1 - \lambda_0)$, and the expected duration in the pipeline $1/(\rho + \Delta\mu)$, giving the lower bound

$$\text{demographic cost of delay} \geq g_\lambda \cdot \Delta\mu \cdot \frac{1 - \lambda_0}{\rho + \Delta\mu}.$$

The congestion cost of raising speed is bounded above by the maximal displacement spike $\Phi_{\max} = \bar{\imath}\kappa/4$ scaled by the unemployment weight relative to the old-worker exit weight, which yields

$$\text{congestion cost of speed} \leq \frac{\chi^U}{\chi_D^o} \cdot \frac{\kappa}{4\rho(Z)}.$$

Therefore a sufficient condition for $W_\lambda > 0$ is precisely

$$g_\lambda \cdot \Delta\mu \cdot \frac{1 - \lambda_0}{\rho + \Delta\mu} > \frac{\chi^U}{\chi_D^o} \cdot \frac{\kappa}{4\rho(Z)}.$$

When this inequality holds, the demographic cost of waiting dominates the congestion cost of speed and earlier adoption is welfare-improving.

D Extension: Age-Specific Incidence of Displacement

D.1 Proof of Proposition 5

Under the extension,

$$\frac{\Phi^y(t;\kappa)}{\Phi(t;\kappa)} = \lambda(t)\cdot\frac{1+\varphi\eta(t)}{1+\varphi\eta(t)\lambda(t)}.$$

The ratio exceeds $\lambda(t)$ if and only if

$$\frac{1+\varphi\eta(t)}{1+\varphi\eta(t)\lambda(t)} > 1 \quad\Longleftrightarrow\quad \varphi\eta(t)\big(1-\lambda(t)\big) > 0.$$

For $\varphi > 0$ and $\lambda(t) \in (0,1)$, this reduces to $\eta(t) > 0$. Since $\eta(t) = 1 - 2a(t)$, the sign switch occurs at $a(t^*) = 1/2$. Solving the logistic path (2) for $a(t^*) = 1/2$ yields

$$t^* = \frac{\ln\big((1-a_0)/a_0\big)}{\kappa}.$$

This proves part (i): the displacement flow is young-skewed before t^*, neutral at t^*, and old-skewed thereafter.

For part (ii), note that the front-loading channel changes only the age composition of inflows. It does not alter the type-specific hazard ranking once workers enter unemployment: $\rho^o(Z) = \theta\rho(Z) < \rho^y(Z)$ and $\mu^o(q, w^N) \geq \mu^y(q, w^N)$ continue to hold pointwise. Therefore the competing-risks argument used in Appendix B still implies a lower eventual retraining probability and a higher eventual permanent-exit probability for old workers conditional on displacement. The extension can therefore generate a young-skewed early displacement flow together with an old-skewed terminal exit stock.