A. Introduction

In May 2026, Anthropic withheld public access to Claude Mythos Preview after it uncovered over 10,000 high- or critical-severity software vulnerabilities. Meanwhile, Lin and Parker (2025) found that 52% of US workers were worried about AI's workplace impacts.

This essay will argue that a negative income tax (NIT) style income floor, financed by a suitable combination of outcome taxation with a token tax extension, is desirable in limiting inequality in pay and paid hours of work caused by AI displacement.

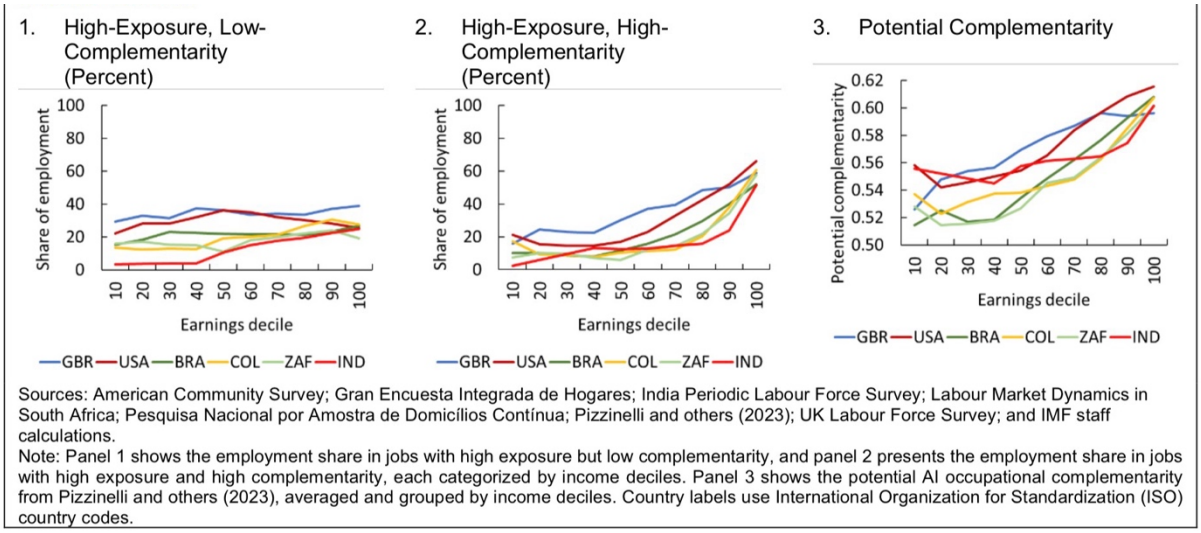
B. Background: AI and Inequality

The essay question assumes a scenario of widespread automation and subsequent displacement. Whereas the ILO finds job transformation more likely than displacement (Gmyrek et al., 2025), the scenario under examination is plausible, but not strongly supported by existing evidence (Table 1).

Study	Scope and method	Key finding on displacement
Eloundou et al. (2024)	US occupations; task-level LLM exposure ratings	Around 80% of workers have at least 10% of tasks exposed to LLMs; 19% have at least 50% exposed. Exposure is highest in higher-wage occupations.
Cazzaniga et al. (2024), IMF	125 countries; exposure combined with complementarity	Almost 40% of global employment (60% in advanced economies) is exposed; roughly half of exposed jobs are complementary, half substitutable.
Gmyrek et al. (2025), ILO	Global refined occupational exposure index	One in four jobs is exposed to generative AI, but transformation of tasks is far more likely than full automation.
Acemoglu (2025)	Task-based macro model calibrated to exposure estimates	Aggregate TFP gains of under 1% over a decade; modest effects on wages and employment, with limited scope for large-scale displacement.
Brynjolfsson, Chandar and Chen (2025)	US payroll microdata, 2022–2025	Employment of workers aged 22–25 in the most AI-exposed occupations fell about 13% relative to less-exposed peers, while older workers were unaffected.

Historically, no evidence suggests that automation has caused aggregate employment to fall over a sustained period of time, because displacement can be offset by productivity-induced expansion of labour demand and the creation of new tasks (Autor et al., 2020). Nonetheless, displacement remains possible. AI usage costs have declined rapidly, and performance has

surpassed human benchmarks in various skills. The IMF finds the share of low-complementarity, high-exposure occupations to be broadly similar across income quantiles, implying potential for mass displacement (Cazzaniga et al., 2024; Figure 1).



Sources: American Community Survey; Gran Encuesta Integrada de Hogares; India Periodic Labour Force Survey; Labour Market Dynamics in South Africa; Pesquisa Nacional por Amostra de Domicílios Contínua; Pizzinelli and others (2023); UK Labour Force Survey; and IMF staff calculations.
Note: Panel 1 shows the employment share in jobs with high exposure but low complementarity, and panel 2 presents the employment share in jobs with high exposure and high complementarity, each categorized by income deciles. Panel 3 shows the potential AI occupational complementarity from Pizzinelli and others (2023), averaged and grouped by income deciles. Country labels use International Organization for Standardization (ISO) country codes.

Although aggregate employment may recover, inequality can still persist. As estimated AI complementarity is positively correlated with income, high-paid senior workers may use AI to become more productive, benefiting from higher wages and hours. Conversely, substitutable workers may experience reduced earnings through several margins:

$$Y_i = w_i h_i$$

$$h_i = 0 \quad \text{unemployment}$$

$$h_i \downarrow \quad \text{fewer hours}$$

$$w_i \downarrow \quad \text{lower pay}$$

$$w_i' < w_i \quad \text{lower re-employment wage}$$

$$h_i' < h_i \quad \text{involuntary part-time after re-employment}$$

Thus, policy design can be decomposed into two questions:

1. How a fixed transfer budget should be allocated to maximise social welfare.
2. How the required revenue should be raised at minimum efficiency cost.

C. Redistribution via a Negative Income Tax

Job polarisation since the 1980s saw middle-skill, middle-wage workers displaced, transitioning asymmetrically into lower-paid occupations or exiting the labour force (Autor, Mindell and Reynolds, 2020; Figure 2). Transition to low-skill occupations may also widen hours inequality: workers in low-skill occupations are disproportionately under-employed, working fewer hours despite wanting more work.

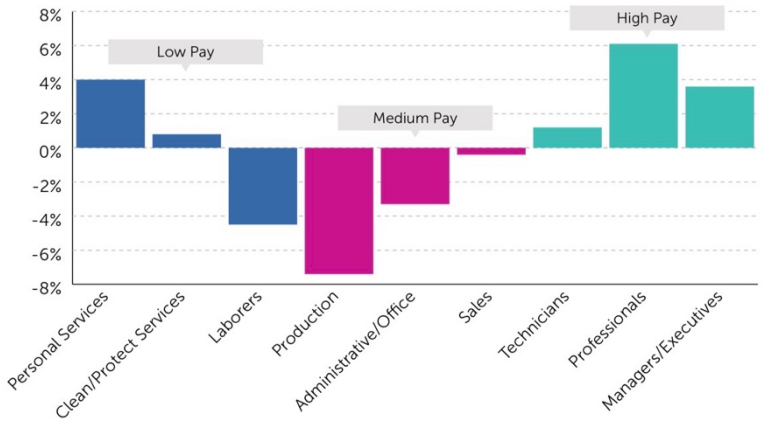


Figure is constructed using U.S. Census of Population data for 1980, 1990, and 2000, and pooled American Community Survey (ACS) data for years 2014 through 2016, sourced from IPUMS (Ruggles et al., 2018). Sample includes working-age adults ages 16 – 64 excluding those in the military. Occupational classifications are harmonized across decades using the classification scheme developed by Dorn (2009).

Thus, to limit inequality in earnings, redistribution of incomes is required. This essay considers a negative income tax (NIT) as its main transfer policy:

$$B(Y) = max\{G - rY, 0\}$$

For incomes below the breakeven threshold,

$$B'(Y) = -r < 0$$

so payments fall as earnings rise, with maximum support ($B(0) = G$) and breakeven income

$$\bar{Y} = \frac{G}{r}$$

Suppose the government allocates a fixed transfer budget R to maximise social welfare:

$$\max_{G,r} W = \sum_i \omega_i u(c_i), \quad u' > 0, \; u'' < 0$$

subject to

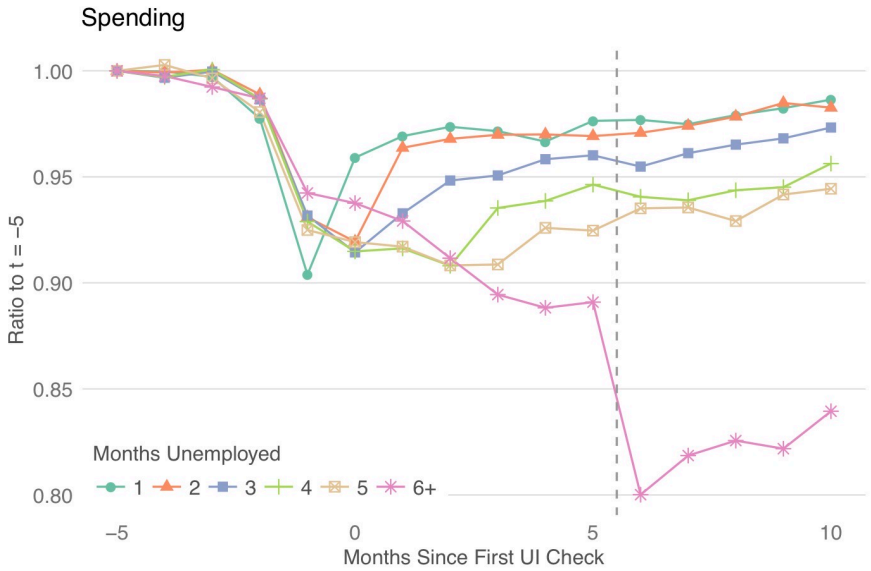
$$\sum_i B(Y_i) \leq R, \quad c_i = Y_i + B(Y_i)$$

Holding r fixed, the optimal income threshold satisfies

$$\frac{1}{N_B} \sum_{Y_i < \bar{Y}^*} \omega_i u'(c_i) = \lambda$$

where λ is the marginal social value of the transfer budget. Since $u'' < 0$, transfers to lower-income individuals generate greater marginal welfare gains. The NIT therefore concentrates a fixed redistributive budget where its social value is greatest.

Furthermore, AI displacement can reach more educated, white-collar entry-level workers. Although these workers may possess the human capital to transition into new AI-complementary jobs, imperfect unemployment insurance (UI) can create more lower-quality matches. Existing UI has finite durations, creating an income cliff at exhaustion (Ganong and Noel, 2019; Figure 3). When workers are risk averse, inadequate UI can push them to choose accessible low-productivity jobs, whereas supported workers are willing to seek riskier high-productivity jobs.

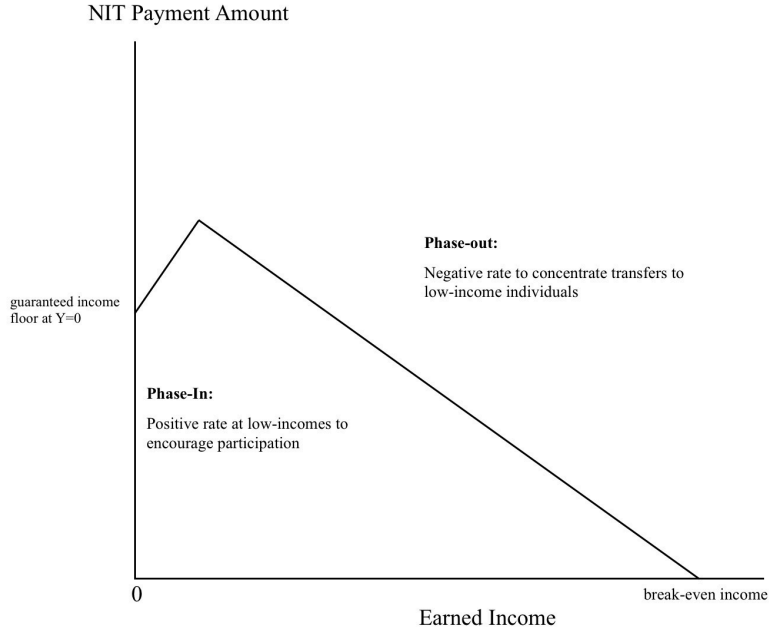


Nekoei and Weber (2017) find longer UI eligibility to increase re-employment wages by improving the quality of worker–job matches. However, a single optimal universal UI duration

may be infeasible in the case of AI-induced displacement. Due to AI's broad exposure, the skill gap to emerging occupations is different for each worker, leading to different required durations of unemployment for reallocation to high-productivity occupations. Therefore, the guaranteed, unconditional income floor provided by the NIT can support ambiguous durations of unemployment for different workers, enabling retraining and selective search towards emerging occupations.

Meanwhile, the NIT is commonly criticised for its disincentive to work. Chetty (2008) finds 40% of prolonged unemployment under higher UI to reflect moral hazard and distorted search incentives. If massive displacement occurs and extensive-margin responses dominate, a classical NIT weakens participation incentives, because the large payment at unemployment is immediately withdrawn with income (Saez, 2002).

In this case, a non-linear withdrawal rate may be preferable, with initially negative marginal tax rates to preserve re-employment incentives, resembling an EITC (Figure 4). However, this design only improves incentives provided jobs exist; it cannot solve an inherent shortage of vacancies. If reallocation fails to limit inequality, the NIT's rationale becomes purely redistributive.



D. The Efficiency Benchmark of Outcome Taxation

Under the Diamond–Mirrlees (1971) production-efficiency result, if all AI-generated gains were observable and taxable, the government should tax outcomes instead of inputs to preserve productive efficiency, maximising output before redistribution occurs. As AI integration is applicable in many industries, gains may dissipate across various outcomes:

$$G_{AI} = G_w + G_\pi + G_k + G_c$$

where G_w denotes gains to complementary workers, G_π profits or rents, G_k capital/wealth gains and G_c consumer gains. A practical redistributive system would need to draw upon a broad set of outcomes, including rent, capital, income or wealth (Table 2).

Instrument	Gain captured	Efficiency and feasibility
Corporate and excess-profits taxation	G_π	Targets rents from AI-enabled market power; production-efficient if levied on pure rents, but profit shifting across jurisdictions erodes the base.
Capital gains and dividend taxation	G_k	Captures appreciation of AI-owning firms; lock-in and deferral responses, and gains concentrated among top wealth holders.
Net wealth taxation	G_k	Broadest base for concentrated ownership gains; valuation and avoidance difficulties, and mixed international experience.
Progressive labour income taxation	G_w	Taxes wage gains of AI-complementary workers; well-established administration, but intensive-margin distortions at top rates.
Broad consumption taxation (VAT)	G_c	Captures gains passed to consumers through lower prices; large, stable base, but regressive unless paired with the NIT.

E. The Unusual Case for AI: the Token Tax

A token tax is a surcharge directly on observable commercial AI usage:

$$T = \tau PQ$$

where Q denotes taxable AI usage, P its price, and τ the tax rate. This could provide a relatively immediate and observable revenue source for redistribution as AI adoption expands.

Normally undesirable under productive efficiency, an input tax increases the marginal cost of all AI usage, delaying productivity benefits. However, this section will argue that a token tax is justified specifically under rapid AI-induced mass displacement. Acemoglu et al. (2020) argue that existing tax systems may already make labour relatively expensive compared to capital, encouraging excessive automation. In addition, Beraja and Zorzi (2025) show that the

private speed of automation can exceed the socially optimal speed, because firms do not internalise the welfare loss of borrowing-constrained workers.

As a general-purpose technology (GPT), AI displacement can occur across many occupations at the same time, whereas workers need time to accumulate new AI skills, and new vacancies are created later. In a standard DMP matching model, when unemployment rises relative to vacancies, labour-market tightness decreases:

$$\theta = \frac{V}{U}$$

Lower tightness reduces the job-finding rate:

$$f = \frac{M}{U}, \quad M = AU^{\alpha}V^{1-\alpha} \quad \Rightarrow \quad f = A\theta^{1-\alpha}, \ \frac{\partial f}{\partial \theta} > 0$$

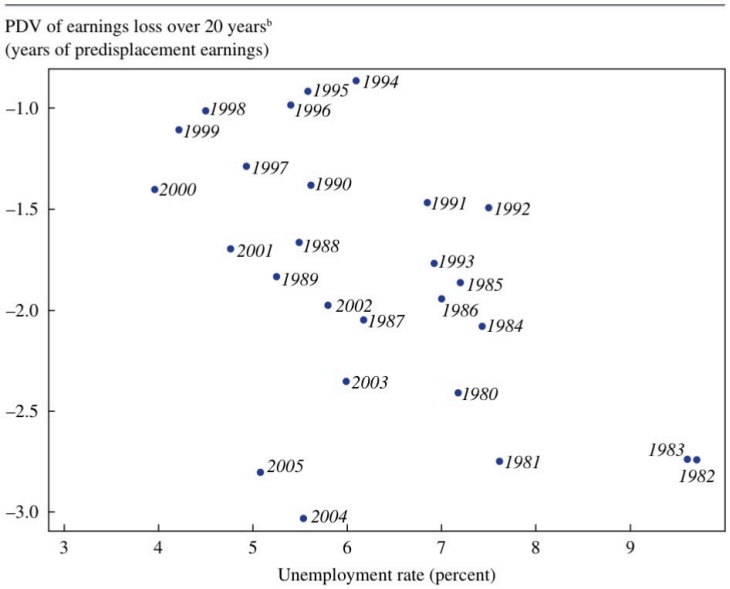
Unemployment increases:

$$u_{t+1} = u_t + s_t(1 - u_t) - f_t u_t; \quad s_t \uparrow \text{while } f_t \text{ adjusts slowly } \Rightarrow \ u_{t+1} \uparrow$$

Under lower tightness, the duration of unemployment may also increase:

$$E[D] = \frac{1}{f}; \quad \theta \downarrow \ \Rightarrow \ f \downarrow \ \Rightarrow \ E[D] \uparrow$$

Inequality in hours worked arises immediately, as higher unemployment and prolonged duration result in more displaced workers with zero hours of paid work. Moreover, workers displaced during weaker labour markets experienced larger and more persistent earnings losses than workers displaced in stronger labour markets (Davis and von Wachter, 2011; Figure 5). As unemployment lengthens and fewer vacancies are available relative to unemployed workers, workers' perceived difficulty of re-employment increases. This may encourage acceptance of low-quality matches, amplifying persistent inequality.



Therefore, a token tax may slow adoption to reduce displacement inflow, limiting immediate unemployment and reducing inequality during the transition. Meanwhile, Hershbein et al. (2026) find that a 10% increase in vacancies raises the probability of switching into a better job by 11–18%; so slower displacement may spread displacement over a longer period of time, preventing labour-market tightness from falling as sharply and increasing good-quality reallocation to reduce persistent inequality in wages and hours.

On the other hand, the token tax is justified only as a correction to excessively rapid AI-induced displacement. Yet, even if labour-market tightness θ is observable, the government cannot directly identify:

$$\frac{\partial R}{\partial \theta} \times \frac{\partial I}{\partial R} \times \frac{\partial W}{\partial I}$$

where R denotes reallocation quality and I inequality. The optimal tax is therefore ambiguous ex ante: policymakers face uncertainty over the extent of AI's impact on reallocation, and over how much taxation slows adoption.

One potential response is an adaptive rule:

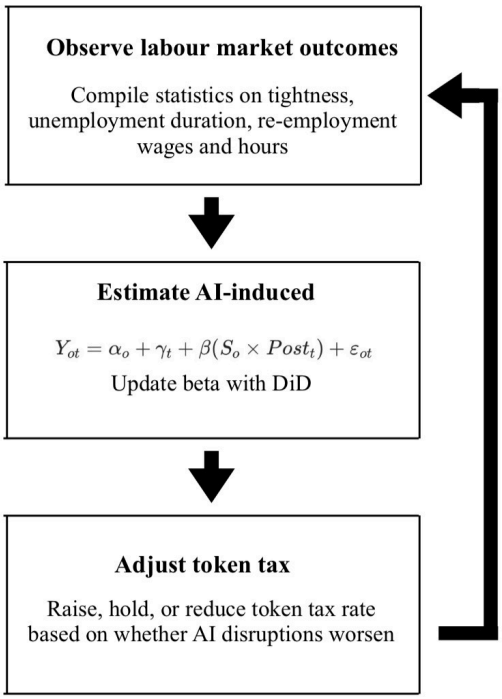
$$\tau_t = \tau(Z_t)$$

where Z_t contains a range of observable labour-market outcomes. The government should raise the tax when AI-related disruption worsens and reduce it as conditions recover. Since these

outcomes are affected by many shocks, AI-specific effects could be estimated using a continuous-exposure difference-in-differences:

$$Y_{ot} = \alpha_o + \gamma_t + \beta(S_o \times Post_t) + \varepsilon_{ot}$$

where S_o denotes AI substitution risk, and Y_{ot} includes outcomes for occupation o at time t, such as vacancies, unemployment duration or re-employment earnings. Regular re-estimation, or an event-study extension, would then allow τ_t to be updated as information on AI-induced disruption builds up (Figure 6).



While the adaptive rule mitigates the information problem, firms can still substitute between commercial APIs and self-hosted models. In addition, if one country taxes AI usage while others do not, AI-integrated production and AI development may shift towards other nations, while domestic firms face higher adoption costs.

F. Conclusion

The existing evidence does not yet establish that AI will generate massive technological unemployment. However, under the scenario posed by the question, widespread displacement could produce inequality in pay and paid hours, particularly if workers reallocate into lower-quality jobs.

This essay thus argues that the appropriate response is not to prevent productive AI adoption, but to redistribute its gains while improving the transition of displaced workers. A negative income tax provides redistribution, concentrating support on those experiencing the largest earnings losses while supporting heterogeneous durations of unemployment and under-employment. It should principally be financed through a range of taxation on the outcomes of AI adoption. However, an ordinarily distortionary token tax may be justified temporarily when unusually rapid displacement deteriorates labour-market conditions, and if private adoption fails to internalise displacement costs. Because information on the extent of these effects is initially unavailable, the tax should be adaptive and decline as labour-market conditions recover. Therefore, the policy objective should not simply be to minimise inequality in the short run, but to prevent rapid technological progress from outpacing reallocation and producing persistent welfare losses among those least able to adjust.

Bibliography

Visuals:

Figure 1: Cazzaniga et al. (2024), Figure 4.

Figure 2: Autor, Mindell and Reynolds (2020), Figure 6, p. 18.

Figure 3: Ganong and Noel (2019), Figure 1, Panel B, p. 2395.

Main body:

Acemoglu, D. (2025) 'The simple macroeconomics of AI', Economic Policy, 40(121), pp. 13–58. Available at: https://doi.org/10.1093/epolic/eiae042.

Acemoglu, D., Manera, A. and Restrepo, P. (2020) 'Does the US tax code favor automation?', Brookings Papers on Economic Activity, Spring, pp. 231–300.

Anthropic (2026) 'Project Glasswing: An initial update', 22 May. Available at: https://www.anthropic.com/research/glasswing-initial-update (Accessed: 30 August 2026).

Autor, D., Mindell, D. and Reynolds, E. (2020) The Work of the Future: Building Better Jobs in an Age of Intelligent Machines. Cambridge, MA: MIT Task Force on the Work of the Future.

Beraja, M. and Zorzi, N. (2025) 'Inefficient automation', Review of Economic Studies, 92(1), pp. 69–96. Available at: https://doi.org/10.1093/restud/rdae019.

Brynjolfsson, E., Chandar, B. and Chen, R. (2025) Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence. Working paper. Stanford, CA: Stanford Digital Economy Lab.

Cazzaniga, M., Jaumotte, F., Li, L., Melina, G., Panton, A.J., Pizzinelli, C., Rockall, E.J. and Mendes Tavares, M. (2024) Gen-AI: Artificial Intelligence and the Future of Work. IMF Staff Discussion Note SDN/2024/001. Washington, DC: International Monetary Fund. Available at: https://doi.org/10.5089/9798400262548.006.

Chetty, R. (2008) 'Moral hazard versus liquidity and optimal unemployment insurance', Journal of Political Economy, 116(2), pp. 173–234. Available at: https://doi.org/10.1086/588585.

Davis, S.J. and von Wachter, T. (2011) 'Recessions and the costs of job loss', Brookings Papers on Economic Activity, Fall, pp. 1–55.

Diamond, P.A. and Mirrlees, J.A. (1971) 'Optimal taxation and public production I: Production efficiency', American Economic Review, 61(1), pp. 8–27.

Eloundou, T., Manning, S., Mishkin, P. and Rock, D. (2024) 'GPTs are GPTs: Labor market impact potential of LLMs', Science, 384(6702), pp. 1306–1308. Available at: https://doi.org/10.1126/science.adj0998.

Ganong, P. and Noel, P. (2019) 'Consumer spending during unemployment: Positive and normative implications', American Economic Review, 109(7), pp. 2383–2424. Available at: https://doi.org/10.1257/aer.20170537.

Gmyrek, P., Berg, J., Kamiński, K., Konopczyński, F., Ładna, A., Nafradi, B., Rosłaniec, K. and Troszyński, M. (2025) Generative AI and Jobs: A Refined Global Index of Occupational Exposure. ILO Working Paper 140. Geneva: International Labour Organization. Available at: https://doi.org/10.54394/HETP0387.

Hershbein, B., Lim, K., Webber, D. and Zabek, M. (2026) Local Labor Market Tightness and Job Quality: Evidence from Job Changers. Finance and Economics Discussion Series 2026-043. Washington, DC: Board of Governors of the Federal Reserve System. Available at: https://doi.org/10.17016/FEDS.2026.043.

Lin, L. and Parker, K. (2025) U.S. Workers Are More Worried Than Hopeful About Future AI Use in the Workplace. Washington, DC: Pew Research Center. Available at: https://www.pewresearch.org/social-trends/2025/02/25/u-s-workers-are-more-worried-than-hopeful-about-future-ai-use-in-the-workplace/.

Nekoei, A. and Weber, A. (2017) 'Does extending unemployment benefits improve job quality?', American Economic Review, 107(2), pp. 527–561. Available at: https://doi.org/10.1257/aer.20150528.

Saez, E. (2002) ‘Optimal income transfer programs: Intensive versus extensive labor supply responses’, Quarterly Journal of Economics, 117(3), pp. 1039–1073. Available at: https://doi.org/10.1162/003355302760193959.