Inefficient Automation*

Martin Beraja
MIT and NBER

Nathan Zorzi
Dartmouth

February 16, 2023

Abstract

How should the government respond to automation? We study this question in a heterogeneous agent model that takes worker displacement seriously. We recognize that displaced workers face two frictions in practice: reallocation is slow and borrowing is limited. We analyze a second best problem where the government can tax automation but lacks redistributive tools to fully alleviate borrowing frictions. The equilibrium is (constrained) inefficient and automation is excessive. The reason is that there is a conflict between how firms and displaced workers value the effects of automation over time. The government finds it optimal to tax automation on efficiency grounds, even when it does not value equity. Slowing down automation increases aggregate consumption and redistributes early on during the transition, precisely when displaced workers value it more. Using a quantitative version of our model, we find that the optimal speed of automation is considerably lower than at the laissez-faire. The optimal policy improves efficiency and achieves substantial welfare gains.

*Martin Beraja: maberaja@mit.edu. Nathan Zorzi: nathan.g.zorzi@dartmouth.edu. We thank our discussants, Pascual Restrepo and Henry Siu. We are grateful to Daron Acemoglu, George-Marios Angeletos, Adrien Bilal, Ricardo Caballero, Diego Comin, Arnaud Costinot, Mariacristina De Nardi, John Grigsby, Jonathon Hazell, Roozbeh Hosseini, Anders Humlum, Chad Jones, Narayana Kocherlakota, Pablo Kurlat, Monica Morlacco, Giuseppe Moscarini, Jeremy Pearce, Elisa Rubbo, Katja Seim, Aleh Tsyvinski, Gustavo Ventura, Jesús Fernández-Villaverde, Conor Walsh, Iván Werning and Christian Wolf for helpful discussions. All errors are our own.

1 Introduction

Automation technologies raise productivity but disrupt labor markets, displacing workers and lowering their earnings (Humlum, 2019; Acemoglu and Restrepo, 2022). The increasing adoption of automation has fueled an active debate about appropriate policy interventions (Lohr, 2022). Despite the growing public interest in this question, the literature has yet to produce optimal policy results that take into account the frictions that workers face in practice when they are displaced by automation.

The existing literature that justifies taxing automation assumes that worker reallocation is frictionless or absent altogether. First, recent work shows that a government that has a preference for redistribution should tax automation to mitigate its distributional consequences (see Guerreiro et al., 2022 and subsequent work by Costinot and Werning 2022; Korinek and Stiglitz 2020). This literature assumes that automation and labor reallocation are instrinsically efficient, and that the government is willing to sacrifice efficiency for equity. Second, a large literature finds that taxing capital in the long-run — and automation, by extension — might improve efficiency in economies with incomplete markets (Aiyagari, 1995; Conesa et al., 2009). This literature abstracts from worker displacement and labor reallocation.

In this paper, we take worker displacement seriously and study how a government should respond to automation. In particular, we recognize that workers face two important frictions when they reallocate or experience earnings losses. First, reallocation is slow: workers face barriers to mobility and may go through unemployment or retraining spells before finding a new job (Jacobson et al., 2005; Lee and Wolpin, 2006). Second, credit markets are imperfect: workers have a limited ability to borrow against future incomes (Jappelli and Pistaferri, 2017), especially when moving between jobs (Chetty, 2008).

We show that these frictions result in *inefficient* automation. A government should tax automation — even if it does not value equity — when it lacks redistributive instruments to fully alleviate borrowing frictions. The optimal policy *slows down* automation while workers reallocate but does not tax it in the long-run. Quantitatively, we find important welfare gains from slowing down automation.

We incorporate reallocation and borrowing frictions in a dynamic model with endogenous automation and heterogeneous agents. Occupations use labor as an input. Firms invest in automation to expand their productive capacity. Automated occupations become less labor intensive, which displaces workers but increases output as labor reallocates to non-automated occupations. Displaced workers face reallocation frictions: they receive random opportunites to move between occupations, experience a temporary period of unemployment or retraining when they do so (Alvarez and Shimer, 2011), and incur a productivity loss due to the specificity of their skills (Adão et al., 2020). Workers also face financial frictions: they are not insured against the risk that their occupation is automated and face borrowing constraints (Huggett, 1993; Aiyagari, 1994). This baseline model has the minimal elements needed to study our question. We enrich it for our quantitative analysis.

Displaced workers experience earnings losses when their occupation is automated, but expect their income to increase as they slowly reallocate and find a new job. This creates a motive for borrowing to smooth consumption during this transition. When borrowing and reallocation frictions are sufficiently severe, displaced workers are pushed against their borrowing constraints.[1] Their consumption profiles are steeper than those of unconstrained workers who price the firms' equity. That is, automated workers are effectively more impatient than firms. Thus, there is a *conflict* between how firms and displaced workers value the effects of automation *over time*. This creates a motive for policy intervention on efficiency grounds.

In principle, the government could implement a first best if it was able to fully alleviate borrowing constraints using redistributive transfers. This is unlikely in practice, which motivates us to study second best interventions.[2] In particular, we analyze the constrained Ramsey problem of a government that can tax automation and implement active labor market interventions but is unable to fully alleviate the borrowing constraints of displaced workers by redistributing income.[3]

[1] This is consistent with the evidence. The earnings of displaced workers fall but later partially recover (Jacobson et al., 1993); including for those exposed to technological change (Braxton and Taska, 2023). Moreover, workers who loose their job indeed attempt to borrow (Sullivan, 2008), but are often unable to fully smooth consumption (Landais and Spinnewijn, 2021) or finance their retraining (Humlum et al., 2022) while unemployed.

[2] Governments often do not have have access to such rich instruments, which is precisely what motivates the public finance literature (Piketty and Saez, 2013). Moreover, the taxes required to pay for the transfers could tighten constraints for other workers (Aiyagari and McGrattan, 1998) and carry large dead-weight losses (Guner et al., 2021), and the take-up of transfers could be low (Schochet et al., 2012). We allow for various forms of social insurance in our quantitative model.

[3] These instruments are already used in many countries. For example, US taxes vary by type of capital and in fact *favor* automation (Acemoglu et al., 2020). South Korea reduced tax credits on automation investments, Nevada imposed an excise tax on autonomous vehicles, and the Grand

We have two main theoretical results. Our first result shows that the equilibrium is generically constrained inefficient, as defined by Geanakoplos and Polemarchakis (1985). Firms fail to internalize the aggregate and distributional effects of automation over time on displaced workers who are borrowing constrained. Taxing automation and implementing active labor market interventions makes automated workers strictly better off and leaves non-automated workers indifferent — a Pareto improvement. The policy increases aggregate consumption and redistributes early on in the transition, precisely when displaced workers value it more.

Our second result characterizes optimal policy for a given set of Pareto weights. To focus on the new *efficiency* rationale that we propose, we consider weights that remove any *equity* motive. These weights ensure that the government would not distort an efficient economy to redistribute.

We show that taxing automation is optimal on efficiency grounds alone. In particular, the government should *slow down* automation while labor reallocation takes place but should not intervene in the long-run. The optimal policy not only improves efficiency but also equity when the government values it. There is no trade-off, in contrast to the literature on the taxation of automation on equity grounds. As an extension, we also consider a third best problem where the government can tax automation but cannot implement active labor market interventions. This is motivated by the fact that such interventions have mixed results (Card et al., 2018) or unintended effects (Crépon and van den Berg, 2016). The rationale for taxing automation on efficiency grounds is reinforced, as borrowing constrained workers rely excessively on mobility to self-insure.

We conclude the paper with a quantitative exploration. Our goal is to evaluate the efficiency and welfare gains from slowing down automation, while allowing for various redistributive instruments. Our theoretical analysis found that workers' consumption profiles are key for optimal policy. These profiles are determined by reallocation frictions and the ability of workers to smooth consumption. Thus, we enrich our baseline model to ensure it performs well along these dimensions. First, we introduce idiosyncratic mobility shocks (Artuç et al., 2010), which leads to a dynamic discrete choice for reallocation and gross flows across occupations (Moscarini and Vella, 2008). Second, we add uninsured earnings risk (Floden and

Council of Geneva in Switzerland proposed to tax automated cashiers. See Kovacev (2020) for a detailed review.

Lindé, 2001), which produces a realistic distribution of savings. We also allow for unemployment benefits (Krueger et al., 2016) and non-linear income taxation (Heathcote et al., 2017) to account for existing insurance that helps workers. We calibrate the model to match several key moments of the US economy. In particular, we match the dynamics of occupation-level wages since 1980 in Cortes (2016).

We find that the constrained planner slows down the speed of automation so as to increase its half-life from 15 years at the laissez-faire to 22 years at the optimum. The optimal tax reduces investments in automation especially over the first decade of the transition. The tax starts at roughly 5%, raises progressively to 7% over a decade and then gradually declines, reaching roughly zero in year 25. Automated workers benefit substantially — their welfare increases by 0.80% in consumption equivalent terms — whereas non-automated workers and new generations are worse off — their welfare falls by 0.19% and 0.08%, respectively. The optimal policy offsets more than half of the gap in welfare between automated and non-automated workers at the laissez-faire. Overall, the policy raises social welfare by 0.20%.

We then consider several robustness checks and an alternative policy. First, we target a narrower definition of liquid assets. Automated workers are more likely to become borrowing constrained. They benefit more from slowing down automation and the total welfare gains increase. Second, we target a lower occupational mobility rate to reflect its decline in recent decades. The consumption of automated workers is lower than in our benchmark as they reallocate less, but the slope of their consumption profile is not meaningfully affected. Therefore, they benefit more from the intervention but the total welfare gains are mostly unchanged. Finally, as an alternative policy, we allow the government to partially insure automated workers by providing wage supplements — similar to Trade Adjustment Assistance for Workers (TAA) in the US. In present discounted terms, the government would need to give about $20,000 to the average automated worker to deliver the same welfare gains to them as the optimal tax on automation. The aggregate fiscal cost of this policy would be several orders of magnitude larger than the amount currently budgeted for TAA. This suggests that slowing down automation delivers welfare gains that would be costly to replicate with wage supplements alone.

Our paper relates to several strands of the literature. We contribute to the literature on the labor market impact of automation (Acemoglu and Restrepo, 2018; Martinez, 2018; Humlum, 2019; Moll et al., 2021; Hémous and Olsen, 2022) by studying optimal policy in an economy with frictions and quantifying the gains from slowing down automation. Moreover, we show that taxing automation improves *both* efficiency and equity, while there is a trade-off in the efficient economies studied in the literature (Guerreiro et al., 2022; Costinot and Werning, 2022; Thuemmel, 2018; Korinek and Stiglitz, 2020).

The rationale we propose for taxing automation also complements a large literature on capital taxation due to equity considerations (Judd, 1985; Chamley, 1986), dynamic inefficiency (Diamond, 1965; Aguiar et al., 2021), or pecuniary externalities when markets are incomplete (Conesa et al., 2009; Dávila et al., 2012; Dávila and Korinek, 2018). Optimal policies in our model also address pecuniary externalities. However, these externalities are distinct from the type encountered in the incomplete markets literature. They rely neither on the presence of uninsured idiosyncratic risk, nor on endogenous borrowing constraints. In addition, the literature on pecuniary externalities has almost exclusively studied static (or two-period) models or long-run stationary equilibria. The *timing* of these externalities plays no role in optimal policy. In contrast, the rationale for intervention that we propose applies during the *transition* to the long run, and the timing of externalities is central to optimal policy.

The mechanism that we present applies to any changes in labor demand that displace labor, including creative destruction (Caballero and Hammour, 1996) and offshoring (Hummels et al., 2018). We show that slowing down the adoption of automation technologies can improve efficiency when displaced workers are borrowing constrained. As such, our paper complements a literature studying the optimal speed of structural reforms and trade liberalization (Aghion and Blanchard, 1994; Caballero and Hammour, 1996; Neary, 1982; Mussa, 1984).

Methodologically, our quantitative model combines two state-of-the-art frameworks: (i) dynamic discrete choice models with mobility shocks (Artuç et al., 2010) used for studying the impact of technologies and trade; and (ii) heterogeneous-agent models (Huggett, 1993; Aiyagari, 1994) used for analyzing consumption and insurance. Our analysis also contributes to the public finance literature studying optimal taxation (Heathcote et al., 2017) and social insurance (Imrohoroglu et al., 1995; Golosov and Tsyvinski, 2006) in dynamic models with heterogeneous agents.

2 Model

Time is continuous and there is no aggregate uncertainty. Periods are indexed by $t \geq 0$. The economy consists of a representative firm producing a final good and a continuum of workers with unit mass. We first describe the problem of the firm which chooses automation and labor demands. We then describe the workers' problem, including the assets they trade, the frictions they face and their sources of income. Finally, we define a competitive equilibrium.

2.1 Firm

The firm produces the final good by combining the output of two occupations. Both occupations use labor as an input. The first occupation can be automated (e.g., a routine-intensive occupation) whereas the second cannot. At time $t = 0$, the firm chooses the degree of automation α in the automatable occupation.[4] We denote automated and non-automated occupations by $h = \{A, N\}$. At time $t \geq 0$, the firm chooses labor demands $\{\mu_t^A, \mu_t^N\}$ in both occupations.

Technology. Aggregate output is produced by combining the output y_t^h of the two occupations with a neoclassical technology

$$Y_t = G\left(y_t^A, y_t^N\right). \tag{2.1}$$

The occupations' outputs are

$$y_t^h = \begin{cases} F\left(\mu_t^A; \alpha\right) & \text{if automated } (h = A) \\ F^{\star}\left(\mu_t^N\right) = F\left(\mu_t^N; 0\right) & \text{otherwise } (h = N) \end{cases}, \tag{2.2}$$

for some production function $F(\cdot)$ with (weakly) decreasing returns to scale in labor. Automation is labor-displacing: it decreases the marginal product of labor in the automated occupation.[5] Moreover, occupations are (weak) complements, so

[4] For now, automation is chosen once and for all. We introduce gradual investment later on. This allows us to clarify that the optimal policy is to *slow down* automation while labor reallocates.

[5] It should be noted that some forms of automation might complement labor within occupations too. We focus on automation technologies that displace labor, such as industrial robots, certain types of artificial intelligence, autonomous vehicles, automated cashiers, etc.

automation increases the marginal product of the non-automated occupation. We formalize these assumptions below.

Assumption 1 (Technology)**.** *The marginal product of labor $\partial_{\mu} F(\mu;\alpha)$ decreases with automation α, and $\partial^2_{A,N} G\left(y^A, y^N\right) \geq 0$ so that occupations are complements.*

Automation increases output and can improve *aggregate* labor productivity, but it comes at a cost $\mathcal{C}(\alpha)$.[6] For example, the technology requires some continued investment due to depreciation (as in our quantitative model). We define the aggregate production function net of the cost of investing in automation

$$G^{\star}\left(\mu^A, \mu^N; \alpha\right) \equiv G\left(F\left(\mu^A; \alpha\right), F\left(\mu^N; 0\right)\right) - \mathcal{C}(\alpha). \tag{2.3}$$

We refer to $G^{\star}(\cdot)$ as *output* in the following.

Example. We illustrate the production function (2.3) with an example based on the model of Acemoglu and Restrepo (2018). There is a mass ϕ of automatable occupations ($h = A$) and a mass $1-\phi$ of non-automatable ones ($h = N$). Occupations operate a technology where automation and labor are perfect substitutes

$$y^A = F\left(\mu^A; \alpha\right) = \alpha + \mu^A \quad \text{and} \quad y^N = F^{\star}\left(\mu^N\right) = \mu^N.$$

The aggregate production function is

$$G^{\star}\left(\mu^A, \mu^N; \alpha\right) = \left[\phi\left(\alpha + \mu^A\right)^{\frac{\nu-1}{\nu}} + (1-\phi)\left(\mu^N\right)^{\frac{\nu-1}{\nu}}\right]^{\frac{\nu}{\nu-1}} - \delta\alpha,$$

where $\nu < 1$ is the elasticity of substitution across occupations and δ is the marginal cost of automation.

Optimization. The firm chooses the degree of automation α and labor demands $\{\mu^h_t\}$ to maximize the value of its equity

$$\max_{\alpha \geq 0} \int_0^{+\infty} Q_t \Pi_t(\alpha)\, dt \tag{2.4}$$

[6] A larger α lowers the marginal product of labor *within* the automated occupation but can raise the *aggregate* marginal product of labor (Appendix A.7). This is the case in the quantitative model.

where $\{Q_t\}$ is the equilibrium stochastic discount factor, and

$$\Pi_t(\alpha) \equiv \max_{\mu^A,\mu^N \geq 0} G^\star\left(\mu^A, \mu^N; \alpha\right) - \mu^A w_t^A - \mu^N w_t^N \tag{2.5}$$

are profits given wages $\left\{w_t^h\right\}$ and the price of the final good (normalized to 1).

We impose a regularity condition that ensures that automation is positive and finite in equilibrium. This is needed for a meaningful discussion of automation.

Assumption 2 (Interior solution)**.** *The production function* $G^\star\left(\mu^A, \mu^N; \alpha\right)$ *is concave in* α *and satisfies* $\partial_\alpha G^\star\left(\mu^A, \mu^N; \alpha\right)\big|_{\alpha=0} > 0$ *and* $\lim_{\alpha\to+\infty} \partial_\alpha G^\star\left(\mu^A, \mu^N; \alpha\right) = -\infty$ *for any* $0 \leq \mu^A \leq \frac{1}{2}$ *and* $\mu^N \geq \frac{1}{2}$.

2.2 Workers

Workers consume and save in financial assets. They supply inelastically one unit of labor and choose to reallocate across occupations.

Preferences. Workers' preferences over consumption flows $\{c_t\}$ are represented by

$$U = \mathbb{E}_0\left[\int_0^{+\infty} \exp\left(-\rho t\right) u\left(c_t\right) dt\right] \tag{2.6}$$

for some discount rate $\rho > 0$ and some isoelastic utility $u(c) \equiv \frac{c^{1-\sigma}}{1-\sigma}$ with $\sigma > 0$.

Reallocation frictions. We assume that the process of labor reallocation is *slow*. At time $t = 0$, workers are equally distributed across occupations, so there is a mass $1/2$ in automated and non-automated occupations. Workers are given the opportunity to reallocate to a new occupation with intensity λ. If they do so, they enter their new occupation with probability $1 - \iota$ or a temporary state of non-employment with probability i. Workers exit non-employment at rate $\kappa > 0$, at which point they enter their new occupation. The non-employment state can be interpreted either as unemployment due to search frictions or as temporary exit from the labor force while workers retrain.[7] Finally, we assume that workers in-

[7] Workers' mobility decision is purely time-dependent, which delivers tractable expressions. We allow for state-dependent mobility in our quantitative model (Section 5). We also allow for one other reason for slow labor reallocation (new generations gradually replacing older ones).

cur a permanent productivity loss $\theta \in (0, 1]$ after they have reallocated. This loss captures the lack of transferability of skills across occupations.

To retain tractability and abstract from idiosyncratic insurance considerations at this point, we assume that workers initially employed in each occupation form a large household.[8] This allows them to achieve full risk sharing against the risks of being allowed to reallocate (at rate λ), becoming unemployed (probability ι), and exiting unemployment (at rate κ). In what follows, we refer to each large household as automated $(h = A)$ or non-automated $(h = N)$ workers.

Assets. We suppose that financial markets are incomplete: workers cannot trade contingent securities against the risk that their initial occupation is automated.[9] Workers trade bonds and the firm's equity. Bonds are in zero net supply, and workers have no bonds initially. There is a unit of equity, which is initially in the hands of a competitive mutual fund that trades the same two assets.[10] Workers hold an equal and fixed share in this mutual fund, which rebates profits lump sum to them.

Budget constraint. A worker's flow budget constraint is

$$da_t^h = \left(\hat{\mathcal{Y}}_t^h + \Pi_t + r_t a_t^h - c_t^h \right) dt, \tag{2.7}$$

where a_t^h is bond holdings, $\hat{\mathcal{Y}}_t^h$ is labor income, Π_t is the profits rebated by the mutual fund, and $r_t \geq 0$ is the return on savings. To save on notation, the budget constraint (2.7) implicitly assumes that workers only save in bonds. This is without loss of generality, as workers will be indifferent between saving in bonds or equity

[8] This assumption prevents an artificial dispersion in the distribution of assets and implies that a worker's reallocation history is irrelevant. We relax this assumption in our quantitative model.

[9] We rule out complete markets for two reasons: financial markets participations is limited in practice (Mankiw and Zeldes, 1991); and workers' equity holdings are typically not hedged against their employment risk (Poterba, 2003). The absence of contingent securities is precisely what motivates the literature on the regulation of automation. The equilibrium would be efficient if workers could trade contingent securities before occupations become automated.

[10] This assumption ensures that (displaced) workers cannot sell equity to self-insure when their occupation is automated. In practice, almost all the firm equity in the US is held by the wealthiest 10% of households (Survey of Consumer Finances, 2022) — not the typical displaced worker. Accordingly, an alternative approach would have been to introduce a third agent (i.e., a Ricardian investor) who trades and holds equity but does not supply labor. We did not pursue this route to avoid introducing an additional agent.

in equilibrium.[11] Labor income $\hat{\mathcal{Y}}_t^h$ is

$$\hat{\mathcal{Y}}_t^h = \begin{cases} w_t^A \left(1 - u_t - \tilde{\mu}_t\right) + \left(1 - \theta\right) w_t^N \tilde{\mu}_t & \text{if } h = A \\ w_t^N & \text{if } h = N, \end{cases} \tag{2.8}$$

where u_t and $\tilde{\mu}_t$ are the shares of automated workers who are unemployed or have become employed in the non-automated occupation, respectively. Expression (2.8) already uses the fact that, in equilibrium, non-automated workers do not reallocate. The expression also assumes that unemployed workers earn no income.[12]

Borrowing friction. Workers are subject to a borrowing constraint

$$a_t^h \geq \underline{a} \tag{2.9}$$

where the borrowing limit is $\underline{a} \leq 0$.

Optimization. The households maximize utility (2.6) by choosing consumption c_t^h, bonds a_t^h, and reallocation m_t^h, subject to the following constraints. First, they must satisfy the budget constraint (2.7) and borrowing constraint (2.9). Second, their labor income is given by (2.8). Third, workers' labor supply across occupations is consistent with their reallocation choice m_t^h, given reallocation frictions. Since only automated workers find it optimal to reallocate, in the following we use $m_t \equiv m_t^A$ and implicitly set $m_t^N = 0$. The laws of motion for the share of automated workers who are unemployed (u_t) or employed in the non-automated occupation $(\tilde{\mu}_t)$ are

$$du_t = \left[\lambda \iota \left(1 - u_t - \tilde{\mu}_t\right) m_t - \kappa u_t\right] dt \tag{2.10}$$

$$d\tilde{\mu}_t = \left[\lambda \left(1 - \iota\right) \left(1 - u_t - \tilde{\mu}_t\right) m_t + \kappa u_t\right] dt, \tag{2.11}$$

with $u_0 = \tilde{\mu}_0 = 0$. Next, we impose a regularity condition on reallocation frictions which ensures that reallocation takes place in equilibrium and output does not decrease over time.

[11] In equilibrium, there will be no arbitrage between bonds and equity — i.e., condition (2.14) holds. The reason is that (i) both assets are traded, and (ii) the borrowing constraint (2.9) applies to the sum of bond and equity holdings as in Werning (2015).

[12] Our quantitative model has unemployment benefits and gross flows across occupations.

Assumption 3 (Reallocation frictions)**.** *The productivity loss θ is sufficiently small and the duration of unemployment $1/\kappa$ is sufficiently short that $1-(1-\theta)(1-1/\kappa) < Z^{\star}$ for some $Z^{\star} > 0$ defined in Appendix A.5.*

2.3 Equilibrium

Market clearing in the labor market requires

$$\mu_t^A = \frac{1}{2}\left(1-u_t-\tilde{\mu}_t\right) \quad \text{and} \quad \mu_t^N = \frac{1}{2}\left(1+(1-\theta)\,\tilde{\mu}_t\right) \tag{2.12}$$

for each occupation and all $t \geq 0$. The aggregate resource constraint is

$$G^*\left(\mu_t^A, \mu_t^N; \alpha\right) = \frac{1}{2}\left(c_t^A + c_t^N\right). \tag{2.13}$$

Finally, there is no arbitrage between bonds and equity, as workers and the (competitive) mutual fund can trade both. Thus, the firm discounts future cash-flows with the equilibrium interest rate r_t. The stochastic discount factor in (2.4) is

$$Q_t = \exp\left(-\int_0^t r_s ds\right). \tag{2.14}$$

We define a competitive equilibrium below.

Definition 1 (Competitive equilibrium)**.** A competitive equilibrium consists of a degree of automation α, and sequences for labor demands $\{\mu_t^h\}$, consumption and savings choices $\{c_t^h, a_t^h\}$, reallocation choices $\{m_t^h\}$, interest rate, stochastic discount factor, wages, profits and incomes $\{r_t, Q_t, w_t^h, \Pi_t, \hat{y}_t^h\}$ such that: (i) automation and labor demands are consistent with the firm's optimization; (ii) consumption, savings, and worker reallocation are consistent with workers' optimization; and (iii) the labor market clearing condition (2.12), the resource constraint (2.13), and the no arbitrage condition (2.14) are satisfied.

3 Equilibrium Characterization

We now characterize the laissez-faire equilibrium allocations. We begin with the allocations of labor, and consumption and savings *after* automation has occurred.

We then turn to the equilibrium degree of automation.

3.1 Labor Reallocation and Incomes

Firm optimization implies that wages equal the marginal products of labor

$$w_t^h \equiv \partial_h G^\star \left(\mu_t^A, \mu_t^N; \alpha\right) \tag{3.1}$$

for each $h = A, N$. Automation is labor-displacing and decreases the wage of automated workers. This induces them to reallocate to the non-automated occupation. As workers reallocate, the wedge between marginal products closes and output increases over time. The following proposition shows that automated workers reallocate until a stopping time T^{LF} when the marginal benefit of doing so is zero.

Lemma 1 (Equilibrium labor reallocation)**.** *The equilibrium reallocation of labor is characterized by a stopping time T^{LF} until which automated workers reallocate to non-automated occupations. Formally, $m_t = 1$ for all $t \leq T^{LF}$ and $m_t = 0$ otherwise. The stopping time satisfies the smooth pasting condition*

$$\int_{T^{LF}}^{+\infty} \exp\left(-\rho t\right) \frac{u'\left(c_t^A\right)}{u'\left(c_0^A\right)} \Delta_t dt = 0 \tag{3.2}$$

where

$$\Delta_t \equiv (1-\theta)\left[\iota\left(1 - \exp\left(-\kappa\left(t - T\right)\right)\right) + 1 - \iota\right] w_t^N - w_t^A \tag{3.3}$$

for all $t \geq T$ denotes the output gains from labor reallocation when evaluated at $T = T^{LF}$, since $w_t^h = \partial_h G^\star \left(\mu_t^A, \mu_t^N; \alpha\right)$ in equilibrium.

Proof. See Appendix A.1. □

The flows Δ_t capture the benefits and costs of reallocation. When an automated worker reallocates, they forgo their wage w_t^A and earn no income if they become unemployed (probability ι) or $(1-\theta)\, w_t^N$ if they enter the non-automated occupation (probability $1-\iota$) . As they exit unemployment at rate κ, they earn $(1-\theta)\, w_t^N$ too. The laissez-faire stopping time T^{LF} trades off these benefits and costs.

To complete the characterization, labor allocations across occupations are

$$\mu_t^A = \frac{1}{2} \exp\left(-\lambda \min\{t, T\}\right) \tag{3.4}$$

$$\begin{aligned} \mu_t^N = & \frac{1}{2} + \frac{1}{2}(1-\theta)\left(1 - \exp\left(-\lambda \min\{t,T\}\right)\right) \\ & - \frac{1}{2}(1-\theta)\,\iota \frac{\lambda}{\lambda-\kappa} \exp\left(-\kappa t\right)\left(1 - \exp\left(-\left(\lambda-\kappa\right)\min\{t,T\}\right)\right), \end{aligned} \tag{3.5}$$

evaluated at $T = T^{\mathrm{LF}}$, after solving the differential equations (2.10)–(2.11) and using labor market clearing (2.12).

Turning to earnings, automation drives a wedge between the labor incomes of automated and non-automated workers $\hat{\mathcal{Y}}_t^A - \hat{\mathcal{Y}}_t^N < 0$. In theory, this distributional effect could weaken or strenghten over time as labor reallocates, depending on third derivatives of the production function $G^\star(\cdot)$. Reallocation directly weakens this distributional effect since fewer workers remain employed in the automated occupation. However, reallocation indirectly affects equilibrium wages too. We next provide a sufficient condition for the direct effect to dominate. This ensures that the effect of automation on the labor income gap $\hat{\mathcal{Y}}_t^A - \hat{\mathcal{Y}}_t^N$ weakens over time, which is intuitive. The assumption uses the fact that this gap is approximately equal to $\mu \partial_\mu G^\star(\mu, 1-\mu; \alpha)$ at $\mu = \mu_t^A$ when unemployment spells are short (Assumption 3). We will only impose this assumption in Section 4.5.[13]

Assumption 4 (Labor income gap)**.** *The labor income gap* $\mu \partial_\mu G^\star(\mu, 1-\mu; \alpha)$ *has decreasing differences in* (α, μ).

3.2 Consumption and Savings

We now show that the labor displacement induced by automation creates a motive for borrowing and that workers become borrowing constrained when reallocation and borrowing frictions are sufficiently severe.

Lemma 2 (Binding borrowing constraints)**.** *Workers initially employed in the automated occupation* $(h = A)$ *borrow in equilibrium. They become borrowing constrained if and only if reallocation frictions* (λ, κ) *and borrowing frictions* $(\underline{a})$ *are sufficiently severe.*

[13] For instance, the example in Section 2.1 satisfies this assumption when evaluated in a symmetric allocation $y^A = y^N$, and if $\phi > \frac{1}{2}$ and $\nu < 2$ as in our quantitative model (Section 5).

This is the case when the borrowing limit $\underline{a} \leq 0$ is sufficiently tight that $\underline{a} > a^{\star}(\lambda, \kappa)$ for some threshold $a^{\star}(\cdot)$ defined in Appendix A.2. This threshold satisfies $a^{\star}(\lambda, \kappa) < 0$, i.e., borrowing constraints can bind, if and only if reallocation is slow ($1/\lambda > 0$ or $1/\kappa > 0$).

Proof. See Appendix A.2. □

To understand this result, the left panel of Figure 3.1 depicts the paths of the labor incomes for workers initially employed in each occupation

$$\hat{\mathcal{Y}}_t^h = \underbrace{w_t^h}_{\text{Initial wage}} + \mathbf{1}_{\{h=A\}} \times 2 \times \Bigg[\underbrace{\left(\frac{1}{2} - \mu_t^A\right) \times \left((1-\theta)\, w_t^N - w_t^A \right)}_{\text{Reallocation gains}} - \underbrace{\left(1 - \mu_t^A - \mu_t^N - \left(\frac{1}{2} - \mu_t^A\right)\theta\right) \times w_t^N}_{\text{Unemployment loss}} \Bigg]. \quad (3.6)$$

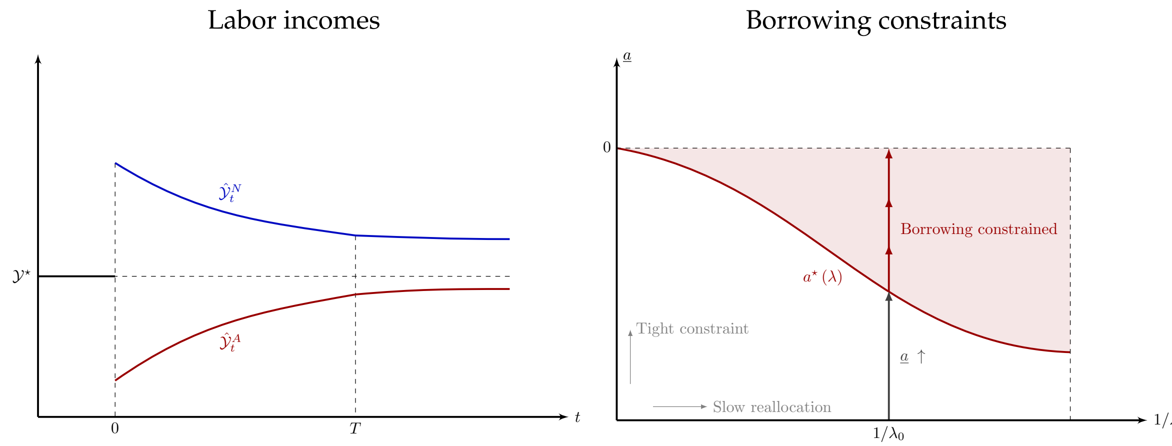
When reallocation is *slow*, automation decreases the income of workers displaced by automation, both directly by lowering the wage $w_t^A = \partial_A G^{\star}(\cdot)$ in their initial occupation and indirectly through unemployment $\left(1 - \mu_t^A - \mu_t^N\right)$. This decrease is not fully persistent though. Their income rises over time as they become employed in the non-automated occupation at a higher wage $(1-\theta)\, w_t^N$. Therefore, automated workers wish to borrow while they slowly reallocate.

Remark 1. *Workers displaced by automation expect their income to partially recover as they slowly reallocate. This creates a motive for borrowing.*

The right panel of the figure illustrates the second part of Lemma 2 in the space of reallocation frictions $(1/\lambda)$ and borrowing frictions $(\underline{a})$ in the particular case where unemployment spells are short $(1/\kappa \to 0)$. When the frictions are sufficiently mild, workers are never borrowing constrained, i.e., the white region in the figure. This region includes two limit cases in the literature. First, suppose that labor reallocation is *instantaneous* $(1/\lambda \to 0, 1/\kappa \to 0)$ as in Costinot and Werning (2022). In this case, there is no motive for borrowing, since income changes are fully permanent, and borrowing frictions are irrelevant. That is, slow reallocation is *necessary* for borrowing constraints to bind. Second, suppose that there are no borrowing frictions $(\underline{a} \to -\infty)$ as in Guerreiro et al. (2022).[14] In this case, automa-

[14] In Guerreiro et al. (2022), reallocation takes place (entirely) through new generations replacing older ones. We introduce overlapping generations in Section 4.6.2 and in our quantitative model.

Figure 3.1: Laissez-faire: labor incomes and borrowing constraints



tion still creates a motive for borrowing but workers are never constrained. As reallocation and borrowing frictions become more severe, borrowing constraints eventually bind $\underline{a} > a^\star(\cdot)$, i.e., the colored region in the figure.[15]

Turning to consumption, automated workers are worse off and have a higher marginal utility, i.e., $u'\left(c_t^A\right) > u'\left(c_t^N\right)$. Furthermore, binding borrowing constraints imply that automated workers have steeper consumption profiles, i.e., $u'\left(c_t^A\right)/u'\left(c_0^A\right) < u'\left(c_t^N\right)/u'\left(c_0^N\right) = \exp\left(-\int_0^t r_s ds\right)$.

Evidence on displaced workers. The literature on the consequences of job loss has documented that the earnings of displaced workers initially fall and later partially recover (Jacobson et al., 1993), which is consistent with Remark 1. This holds in particular for workers who switch occupations due to technological change (Braxton and Taska, 2023). Moreover, workers who loose their job indeed attempt to borrow (Sullivan, 2008) but often cannot fully smooth consumption due to borrowing constraints (Landais and Spinnewijn, 2021). Humlum et al. (2022) find that borrowing constraints affect the retraining decisions of unemployed workers. While we abstract from ex-ante heterogeneity across workers, our mechanism is more likely to be relevant when automation impacts workers with small liquidity buffers. For example, industrial robots, automated cashiers, or autonomous vehicles tend to displace low-to-middle income routine workers who are more likely to

[15] It should be noted that the threshold $a^\star(\lambda, \kappa)$ is non-monotonic in its arguments. In particular, $\lim_{1/\lambda \to +\infty} a^\star(\lambda, \kappa) = 0$ when workers cannot reallocate.

be hand-to-mouth. In contrast, artificial intelligence for natural language processing tends to affect higher income skilled workers who can borrow more easily.[16]

3.3 Automation

We now turn to the equilibrium automation choice.

Lemma 3 (Equilibrium automation)**.** *The degree of automation α^{LF} is unique and interior, and satisfies*

$$\int_0^{+\infty} \exp\left(-\int_0^t r_s ds\right) \Delta_t^{\star} dt = 0 \tag{3.7}$$

where

$$\Delta_t^{\star} \equiv \partial_\alpha G^{\star}\left(\mu_t^A, \mu_t^N; \alpha\right) \quad \text{for all } t \geq 0 \tag{3.8}$$

denotes the output gains from automation, and

$$Q_t = \exp\left(-\int_0^t r_s ds\right) = \exp\left(-\rho t\right) \frac{u'\left(c_t^N\right)}{u'\left(c_0^N\right)} \tag{3.9}$$

is the equilibrium stochastic discount factor used by the firm. The output gains from automation $\Delta_t^{\star}$ increase over time in equilibrium.

Proof. See Appendix A.3. □

The firm maximizes the present discounted value of output. No arbitrage between equity and bonds implies that the firm values cash-flows over time using the interest rate $\exp\left(-\int_0^t r_s ds\right)$, which equals the marginal rate of substitution over time (MRS) of non-automated workers $\exp\left(-\rho t\right) u'\left(c_t^N\right) / u'\left(c_0^N\right)$ in equilibrium since they are not borrowing constrained.

The firm trades off the benefits and costs of automation over time, which are captured in the output gains $\Delta_t^{\star}$. These gains build up over time as labor reallocates from the automated to the non-automated occupation. The reason is that

[16] The mechanism might, in theory, also apply to *increases* in labor demand in an occupation or sector, as workers would borrow in anticipation of higher wages. However, this type of anticipatory effect is likely to be weak (Poterba, 1988). Indeed, we find in our quantitative model that workers borrow substantially more after a fall in their occupation's wage compared to an increase in the other occupation's (result available upon request).

the marginal product of the automated occupation increases, by concavity of the technology and complementarity across occupations (Assumption 1).

4 Excessive Automation

In this section, we show that automation is *excessive* at the laissez-faire and characterize optimal policy. We first specify the set of policy instruments available to the government (Section 4.1). We then state the constrained Ramsey problem (Section 4.2), and discuss the aggregate and distributional effects from automation (Section 4.3). Next, we show that the equilibrium is constrained inefficient and that taxing automation Pareto improves upon the laissez-faire (Section 4.4). We then show that the government finds it optimal to tax automation purely on efficiency grounds (Section 4.5), even when it has no preference for redistribution. Finally, we present various extensions (Section 4.6). For tractability and to obtain more compact expressions, we assume in the following that workers cannot borrow $\underline{a} \to 0$.

4.1 Policy Instruments

A government that has access to a sufficiently rich set of lump-sum transfers to fully undo borrowing frictions could, in theory, implement a first best. For example, the government could use targeted lump-sum transfers $\left\{T_t^h\right\}$ (indexed by worker and time) to help displaced workers. In practice, such rich interventions are unlikely. The literatures on optimal taxation (Piketty and Saez, 2013) and the regulation of automation precisely rule out such transfers, in part due to their informational requirements.[17] This motivates us to study second best policy interventions.

We assume that the government has access to a simple set of instruments that depend on calendar time alone: a linear tax on automation τ^α, and active labor market interventions (Card et al., 2018) that tax or subsidize labor reallocation

[17] Alternatively, the government could implement *symmetric* transfers $\{T_t\}$ to effectively borrow on behalf of the workers. However, the associated debt needs to be repaid later by taxing them. This future tax burden could tighten borrowing constraints (Aiyagari and McGrattan, 1998) and carry large distorsions (Guner et al., 2021), limiting or reversing the benefits of the transfers. The transfers need to be generous enough to ensure that *no* worker is constrained — a scenario that the literature on heterogeneous agents has not seriously considered. The size of transfers is further limited by the fact that future higher taxes could push the poorest workers into default.

$\{\varsigma_t\}$.[18] These instruments are already used in many economies and do not require the government to know which occupations are automated or which workers are displaced. For instance, US taxes vary by type of capital (e.g., equipment, software, structures) and industry (due to differential depreciation allowances), and seem to be favoring automation instead of taxing it (Acemoglu et al., 2020). Concrete policies discriminating against automation technologies (Kovacev, 2020) include: (i) South Korea's reduction in the automation tax credit aimed at protecting workers in high-tech manufacturing; (ii) Nevada's excise tax on transportation companies using autonomous vehicles that would displace human drivers; and (iii) the Swiss canton of Geneva's proposed tax on retail stores installing automated cashiers. That said, identifying technologies that displace labor could be more challenging in other instances (e.g., artificial intelligence algorithms).

4.2 The Constrained Ramsey Problem

We consider the problem of a government that values automated and non-automated workers, and assigns them Pareto weights $\{\eta^A, \eta^N\}$. The government effectively controls two choices with its tax on automation and active labor market interventions: the degree of automation α; and the reallocation of workers, as governed by the stopping time T.[19] All other choices must be consistent with workers' and the firm's optimality.

Lemma 4 (Primal problem)**.** *The government maximizes the social welfare function*

$$\mathcal{U} = \sum_h \eta^h \int_0^{+\infty} \exp(-\rho t)\, u\left(c_t^h\right) dt \tag{4.1}$$

by choosing $\{\alpha, T, \mu_t^A, \mu_t^N, c_t^A, c_t^N\}$, subject to the laws of motion (3.4)–(3.5) for labor $\{\mu_t^A, \mu_t^N\}$, and the consumption allocations $c_t^h = \hat{\mathcal{Y}}_t^h + \Pi_t$ for workers initiallly employed in occupations $h = \{A, N\}$, where labor incomes $\hat{\mathcal{Y}}_t^h$ are given by (3.6) and profits Π_t are given by (2.5).

[18] To abstract from income effects, we assume that the large families reimburse lump sum any reallocation taxes or subsidies it perceives. The latter can take the form of credits for retraining programs or unemployment insurance (when positive), or penalties such as imperfect vesting of retirement funds (when negative).

[19] Formally, the government would control reallocation choices $\left\{m_t^h\right\}$. To save on notation, we directly impose that the optimal reallocation policy is a stopping time T for automated workers.

It is worth noting that the only difference between this constrained problem and the unconstrained (first best) Ramsey problem lies in the set of implementable consumption allocations. In the constrained problem, workers must consume their income, since borrowing is not possible ($\underline{a} \to 0$). In the first best problem, any consumption allocation that satisfies the resource constraint (2.13) is feasible.

4.3 Aggregate vs. Distributional Effects

Consider the effect of a policy intervention $\{\delta\alpha, \delta T\}$ on the government's objective $\mathcal{U}$ starting from the laissez-faire. The change in welfare is

$$\delta\mathcal{U} = \eta^N \times u'\left(c_0^N\right) \times \int_0^{+\infty} \underbrace{\exp\left(-\rho t\right)\frac{u'\left(c_t^N\right)}{u'\left(c_0^N\right)}}_{=\exp\left(-\int_0^t r_s ds\right)} \times \delta c_t^N dt$$
$$+\eta^A \times u'\left(c_0^A\right) \times \int_0^{+\infty} \underbrace{\exp\left(-\rho t\right)\frac{u'\left(c_t^A\right)}{u'\left(c_0^A\right)}}_{\text{How automated workers value flows}} \times \delta c_t^A dt \tag{4.2}$$

with

$$\delta c_t^h \equiv \delta\alpha \times \left(\Delta_t^\star + \Sigma_t^{h,\star}\right) + \delta T \times \mathbf{1}_{t>T}\left(\Delta_t + \Sigma_t^h\right), \tag{4.3}$$

where the flows $\Delta_t^\star$ and Δ_t are the aggregate output gains from automation (3.8) and reallocation (3.3), and the flows $\Sigma_t^{h,(\star)}$ capture purely distributional effects of automation and reallocation through its effect on the relative incomes of workers. By definition, these distributional terms sum up to zero $\Sigma_t^{A,(\star)} + \Sigma_t^{N,(\star)} = 0$ at all times.

No borrowing constraints. Consider first the case where borrowing constraints do not bind. Decreasing automation $\delta\alpha < 0$ has no first order aggregate effects on welfare. The reasons is that the firm and workers agree on how to value the output gains of automation $\Delta_t^\star$ over time. The workers' MRS coincide with the equilibrium interest rate $\exp\left(-\rho t\right) u'\left(c_t^h\right)/u'\left(c_0^h\right) = \exp\left(-\int_0^t r_s ds\right)$ for $h = A, N$ and the firm was already optimizing in equilibrium, i.e., equation (3.7) holds.[20] However, decreasing automation has distributional consequences: it makes automated

[20] A similar logic implies that varying reallocation δT has no aggregate effects on welfare either.

workers better-off relative to non-automated workers. Automated workers value changes in consumption more since $u'\left(c_t^A\right) > u'\left(c_t^N\right)$. This provides a redistributive motive for taxing automation when the government values equity, e.g., when it uses utilitarian weights $\eta^h \equiv 1/2$. This motive for intervention has been the focus of the exisiting literature (Guerreiro et al., 2022; Costinot and Werning, 2022).

Binding borrowing constraints. Suppose instead that reallocation and borrowing frictions are sufficiently important that borrowing constraints bind (Lemma 2). Automated workers who are displaced are effectively more impatient than the firm, since $u'\left(c_t^A\right)/u'\left(c_0^A\right) < u'\left(c_t^N\right)/u'\left(c_0^N\right)$. There is now a *conflict* between how the firm and displaced workers value the effects of automation *over time*. As we show in the next sections, this creates room for Pareto improvements and a new motive for taxing automation on efficiency grounds.

4.4 Constrained Inefficiency

We now establish that the equilibrium is *generically* constrained inefficient in the sense of Geanakoplos and Polemarchakis (1985). The government can implement a Pareto improvement by varying automation $(\delta\alpha)$ and reallocation (δT). This is the case in virtually *any* economy: if this happen not to be the case, then there exists an arbitrarily small perturbation of the production function $G^{\star}(\cdot)$ that again allows for a Pareto improvement.

Proposition 1 (Constrained inefficiency). *Generically, there exists a variation $\{\delta\alpha, \delta T\}$ starting from the laissez-faire which makes automated workers strictly better off $\left(\delta U^A > 0\right)$ and non-automated workers indifferent $\left(\delta U^N = 0\right)$. The Pareto improvement requires taxing automation $(\delta\alpha < 0)$.*

Proof. See Appendix A.4. □

To understand why taxing automation generates a Pareto improvement, we reproduce the main steps of the proof. Consider a reduction in automation $\delta\alpha < 0$. Absent any change in reallocation, non-automated workers would be worse off. To leave them indifferent $\delta U^N = 0$, the government can always compensate them by reducing the mass of workers who enter their occupation $\delta T < 0$, which lifts their wage.

This intervention $\{\delta\alpha, \delta T\}$ strictly improves the welfare of automated workers $\delta U^A > 0$ through both aggregate and distributional effects. The aggregate effect of the intervention on automated workers is

$$\int_0^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^A\right) \left\{\delta\alpha \times \Delta_t^\star + \delta T \times \mathbf{1}_{t>T}\Delta_t\right\} dt > 0. \tag{4.4}$$

The output gains from automation $\Delta_t^\star$ take time to materialize (Lemma 3). The cost is borne early on in the transition (the flows $\Delta_t^\star$ are initially negative) and output increases gradually as workers reallocate (the flows $\Delta_t^\star$ become positive later on). Automated workers are effectively more impatient than the firm since they are borrowing constrained.[21] They value the future output gains less, and the early costs more. Thus, curbing automation $\delta\alpha < 0$ benefits automated workers by increasing aggregate consumption early on in the transition when they value it more. The decrease in reallocation $\delta T < 0$ has no first order effects on their welfare, since automated workers were already reallocating optimally (Lemma 1).

The distributional effect of the intervention on automated workers is

$$\int_0^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^A\right) \left\{\delta\alpha \times \Sigma_t^{A,\star} + \delta T \times \mathbf{1}_{t>T}\Sigma_t^A\right\} dt > 0. \tag{4.5}$$

Curbing automation always benefits automated workers relative to non-automated workers $\delta\alpha \times \Sigma_t^{A,\star} > 0$. On the contrary, reducing reallocation makes them worse-off $\delta T \times \Sigma_t^A < 0$. When automated workers are borrowing constrained, the net effect (4.5) is positive. The reason is again that automated workers are effectively more impatient than non-automated workers. They value more the immediate gains $\delta\alpha \times \Sigma_t^{A,\star} > 0$ (for all $t \geq 0$) compared to the losses $\delta T \times \Sigma_t^A < 0$ which come further in the future after reallocation has stopped (only for $t \geq T$). Therefore, the intervention generates a Pareto improvement through both aggregate and distributional effects.

Remark 2. *Firms fail to internalize the effects of automation on displaced workers who are borrowing constrained. Taxing automation increases aggregate consumption and redistributes precisely when displaced workers value it more.*

The inefficiency that we document relies on the firm and displaced workers *dis-*

[21] This would be the case even if the interest rate was fixed, e.g., if we introduced additional Ricardian investors (domestic or international) with a constant MRS.

agreeing on how they value the aggregate and redistributive effects of automation over time. In practice, the wealthiest 10% of households hold close to 90% of firm equity in the US (Survey of Consumer Finances, 2022). The typical displaced worker does not hold or trade equity. As a result, the incentives of a firm to automate (say a car manufacturer) mostly reflect the incentives of its wealthy investors rather than those of the workers that it displaces (who may be borrowing constrained). That said, our mechanism could be muted if workers were represented in the boardroom directly.

4.5 Optimal Policy Interventions

We now characterize the constrained efficient degree of automation for a given set of Pareto weights. The optimal policy depends on how the government values *efficiency* and *equity*.[22] To see this, consider the social incentive to automate $\partial_\alpha \mathcal{U}$ starting from the laissez-faire. It can be decomposed as

$$\partial_\alpha \mathcal{U} = \underbrace{\sum_h \eta^h \int_0^{+\infty} \exp(-\rho t)\, u'\left(c_t^h\right) \times \left(\Delta_t^\star + \hat{\Sigma}_t^{h,\star}\right) dt}_{\text{Taxing } \alpha \text{ on efficiency grounds}}$$
$$+ \underbrace{\sum_h \eta^h \int_0^{+\infty} \exp(-\rho t)\, u'\left(c_t^h\right) dt \times \int_0^{+\infty} Q_t \Sigma_t^{h,\star} dt}_{\text{Taxing } \alpha \text{ on equity grounds}}, \tag{4.6}$$

where $\hat{\Sigma}_t^{h,\star} \equiv \Sigma_t^{h,\star} - \int_0^{+\infty} Q_s \Sigma_s^{h,\star} ds$. The efficiency component captures two effects: the aggregate effect of curbing automation $\Delta_t^\star$, and the change in the distribution of consumptions *over time* $\hat{\Sigma}_t^{h,\star}$ fixing their present net discounted value. This component is zero in any efficient economy where the MRS are equalized across workers. In turn, the equity component captures how consumption is redistributed across workers in present discounted value. This component depends on the differences across workers in the average marginal utilities over time.

Efficiency motive. To focus on the new *efficiency* rationale that we propose, we first consider a government that uses weights $\eta^{\text{effic,h}} = 1/\int_0^{+\infty} \exp(-\rho t)\, u'\left(c_t^h\right) dt$

[22] Bhandari et al. (2021) and Dávila and Schaab (2022) also provide decompositions of the welfare effects of policy into efficiency and equity components.

evaluated at the laissez-faire. These efficiency weights ensure that the government does not intervene to improve equity.[23] Proposition 2 below shows that taxing automation is optimal on efficiency grounds alone.

Proposition 2 (Taxing automation on efficiency grounds)**.** *Suppose that the government uses efficiency weights* $\eta^{\text{effic},h} = 1/\int_0^{+\infty} \exp(-\rho t)\, u'\left(c_t^h\right) dt$*. Then, taxing automation is optimal.*

Proof. See Appendix A.5. □

The intuition is similar to Proposition 1. The output gains from automation $\Delta_t^\star$ build up over time (Lemma 3). The distributional effects $\Sigma_t^{A,\star} = -\Sigma_t^{N,\star}$ are negative but weaken over time, as automated workers reallocate away from their occupation (Assumption 4). Thefore, taxing automation is optimal $(\partial_\alpha \mathcal{U} < 0)$ because it increases aggregate consumption and redistributes precisely at times when the average worker in the economy (under weights $\eta^{\text{effic,h}}$) values it more.

Finally, when the government's problem (4.1) is convex, Proposition 2 implies that the laissez-faire level of automation is excessive compared to its second best counterpart $\alpha^{\text{SB,effic}} < \alpha^{\text{LF}}$. The reason is that there is a unique global optimum. We will compute the optimum numerically in our quantitative model.

Equity motive. Taxing automation not only improves efficiency but also equity when the government values it. There is no trade-off, in contrast to the literature on the taxation of automation on equity grounds. A utilitarian government $\left(\eta^{\text{util},h} = 1/2\right)$ that values equity would tax automation even more compared to Proposition 2.

4.6 Extensions

We next consider a number of extensions to our analysis.

4.6.1 No Active Labor Market Interventions

In practice, ex post policies can be difficult to implement. Active labor market interventions often produce mixed results (Card et al., 2018), or have unintended

[23] In an efficient economy, the weights boil down to the standard inverse marginal utility weights $\eta^{\text{effic},h} = 1/u'\left(c_0^h\right)$ and the government does not intervene at all.

consequences for untargeted workers (Crépon and van den Berg, 2016). For instance, this would be the case with *gross* flows between occupations, as in our quantitative model. For this reason, we now consider a *third best* problem where the government controls automation but is unable to control labor reallocation. This implies that a Pareto improvement (Proposition 1) is no longer possible.

In addition to the direct effects in (4.6), the government now internalizes the indirect effect of automation due to reallocation $T(\alpha)$. This indirect effect is[24]

$$T'(\alpha) \times \frac{1}{2}\lambda \exp(-\lambda T) \times \int_{T(\alpha)}^{+\infty} \exp(-\rho t) \left\{ \eta^N u'\left(c_t^N\right) - \eta^A u'\left(c_t^A\right) \right\} \times \left(\Delta_t + \Sigma_t^N\right) dt.$$

Taxing automation decreases reallocation since $T'(\cdot) > 0$. This indirect effect can either reinforce or dampen the government's incentives to tax automation, depending on the Pareto weights. For instance, a utilitarian government would tax automation *less* compared to Proposition 2, as this induces more reallocation $\delta T > 0$ and redistributes towards automated workers. In contrast, one can show that a government using efficiency weights (which does not value such redistribution) finds it optimal to tax automation *more* when unemployment spells are not too long (as in Assumption 3).

4.6.2 Slowing Down Automation

An extensive literature argues that taxing capital might improve insurance (Conesa et al., 2009; Dávila et al., 2012) or prevent capital overaccumulation (Aiyagari, 1995) in economies with incomplete markets (or overlapping generations). These two rationales share two features: they rely on the presence on uninsured idiosyncratic risk and optimal policies affect investment in the *long-run*.

The rationale that we propose is conceptually distinct. First, we find that taxing automation is optimal even absent idiosyncratic uncertainty. Second, our mechanism implies that the government should *slow down* automation only while labor reallocation takes place and displaced workers are borrowing constrained, but has no reason to tax automation in the long-run. To clarify this last point, we extend our model along two dimensions that are relevant for studying dynamics over long horizons. Both dimensions are present in our quantitative model. First, we allow

[24] This expression uses Lemma 1 and the fact that $\Sigma_t^A + \Sigma_t^N = 0$.

for gradual investments in automation. We assume that the law of motion of automation is $d\alpha_t = (x_t - \delta\alpha_t)\,dt$ for some depreciation rate δ and gross investment rate x_t, and the cost of automation q_t declines over time. Second, we assume that there are overlapping generations of workers who are born (and die) at rate χ and can choose any occupation at birth. We show below that the government has no motive to intervene in the long-run.

Proposition 3 (No intervention in the long-run)**.** *In the long-run, the equilibrium converges to a first best allocation. In particular, $\alpha_t^{LF}/\alpha_t^{FB} \to 1$ as $t \to +\infty$, where α_t^{FB} is automation at the first best.*

Proof. See Appendix A.6. □

The government can improve neither efficiency nor equity in the long-run. Once labor reallocation is complete, workers' incomes are constant and they have no incentive to borrow. The intertemporal MRS of all workers are identical. Therefore, the firm's automation choice is efficient since it values the returns to automation over time as workers do. Moreover, the entry of new generations equalizes wages across occupations in the long-run. The marginal utilities of all workers are identical, and there is no need for redistribution in the long-run as in Guerreiro et al. (2022).

4.6.3 The Direction of Investments

So far, the firm could only invest in automation. Taxing it thus unequivocally reduces *total* investment. We now allow investments in a Hicks-neutral technology. We assume that aggregate output is

$$\tilde{G}\left(\mu^A, \mu^N; \alpha, A\right) = AG\left(F\left(\mu^A; \alpha\right), F\left(\mu^N; 0\right)\right) - \mathcal{C}(\alpha) - \Phi(A)$$

and the firm chooses automation α and productivity A. Hicks-neutral investments do not cause worker displacement. The adjustment is instantaneous and workers are not borrowing constrained. Therefore, the optimal policy changes the *direction* of investments: taxing automation but subsidizing Hicks-neutral investments.

It is also worth noting that our analysis abstracts from other reasons why the government might want to subsidize investment, e.g., firm credit constraints, externalities, etc. Therefore, our results do not necessarily imply that automation should be taxed *on net*. Rather, they suggest that automation should be taxed *relative to* other investments, e.g., through lower subsidies as in South Korea.

5 Quantitative Model

In the remaining of the paper, we quantitatively evaluate the efficiency rationale for slowing down automation. To this end, we enrich our baseline model along several dimensions that are important for the ability of workers to reallocate and smooth consumption. In particular, we allow for gross flows across occupations, uninsurable idiosyncratic earnings and mobility risks, and some forms of social insurance. We also introduce gradual automation and overlapping generations of workers (as in Section 4.6.2). Appendix B provides further details.

5.1 Firm

Production. There is a continuum of occupations of mass 1. A share ϕ are automatable ($h = A$) and a share $1 - \phi$ are not ($h = N$). Occupations use the technology

$$y_t^A = A^A \left(\alpha + \mu^A\right)^{1-\eta} \quad \text{and} \quad y_t^N = A^N \left(\mu^N\right)^{1-\eta} \tag{5.1}$$

where $\eta \in (0, 1)$ is the span of control, and $A^h > 0$ are occupation-specific productivities.[25] The firm's final good technology is

$$G\left(y_t^A, y_t^N\right) = \left[\phi \left(y_t^A\right)^{\frac{\nu-1}{\nu}} + (1-\phi)\left(y_t^N\right)^{\frac{\nu-1}{\nu}}\right]^{\frac{\nu}{\nu-1}}, \tag{5.2}$$

where $\nu < 1$ is the elasticity of substitution.

Investment. The firm invests in automation. The law of motion of automation is

$$d\alpha_t = \left(x_t - \delta\alpha_t\right) dt, \tag{5.3}$$

where δ is the rate of depreciation, and x_t is the investment rate. The firm incurs a

[25] We normalize the relative productivity of automation to 1. This is without loss of generality since only the ratio between this productivity and the cost automation q_t is relevant.

resource cost q_t per unit of investment x_t. As in Guerreiro et al. (2022), we assume that this cost falls over time $q_t = q^{\text{fin}} + \exp\left(-\psi t\right)\left(q^{\text{init}} - q^{\text{fin}}\right)$, where q^{init} and q^{fin} are the initial and final costs, and $\psi > 0$ is the convergence rate. The initial cost ensures that automation starts at $\alpha_0 = 0$. The government taxes automation linearly at rate $\{\tau_t^x\}$ and rebates the proceedings to the firm.

Dividends. The firm smoothes dividends over time (Leary and Michaely, 2011) and issues debt to finance investment early on.[26] Dividends are given by $\Pi_t^{\text{div}} = \Pi^{\text{fin}} + \exp\left(-\varsigma t\right)\left(\Pi^{\text{init}} - \Pi^{\text{fin}}\right)$ where Π^{init} and Π^{fin} are profits at the initial and final steady states. The convergence rate $\varsigma > 0$ ensures that the firm repays its debt $\int_0^{+\infty} \exp\left(-\int_0^t r_s ds\right)\left(\Pi_t^{\text{div}} - \Pi_t\right) = 0$.

5.2 Workers

There are overlapping generations of workers that are replaced at rate χ.[27] A worker is indexed by five states: their asset holdings (a); their occupation of employment (h); their employment status (e); their permanent productivity component (ξ); and the mean-reverting component of their productivity (z). We let $\mathbf{x} \equiv (a, h, e, \xi, z)$ be the workers' states and π its measure.

Assets and constraints. The asset structure is the same as in our baseline model.[28] In addition, workers have access to annuities which allows them to self-insure against survival risk. Financial markets are otherwise incomplete: workers cannot trade contingent securities against the risk that their occupation becomes automated, against the risk that they are not able to relocate, against unemployment risk, or against idiosyncratic productivity risk. Workers now face the budget constraint

$$da_t\left(\mathbf{x}\right) = \left[\mathcal{Y}_t^{\text{net}}\left(\mathbf{x}\right) + \left(r_t + \chi\right) a_t\left(\mathbf{x}\right) - c_t\left(\mathbf{x}\right)\right] dt \tag{5.4}$$

[26] Assuming that the firm smoothes dividends is conservative with respect to our mechanism. At short horizons, the investment cost of automation exceeds revenues and profits are negative. Disbursing negative dividends to workers would make them more likely to become borrowing constrained, and strengthen the efficiency rationale for taxing automation. Our specification ensures that the cost of investing in automation is smoothed over time and that dividends do not fall, which firms seem to be reluctant to do in practice (Leary and Michaely, 2011).

[27] We introduce overlapping generations because young cohorts account for a substantial share of labor reallocation across occupations (Adão et al., 2020).

[28] The mutual fund now rebates dividends Π_t^{div} to workers, instead of profits.

where $\mathcal{Y}_t^{\text{net}}(\mathbf{x})$ denotes net income and r_t is the return on bonds. Workers still face the borrowing constraint (2.9). They hold $a^{\text{birth}}(\mathbf{x}) = 0$ assets at birth.

Occupational choice. Workers choose their first occupation of employment at birth. They supply labor and are given the opportunity to move between occupations with intensity λ. Moreover, workers are subject to linearly additive taste shocks when choosing between occupations. These taste shocks are independent over time and distributed according to an Extreme Value Type-I distribution with mean 0 and scale parameter $\gamma > 0$, as is standard in the literature (Artuç et al., 2010). In particular, workers choose a *non-automated* occupation with hazard

$$\mathcal{S}_t(\mathbf{x}) = \frac{(1-\phi)\exp\left(\frac{V_t^N(\mathbf{x}'(N;\mathbf{x}))}{\gamma}\right)}{\sum_{h'} \phi^{h'} \exp\left(\frac{V_t^{h'}(\mathbf{x}'(h';\mathbf{x}))}{\gamma}\right)}, \tag{5.5}$$

where $V_t^h(\cdot)$ denotes the continuation value associated to automated ($h = A$) and non-automated ($h = N$) occupations, $\mathbf{x}'$ captures the worker's new state after they choose an occupation, and the parameter γ governs the elasticity of labor supply. Workers who reallocate go through unemployment spells which they exit at rate κ. Upon entering their new occupation, workers experience a permanent productivity loss θ. They experience this loss only the first time they reallocate.

Income. Employed workers ($e = E$) earn gross labor income

$$\mathcal{Y}_t^{\text{labor}}(\mathbf{x}) = \xi \exp(z)\, w_t^h, \tag{5.6}$$

with the productivity consisting of a permanent component (ξ) and a mean-reverting component (z). The permanent component switches from 1 to $1-\theta$ the first time a worker switches occupations. The mean-reverting component of productivity evolves as

$$dz_t = -\rho_z z_t dt + \sigma_z dW_t \tag{5.7}$$

with persistence $\rho_z^{-1} > 0$ and volatility $\sigma_z > 0$. The employment status switches to $e_t = U$ upon reallocation and reverts to $e_t = E$ upon exiting unemployment. All workers are born with $e_t = E$. As in Krueger et al. (2016), unemployed workers

$(e = U)$ receive unemployment benefits that are proportional to the gross labor income they would have earned in their previous occupation. The replacement rate is $b \in [0, 1]$. Workers claim dividends in proportion to their idiosyncratic (mean-reverting) productivity, as in Auclert et al. (2018).[29] Workers net income is

$$\mathcal{Y}_t^{\text{net}}(\mathbf{x}) = \mathcal{T}_t\left(\mathcal{Y}_t^{\text{labor}}(\mathbf{x}) + \exp(z)\,\Pi_t^{\text{div}}\right), \tag{5.8}$$

where $\mathcal{T}_t(y) = \psi_0 y^{1-\psi_1}$ captures non-linear taxation (Heathcote et al., 2017).

5.3 Policy and Equilibrium

The government's flow budget constraint is

$$dB_t = (T_t + r_t B_t - U_t - g_t)\,dt \tag{5.9}$$

where B_t is the government's asset holdings, T_t is total tax revenues, U_t is total unemployment benefits, and g_t is government spending. The resource constraint is now

$$\int c_t(\mathbf{x})\,d\pi_t + g_t + q_t x_t = G\left(y_t^A, y_t^N\right) \tag{5.10}$$

The wages are still given by (3.1). A competitive equilibrium is defined as before.

6 Quantitative Evaluation

We now use the model to evaluate the importance of our mechanism and perform policy experiments. Section 6.1 discusses the calibration. Section 6.2 describes the laissez-faire transition. Section 6.3 discusses policy interventions. Finally, Appendices B–E provide details about our numerical implementation.

6.1 Calibration

We parameterize the model using a mix of external and internal calibration. We interpret our initial stationary equilibrium (before automation) as the year 1980.

[29] This assumption implies that workers claim labor and profit income in proportion to their idiosyncratic (mean-reverting) productivity. It is the most neutral possible, as it ensures that the government has no incentives to tax (or subsidize) automation to reduce workers' income risk.

Table 6.1 shows the parameterization.

External calibration. External parameters are borrowed from the literature. The elasticity of substitution across occupations ν is 0.9 (Goos et al., 2014). The span of control parameter η is 0.15 (Basu and Fernald, 1997; Atkeson and Kehoe, 2007). The depreciation rate δ is 10%, as in Graetz and Michaels (2018). We choose an inverse elasticity of intertemporal substitution $\sigma \to 1$ as in Guerreiro et al. (2022). We set the replacement rate χ to obtain an average active lifespan of 50 years (Nuño and Moll, 2018). We pick the unemployment exit hazard parameter κ to match the average unemployment duration in the U.S., as measured by Alvarez and Shimer (2011). The productivity loss θ when moving between occupations is set to match the earnings losses in Kambourov and Manovskii (2009). As in Auclert et al. (2018), we rule out borrowing $\underline{a} = 0$. We use the annual income process estimated by Floden and Lindé (2001) using PSID data and choose the persistence ρ_z^{-1} and volatility σ_z in our continuous time model accordingly. The replacement rate when unemployed b is 0.4, following Ganong et al. (2020). Government spending relative to consumption g_t/C_t is 50% at the initial steady state. The progressivity of the tax schedule ψ_1 is 0.181, as in Heathcote et al. (2017). We choose the intercept of the tax schedule ψ_0 so that the government can finance $g_t/C_t = 0.5$ at the initial steady state. Finally, the ratio of liquidity to GDP $-B_t/Y_t$ is 0.5 at the initial and final steady states (Survey of Consumer Finances, 1980).[30] During the transition, the government adjusts liquidity B_t (and government spending g_t accordingly) so that the interest rate converges exponentially to its long-run level. The half-life is the same as the one we target internally for the wage gap across occupations (15 years).

Internal calibration. We calibrate eight parameters internally: the discount rate (ρ); the mobility hazard (λ); the scale parameter (γ); the two occupation-specific productivities (A^h); the share of automated occupations (ϕ); the final cost of investment (q^{fin}); and its convergence rate (ψ). We pick these to jointly match eight moments. The discount rate targets an annual real interest rate of 2 percent. The mobility hazard λ targets an occupational mobility rate of 10% per year at the ini-

[30] We obtain this value by adding up checkable deposits, time and savings deposits, and money market funds share (Table B.100, lines 11-13, year 1980) and divide by nominal GDP. This amount of liquidity is almost twice as high as the one in Kaplan et al. (2018), which is conservative with respect to our mechanism.

Table 6.1: Calibration

Parameter	Description	Calibration	Target / Source
Workers			
ρ	Discount rate	0.040	2% real interest rate
σ	EIS (inverse)	1	-
χ	Death rate	1/50	Average working life of 50 years
$\underline{a}$	Borrowing limit	0	Auclert et al. (2018)
Technology			
A^A, A^N	Productivities	$(0.719, 1.710)$	Initial output (1) and symmetric wages
$1-\eta$	Initial labor share	0.85	Span of control (Atkeson and Kehoe, 2007)
δ	Depreciation rate	0.1	Graetz and Michaels (2018)
ϕ	Share of automated occupations	0.537	Routine occs. employment share in 1980
q^{fin}	Final cost of investment	5.621	Final wage gap
ψ	Convergence rate of cost	0.054	Half-life of wage gap
ν	Elasticity of subst. across occs.	0.9	Goos et al. (2014)
Mobility frictions			
λ	Mobility hazard	0.364	Occupational mobility rate in 1980
$1/\kappa$	Average unemployment duration	1/3.2	Alvarez and Shimer (2011)
θ	Productivity loss from relocation	0.18	Kambourov and Manovskii (2009)
γ	Scale parameter	0.0360	Elasticity of labor supply
Government			
ψ_0	Intercept of tax schedule	0.661	Gvt spending g/C (BEA)
ψ_1	Elasticity of tax schedule	0.181	Heathcote et al. (2017)
$-B/Y$	Liquidity / GDP	0.5	Liquid assets / GDP (SCF)
Income process			
ρ_z	Mean reversion	0.09	Floden and Lindé (2001)
σ_z	Volatility	0.205	Floden and Lindé (2001)
b	Replacement rate	0.4	Ganong et al. (2020)

tial steady state, which roughly corresponds to the U.S. level in 1980 in Kambourov and Manovskii (2008). The scale parameter γ targets an elasticity of labor supply of 1 for the stock of workers (i.e., all generations) which lies between the estimates of Wiswall and Zafar (2015) and Hsieh et al. (2019).[31] The occupations' productivities $\{A^h\}$ are such that output is 1 and wages are identical across occupations at the initial steady state. The mass of automated occupations ϕ targets an employment share of 56% in routine occupations in 1980 (Bharadwaj and Dvorkin, 2019). We choose the final cost of investment q^{fin} to obtain a log wage gap of 0.45 between occupations at the final steady state (Cortes, 2016). Similarly, the convergence rate ψ targets a half-life of 15 years for the wage gap (Cortes, 2016).

Untargeted moments. The model matches well several untargeted moments (see Appendix E for details). First, the share of hand-to-mouth workers is roughly 23% at the initial steady state, which lies between the estimates of Kaplan et al. (2014) and Aguiar et al. (2020). Second, the share of routine employment 40 years into the transition (year 2020) is 39% compared to roughly 41% in the data (Bharadwaj and Dvorkin, 2019). Third, we obtain that 67% of output in occupation $h = A$ is produced by automation at the final steady state. For comparison, the McKinsey (2017) report finds that roughly 70% of output previously produced by labor could be automated in occupations most susceptible to automation (making up for 51% of initial employment, compared to 56% in our model). Fourth, the (partial equilibrium) effects of automation on employment in our model are comparable to the firm-level estimates in Bonfiglioli et al. (2022). They find that adopting automation changes employment by -0.54 log-points at the firm level, compared to -0.66 in our model (see Appendix E for details). Finally, consumption increases by only 5.3% over the first 40 years. This is consistent with the view that automation delivers relatively small TFP gains (Acemoglu and Restrepo, 2019).

6.2 Laissez-Faire

We start by simulating the laissez-faire transition. The economy is initially at its steady state with no automation ($\alpha_0 = 0$). In period $t = 0$, the cost of investing in

[31] We compute this elasticity in our model by simulating a 10% wage increase in one of the occupations and leaving the other one unchanged.

automation starts to fall and the economy converges gradually to its new steady state with positive automation.

Figure 6.1 shows this transition (solid lines). The rise in automation displaces workers and reallocates labor away from automated occupations. Despite this reallocation, wages decline gradually in automated occupations (red line) but increase in non-automated occupations (blue line) since the two occupations are complements. The wage gap widens to 0.45 with a half-life of 15 years (both are targeted moments). Finally, automated workers consume less and have steeper consumption profiles — their MRS is lower — as they are more likely to become borrowing constrained.[32]

6.3 Second Best and Welfare

We now solve for the optimal policy and quantify welfare gains. The government maximizes

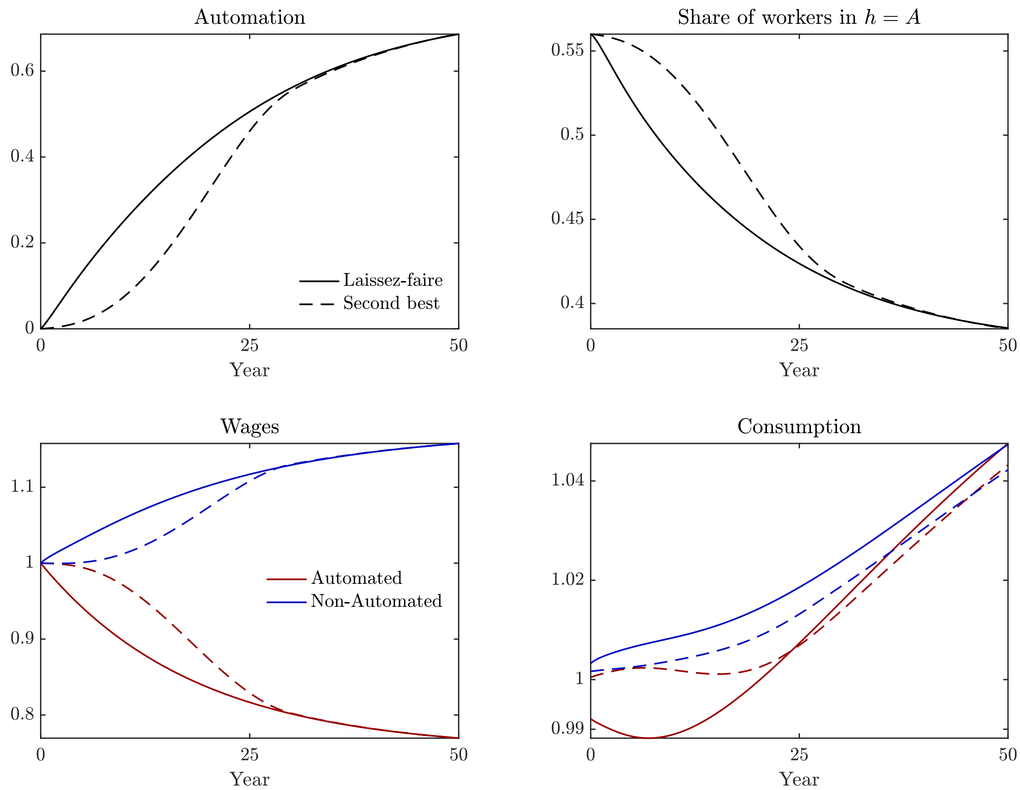
$$\mathcal{W}\left(\boldsymbol{\eta}\right) \equiv \int_{-\infty}^{+\infty} \int \eta_t\left(\mathbf{x}\right) V_t^{\text{birth}}\left(\mathbf{x}\right) d\pi_t\left(\mathbf{x}\right) dt \tag{6.1}$$

where $V_t^{\text{birth}}\left(\mathbf{x}\right)$ is the value of a worker with state $\mathbf{x}$ born in period t, and $\eta_t\left(\mathbf{x}\right)$ are Pareto weights. The government maximizes this objective by choosing taxes on investment $\{\tau_t^x\}$ along the transition. The government uses *efficiency* weights $\eta_t\left(\mathbf{x}\right)$ which are inversely related to workers' marginal utility at birth. These weights are the ones we described in our baseline model (Section 4.5) and ensure that the government has no incentive to redistribute resources. Computational details are provided in Appendix B.3.

Figure 6.1 illustrates the effects of the second best intervention (dashed lines). The optimal policy slows down automation so as to increase its half-life from 15 years at the laissez-faire to 22 years at the optimum. The speed of automation is especially slower over the first decade of the transition. There is less labor reallocation over this period, and the wage gap opens up more slowly. As anticipated in Remark 2, the optimal policy raises consumption and redistributes early on during the transition when displaced workers value it more.

[32] There are two forces at play. First, automated workers can reallocate across occupations and they go through unemployment — which incentivizes them to *borrow*. Second, the wage in automated occupations falls over time — which incentivizes automated workers to *save*. Only the first force is present in our baseline model where automation happens once and for all. This forces dominates in the quantitative model, which is consistent with the empirical evidence (Section 3.2).

Figure 6.1: Allocations



Notes: Solid curves correspond to the laissez-faire and dashed curves to the second best. 'Wages' are skill prices in the two occupations, and 'Consumption' is the average consumption by workers initially employed in a given occupation. Red and blue curves are used to denote automated and non-automated occupations / workers, respectively. Wages and consumptions are normalized by their initial steady state levels.

Table 6.2 reports the welfare gains (in consumption equivalent terms) from the second best intervention. The first column corresponds to our benchmark calibration. Automated workers benefit substantially from slower automation (0.80%). Non-automated workers are hurt by the intervention (−0.19%). Taxing automation goes a long way in improving the welfare of automated workers. At the second best, they are only worse off by −0.60% relative to non-automated workers, compared to −1.58% at the laissez-faire. The intervention lowers slightly the welfare of new generations (−0.08%) since it reduces the value of the firm and hence dividends. Overall, the policy raises social welfare by 0.20%.[33] Figure D.1 in Ap-

[33] These figures are comparable to the ones found in the heterogenous agent literature on optimal taxation (e.g., Heathcote et al., 2017) or the literature on the taxation of automation (e.g., Guerreiro et al., 2022).

Table 6.2: Welfare Gains $\Delta\mathcal{W}$ from Second Best Interventions

	Benchmark	Less liquidity	Less reallocation	More complements
Automated	0.80%	0.91%	0.93%	0.78%
Non-autom.	−0.19%	−0.22%	−0.35%	−0.21%
New gener.	−0.08%	−0.11%	−0.10%	−0.08%
Total	0.20%	0.24%	0.20%	0.19%

Note: 'Benchmark' corresponds to the gains from the optimal taxation of automation under the calibration described in Section 6.1. 'Less liquidity' and 'Less reallocation' denote alternative calibrations where we target a ratio of liquidity to GDP of 0.35 (instead of 0.5) and a separation rate of 7.2% (instead of 10%), respectively. 'More complements' denotes an alternative calibration where the elasticity of substitution across occupations is 0.76 (instead of 0.9).

pendix D plots the tax on investments $\{\tau_t^x\}$ that implement this second best. The optimal tax starts at roughly 5%, raises progressively to 7% over a decade, and then gradually declines to zero in the long-run.

Robustness checks. We then consider additional calibrations to assess the robustness of our results. First, we consider a narrower definition of liquid assets. We now target a ratio $-B_t/Y_t$ of 35% (instead of 50%).[34] All other parameters are recalibrated to match the moments described in Section 6.1. Automated workers are more likely to become borrowing constrained, and their consumption profiles are steeper. Therefore, they benefit more from slowing down automation and the total welfare gains increase. Second, we recognize that occupational mobility has decreased in recent decades. We thus target an occupational mobility rate of 7.2% (instead of 10%) following Moscarini and Vella (2002). This alternative calibration lowers the consumption of automated workers (in levels) compared to our benchmark as they reallocate less, but does not affect meaningfully the slope of their consumption profiles. Accordingly, automated workers benefit more from the intervention but the total welfare gains are mostly unchanged. Third, there is some uncertainty in the literature about the elasticity of substitution between occupations. We thus decrease it to 0.76 (instead of 0.9) based on Gregory et al. (2021) and

[34] We obtain this value by substracting consumer credit and other loans excluding mortgages (Table B.100, lines 34 and 36-37, year 1980) from our previous measure of liquidity (footnote 30).

find very similar welfare gains to our benchmark. Finally, we increase the depreciation rate to 20% (instead of 10%) to capture the fact that some forms of automation like artificial intelligence software could depreciate faster than others like robots. We also decrease the elasticity of labor supply across occupations to 0.4 (instead of 1) following Wiswall and Zafar (2015). We obtain total welfare gains of 0.19% and 0.29%, respectively.

Wage supplements. Government transfers that target automated workers could in principle be an effective tool to respond to automation. In particular, the government could provide wage supplements to automated workers — similar to Trade Adjustment Assistance for Workers (TAA) in the US. This intervention would be financed by taxing non-automated workers — a negative wage supplement. We compute the wage supplements (along the transition) that would make workers indifferent between these supplements and the tax on automation. In present discounted terms, the government would need to give $19,126 to the average automated worker, and would tax $4,622 from the average non-automated worker.[35] Assuming a workforce of 107 million in 1980, the wage supplements to automated workers would cost roughly $1.1 trillion and leave a fiscal deficit (after taxing non-automated workers) of roughly $924 billion. For comparison, the US Congress budgeted $551 million for the TAA program in 2022, or $13.4 billion in present discounted value in our model. These figures show that slowing down automation delivers welfare gains that would be costly to replicate with wage supplements alone.

7 Conclusion

We presented two novel results in economies where workers displaced by automation face reallocation and borrowing frictions. First, automation is inefficient when these frictions are sufficiently severe because there is as conflict between how firms and displaced workers value the effects of automation over time. Second, absent redistributive tools that fully alleviate borrowing frictions, the government should slow down automation while displaced workers reallocate but not tax it in the long-run. The optimal policy raises aggregate consumption and redistributes early

[35] Average earnings are $65k at the initial steady state.

on in the transition precisely when displaced workers value it more. Quantitatively, we found that slowing down automation achieves important welfare gains.

To derive sharp results and clarify the mechanisms at play, our model necessarily abstracted from many features. Some of these are worth discussing now. Tax-codes often subsidize capital and R&D expenditures on the grounds that firms face credit constraints or that there are externalities involved — features that our analysis has ignored. Thus, our results do not necessarily imply that automation technologies ought to be taxed *on net*, as is the case of autonomous vehicles used by transportation companies in Nevada or as proposed for automated cashiers in the Swiss canton of Geneva. Instead, they imply that subsidies on investment in automation should be lowered *temporarily* while the economy adjusts and displaced workers reallocate, which is similar to the lower tax credits for automation in South Korea.

Our quantitative model points to two directions for future work. First, we found that the optimal policy depends on how steep the consumption profiles of workers displaced by automation are. It would be interesting to measure these profiles and compare them to the estimates for the average US worker used in our quantitative exercises. For instance, the profiles could be steeper if automated workers are unemployed for longer while they reallocate. Second, the quantitative model is rich enough to tackle other optimal policy questions where the dynamics of labor reallocation and asset markets imperfections are relevant, such as how governments should manage declining regions or the economy's adjustment to international trade.

References

ACEMOGLU, D., A. MANERA, AND P. RESTREPO (2020): "Does the US Tax Code Favor Automation?" *Brookings Papers on Economic Activity*.

ACEMOGLU, D. AND P. RESTREPO (2018): "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, 108, 1488–1542.

——— (2019): "Automation and New Tasks: How Technology Displaces and Reinstates Labor," *Journal of Economic Perspectives*, 33, 3–30.

——— (2022): "Tasks, Automation, and the Rise in US Wage Inequality," *Forthcoming, Econometrica*.

ADÃO, R., M. BERAJA, AND N. PANDALAI-NAYAR (2020): "Technological Transitions with Skill Heterogeneity Across Generations," Tech. rep., National Bureau of Economic Research.

AGHION, P. AND O. J. BLANCHARD (1994): "On the Speed of Transition in Central Europe," in *NBER Macroeconomics Annual 1994, Volume 9*, National Bureau of Economic Research, Inc, NBER Chapters, 283–330.

AGUIAR, M. A., M. AMADOR, AND C. ARELLANO (2021): "Micro Risks and Pareto Improving Policies with Low Interest Rates," NBER Working Paper 28996.

AGUIAR, M. A., M. BILS, AND C. BOAR (2020): "Who Are the Hand-to-Mouth?" NBER Working Papers 26643, National Bureau of Economic Research, Inc.

AIYAGARI, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 109, 659–684.

——— (1995): "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy*, 103, 1158–1175.

AIYAGARI, S. R. AND E. R. MCGRATTAN (1998): "The Optimum Quantity of Debt," *Journal of Monetary Economics*, 42, 447–469.

ALVAREZ, F. AND R. SHIMER (2011): "Search and Rest Unemployment," *Econometrica*, 79, 75–122.

ARTUÇ, E., S. CHAUDHURI, AND J. MCLAREN (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100, 1008–1045.

ATKESON, A. AND P. J. KEHOE (2007): "Modeling the Transition to a New Economy: Lessons from Two Technological Revolutions," *American Economic Review*, 97, 64–88.

AUCLERT, A., M. ROGNLIE, AND L. STRAUB (2018): "The Intertemporal Keynesian Cross," NBER Working Paper 25020.

BASU, S. AND J. G. FERNALD (1997): "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, 105, 249–283.

BERAJA, M. AND N. ZORZI (2022): "Inefficient Automation," Working Paper 30154, National Bureau of Economic Research.

BHANDARI, A., D. EVANS, M. GOLOSOV, AND T. SARGENT (2021): "Efficiency, Insurance, and Redistribution Effects of Government Policies," *Working paper*.

BHARADWAJ, A. AND M. A. DVORKIN (2019): "The Rise of Automation: How Robots May Impact the U.S. Labor Market," Regionaal Economist Second Quarter 2019, Federal Reserve Bank of St. Louis.

BONFIGLIOLI, A., R. CRINO, H. FADINGER, AND G. GANCIA (2022): "Robot Imports and Firm-Level Outcomes," *Mimeo*.

BRAXTON, J. C. AND B. TASKA (2023): "Technological Change and the Consequences of Job Loss," *American Economic Review*, 113, 279–316.

CABALLERO, R. J. AND M. L. HAMMOUR (1996): "On the Timing and Efficiency of Creative Destruction," *The Quarterly Journal of Economics*, 111, 805–852.

CALVO, G. A. AND M. OBSTFELD (1988): "Optimal Time-Consistent Fiscal Policy with Finite Lifetimes," *Econometrica*, 56, 411–432.

CARD, D., J. KLUVE, AND A. WEBER (2018): "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations," *Journal of the European Economic Association*, 16, 894–931.

CARROLL, C. D. (2006): "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics Letters*, 91, 312–320.

CHAMLEY, C. (1986): "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54, 607–622.

CHETTY, R. (2008): "Moral Hazard versus Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, 116, 173–234.

CONESA, J. C., S. KITAO, AND D. KRUEGER (2009): "Taxing Capital? Not a Bad Idea after All!" *American Economic Review*, 99, 25–48.

CORTES, G. M. (2016): "Where have the middle-wage workers gone? A study of polarization using panel data," *Journal of Labor Economics*, 34, 63–105.

COSTINOT, A. AND I. WERNING (2022): "Robots, Trade, and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation," *Forthcoming, The Review of Economic Studies*.

CRÉPON, B. AND G. J. VAN DEN BERG (2016): "Active Labor Market Policies," *Annual Review of Economics*, 8, 521–546.

DÁVILA, E. AND A. KORINEK (2018): "Pecuniary Externalities in Economies with

Financial Frictions," *Review of Economic Studies*, 85, 352–395.

DÁVILA, E. AND A. SCHAAB (2022): "Welfare Assessments with Heterogeneous Individuals," *Available at SSRN 4102027*.

DÁVILA, J., J. H. HONG, P. KRUSELL, AND J.-V. RÍOS-RULL (2012): "Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks," *Econometrica*, 80, 2431–2467.

DIAMOND, P. A. (1965): "National Debt in a Neoclassical Growth Model," *The American Economic Review*, 55, 1126–1150.

FLODEN, M. AND J. LINDÉ (2001): "Idiosyncratic Risk in the United States and Sweden: Is There a Role for Government Insurance?" *Review of Economic Dynamics*, 4, 406–437.

GANONG, P., P. NOEL, AND J. VAVRA (2020): "US Unemployment Insurance Replacement Rates During the Pandemic," *Journal of Public Economics*, 191.

GEANAKOPLOS, J. AND H. M. POLEMARCHAKIS (1985): "Existence, Regularity, and Constrained Suboptimality of Competitive Allocations When the Asset Market Is Incomplete," Cowles Foundation Discussion Papers 764.

GOLOSOV, M. AND A. TSYVINSKI (2006): "Designing Optimal Disability Insurance: A Case for Asset Testing," *Journal of Political Economy*, 114, 257–279.

GOOS, M., A. MANNING, AND A. SALOMONS (2014): "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring," *American Economic Review*, 104, 2509–2526.

GRAETZ, G. AND G. MICHAELS (2018): "Robots at Work," *The Review of Economics and Statistics*, 100, 753–768.

GREGORY, T., A. SALOMONS, AND U. ZIERAHN (2021): "Racing With or Against the Machine? Evidence on the Role of Trade in Europe," *Journal of the European Economic Association*, 20, 869–906.

GUERREIRO, J., S. REBELO, AND P. TELES (2022): "Should Robots Be Taxed? [Skills, Tasks and Technologies: Implications for Employment and Earnings]," *Review of Economic Studies*, 89, 279–311.

GUNER, N., R. KAYGUSUZ, AND G. VENTURA (2021): "Rethinking the Welfare State," CEPR Discussion Papers 16275, C.E.P.R. Discussion Papers.

HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2017): "Optimal Tax Progressivity: An Analytical Framework," *The Quarterly Journal of Economics*, 132, 1693–1754.

HÉMOUS, D. AND M. OLSEN (2022): "The Rise of the Machines: Automation, Horizontal Innovation, and Income Inequality," *American Economic Journal: Macroeconomics*, 14, 179–223.

HSIEH, C.-T., E. HURST, C. I. JONES, AND P. J. KLENOW (2019): "The Allocation of Talent and U.S. Economic Growth," *Econometrica*, 87, 1439–1474.

HUGGETT, M. (1993): "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, 17, 953–969.

HUMLUM, A. (2019): "Robot Adoption and Labor Market Dynamics," *Mimeo*.

HUMLUM, A., P. P. JØRGENSEN, AND J. R. MUNCH (2022): "Changing Tracks: Human Capital Investment after Loss of Earning Ability," .

HUMMELS, D., J. R. MUNCH, AND C. XIANG (2018): "Offshoring and labor markets," *Journal of Economic Literature*, 56, 981–1028.

IMROHOROGLU, A., S. IMROHOROGLU, AND D. H. JOINES (1995): "A Life Cycle Analysis of Social Security," *Economic Theory*, 6, 83–114.

ITSKHOKI, O. AND B. MOLL (2019): "Optimal Development Policies With Financial Frictions," *Econometrica*, 87, 139–173.

JACOBSON, L. S., R. J. LALONDE, AND D. G. SULLIVAN (1993): "Earnings Losses of Displaced Workers," *American Economic Review*, 83, 685–709.

——— (2005): "Is Retraining Displaced Workers a Good Investment?" *Economic Perspectives*, 29, 47–66.

JAPPELLI, T. AND L. PISTAFERRI (2017): *The Economics of Consumption: Theory and Evidence*, no. 9780199383153 in OUP Catalogue, Oxford University Press.

JUDD, K. L. (1985): "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 28, 59–83.

KAMBOUROV, G. AND I. MANOVSKII (2008): "Rising Occupational And Industry Mobility In The United States: 1968-97," *International Economic Review*, 49, 41–79.

——— (2009): "Occupational Specificity Of Human Capital," *International Economic Review*, 50, 63–115.

KAPLAN, G., B. MOLL, AND G. L. VIOLANTE (2018): "Monetary Policy According to HANK," *American Economic Review*, 108, 697–743.

KAPLAN, G., G. L. VIOLANTE, AND J. WEIDNER (2014): "The Wealthy Hand-to-Mouth," *Brookings Papers on Economic Activity*, 45, 77–153.

KORINEK, A. AND J. E. STIGLITZ (2020): "Steering Technological Progress," Mimeo, University of Virginia.

KOVACEV, R. (2020): "A Taxing Dilemma: Robot Taxes and the Challenges of Effective Taxation of AI, Automation and Robotics in the Fourth Industrial Revolution," *Ohio St. Tech. LJ*, 16, 182.

KRUEGER, D., K. MITMAN, AND F. PERRI (2016): "Macroeconomics and Household Heterogeneity," vol. 2 of *Handbook of Macroeconomics*, chap. 0, 843–921.

LANDAIS, C. AND J. SPINNEWIJN (2021): "The Value of Unemployment Insurance," *The Review of Economic Studies*, 88, 3041–3085.

LEARY, M. T. AND R. MICHAELY (2011): "Determinants of Dividend Smoothing: Empirical Evidence," *The Review of Financial Studies*, 24, 3197–3249.

LEE, D. AND K. I. WOLPIN (2006): "Intersectoral Labor Mobility and the Growth of the Service Sector," *Econometrica*, 74, 1–46.

LOHR, S. (2022): "Economists Pin More Blame on Tech for Rising Inequality," *The New York Times*.

MANKIW, N. G. AND S. ZELDES (1991): "The Consumption of Stockholders and Nonstockholders," *Journal of Financial Economics*, 29, 97–112.

MARTINEZ, J. (2018): "Automation, Growth and Factor Shares," in *SED Meeting Papers*, Society for Economic Dynamics, vol. 736.

MCKINSEY (2017): "A Future that Works: Automation, Employment, and Productivity," *McKinsey Global Institute*.

MOLL, B., L. RACHEL, AND P. RESTREPO (2021): "Uneven Growth: Automation's Impact on Income and Wealth Inequality," NBER Working Papers 28440.

MOSCARINI, G. AND F. VELLA (2002): "Aggregate worker reallocation and occupational mobility in the United States: 1971-2000," IFS Working Papers W02/18, Institute for Fiscal Studies.

MOSCARINI, G. AND F. G. VELLA (2008): "Occupational Mobility and the Business Cycle," NBER Working Papers 13819.

MUSSA, M. (1984): "The Adjustment Process and the Timing of Trade Liberalization," NBER Working Papers 1458, National Bureau of Economic Research, Inc.

NEARY, J. P. (1982): "Intersectoral Capital Mobility, Wage Stickiness, and the Case for Adjustment Assistance," in *Import Competition and Response*, National Bureau of Economic Research, Inc, NBER Chapters, 39–72.

NUÑO, G. AND B. MOLL (2018): "Social optima in economies with heterogeneous agents," *Review of Economic Dynamics*, 28, 150–180.

PIKETTY, T. AND E. SAEZ (2013): "Chapter 7 - Optimal Labor Income Taxation," vol. 5 of *Handbook of Public Economics*, 391–474.

POTERBA, J. M. (1988): "Are Consumers Forward Looking? Evidence from Fiscal Experiments," *American Economic Review*, 78, 413–418.

——— (2003): "Employer Stock and 401(k) Plans," *American Economic Review*, 93, 398–404.

ROUWENHORST, G. (1995): *Asset Pricing Implications of Equilibrium Business Cycle Models'*, vol. Frontiers of Business Cycle Research, Princeton University Press.

SCHOCHET, P. Z., R. D'AMICO, J. BERK, S. DOLFIN, AND N. WOZNY (2012): "Estimated Impacts for Participants in the Trade Adjustment Assistance (TAA) Program Under the 2002 Amendments," Tech. rep., Mathematica Policy Research.

SULLIVAN, J. X. (2008): "Borrowing During Unemployment: Unsecured Debt as a Safety Net," *Journal of Human Resources*, 43, 383–412.

THUEMMEL, U. (2018): "Optimal Taxation of Robots," CESifo Working Paper 7317.

WERNING, I. (2015): "Incomplete Markets and Aggregate Demand," NBER Working Papers 21448, National Bureau of Economic Research, Inc.

WISWALL, M. AND B. ZAFAR (2015): "Determinants of College Major Choice: Identification using an Information Experiment," *Review of Economic Studies*, 82, 791–824.

YOUNG, E. R. (2010): "Solving the incomplete markets model with aggregate uncertainty using the Krusell-Smith algorithm and non-stochastic simulations," *Journal of Economic Dynamics and Control*, 34, 36–41.

Online Appendix for:
Inefficient Automation

This online appendix contains the proofs and derivations of all theoretical results for the article "Inefficient Automation," as well as a detailed description of the quantitative model and how it is solved numerically.

Any references to equations, figures, tables, assumptions, propositions, lemmas, or sections that are not preceded "A.," "B.," "C." or "D." refer to the main article.

A Proofs and Derivations

A.1 Proof of Lemma 1

Fix some period $T \geq 0$. Consider the decision of automated workers to reallocate, i.e. the choice of $\{m_t\}$. Using a standard variational argument, it is optimal to reallocate for all workers who are able to $(m_t = 1)$ if and only if the present discounted value of their labor income is higher in non-automated occupations

$$\int_T^{+\infty} \exp\left(-\rho\left(t-T\right)\right) u'\left(c_t^A\right) \Delta_t dt > 0, \tag{A.1}$$

where

$$\Delta_t \equiv (1-\theta)\left[\iota\left(1-\exp\left(-\kappa\left(t-T\right)\right)\right)+1-\iota\right] w_t^N - w_t^A \tag{A.2}$$

captures the marginal increase in output from reallocating an additional worker, since $w_t^h = \partial_h G^\star\left(\mu^A, \mu^N; \alpha\right)$ in equilibrium. These workers do not reallocate $(m_t = 0)$ if and only if the inequality (A.1) is reversed. Any $m_t \in [0,1]$ is optimal otherwise.

In equilibrium, reallocation takes the following form. Workers reallocate until $T^{\mathrm{LF}} \geq 0$, i.e., $m_t = 1$ for all $t \in \left[0, T^{\mathrm{LF}}\right)$, and they stop reallocating afterwards, i.e., $m_t = 0$ for all $t \geq T^{\mathrm{LF}}$. The reason is that the wage in automated occupations w_t^A increases over time as workers leave this occupation (by decreasing returns), and the wage in non-automated occupations w_t^N decreases as workers enter this occupation.

We next show that reallocation does take place in equilibrium, i.e., $T^{\mathrm{LF}} > 0$. It suffices to show that workers find it optimal to reallocate at $t = 0$. That is,

$$\int_0^{+\infty} (1-\theta)\left[\iota\left(1-\exp\left(-\kappa t\right)\right)+1-\iota\right] \frac{\exp\left(-\rho t\right) u'\left(\tilde{c}_t^A\right) \tilde{w}_t^N}{\int_0^{+\infty} \exp\left(-\rho s\right) u'\left(\tilde{c}_s^A\right) \tilde{w}_s^A ds} dt > 1 \tag{A.3}$$

where $\left\{\tilde{c}_t^A\right\}$ and $\left\{\tilde{w}_t^h\right\}$ are counterfactual sequences of consumption and wages associated with $T = 0$ and $\alpha = \alpha^{\mathrm{LF}}$. Consumption and wages are constant over time when $T = 0$, so the inequality (A.3) holds if and only if

$$\frac{(1-\theta)\left(1-\iota\right)\partial_N G^\star\left(\frac{1}{2},\frac{1}{2};\alpha\right)}{\partial_A G^\star\left(\frac{1}{2},\frac{1}{2};\alpha\right)} \frac{\rho\left(1-\iota\right)+\kappa}{\left(1-\iota\right)\left(\rho+\kappa\right)} > 1, \tag{A.4}$$

where $\alpha = \alpha^{\mathrm{LF}}$. This necessarily holds by Assumption 3 when $Z^\star$ is sufficiently small since $\partial_N G^\star\left(\frac{1}{2}, \frac{1}{2}; \alpha\right) > \partial_A G^\star\left(\frac{1}{2}, \frac{1}{2}; \alpha\right)$ with automation $\alpha > 0$ by Assumption 1.

A.2 Proof of Lemma 2

We begin by showing that automated workers borrow and non-automated workers save in equilibrium. We then show that automated workers become borrowing constrained when borrowing and reallocation frictions are sufficiently severe, and characterize the threshold $a^\star(\lambda, \kappa)$. We focus on the case $\underline{a} < 0$ since the statement is obviously true in the case where $\underline{a} = 0$.

Assets. It suffices to prove that $da_t^N \geq da_t^A$ for any period t where $a_t^N = a_t^A$ with strict inequality in period $t = 0$. The reason is that the equilibrium is continuous in time t, so the sequence of assets of automated and non-automated would intersect before the inequality reverses. This would imply that automated workers borrow and non-automated workers save as $a_t^N + a_t^A = 0$ in equilibrium.

To derive a contradiction, suppose instead that $da_t^N < da_t^A$ when $a_t^N = a_t^A = 0$. Then, there exists some $S > t$ such that $a_S^A > 0$ and $a_S^N < 0$ but all workers are still unconstrained $a_S^h > \underline{a}$. In this case, workers' consumptions satisfy the Euler equation

$$c_s^h = c_t^h \exp\left(\frac{1}{\sigma}\left(\int_t^s (r_\tau - \rho)\, d\tau\right)\right) \tag{A.5}$$

for all $s \in [t, S)$. Using the market clearing condition (2.13), it must also be that

$$\exp\left(\frac{1}{\sigma}\left(\int_t^s (r_\tau - \rho)\, d\tau\right)\right) = \frac{\frac{1}{2}\left(c_s^A + c_s^N\right)}{\frac{1}{2}\left(c_t^A + c_t^N\right)} = \frac{C_s}{C_t} \equiv \frac{G^\star\left(\mu_s^A, \mu_s^N; \alpha\right)}{G^\star\left(\mu_t^A, \mu_t^N; \alpha\right)}, \tag{A.6}$$

for all $s \in [t, S)$. Using the budget constraint (2.7), consumption is

$$c_t^h = \frac{\int_t^S \exp\left(-\int_t^s r_\tau d\tau\right)\left(\hat{\mathcal{Y}}_s^h + \Pi_s\right) ds + a_t^h - \exp\left(-\int_t^S r_\tau d\tau\right) a_S^h}{\int_t^S \exp\left(-\int_t^s r_\tau d\tau\right) \exp\left(\frac{1}{\sigma}\int_t^s (r_\tau - \rho)\, d\tau\right) ds}, \tag{A.7}$$

so assets accumulate according to

$$da_t^h = \left(\hat{\mathcal{Y}}_t^h + \Pi_t - \frac{\int_t^S \exp\left(-\int_t^s r_\tau d\tau\right)\left(\hat{\mathcal{Y}}_s^h + \Pi_s\right) ds}{\int_t^S \exp\left(-\int_t^s r_\tau d\tau\right) \exp\left(\frac{1}{\sigma}\int_t^s \left(r_\tau - \rho\right) d\tau\right) ds} \right.$$
$$\left. + \Gamma_{t,S} a_t^h - \Gamma_{t,S}^\star a_S^h \right) dt \tag{A.8}$$

for some $\Gamma_{t,S}, \Gamma_{t,S}^\star > 0$ that depend on the sequence of interest rates. Using (A.8),

$$\frac{d\left(a_t^N - a_t^A\right)}{C_t} = \left(z_t - \frac{\int_t^S \exp\left(-\int_t^s r_\tau d\tau\right) \frac{C_s}{C_t} z_s ds}{\int_t^S \exp\left(-\int_t^s r_\tau d\tau\right) \exp\left(\frac{1}{\sigma}\int_t^s \left(r_\tau - \rho\right) d\tau\right) ds} dt \right.$$
$$\left. - \Gamma_{t,S}^\star \left(\frac{a_S^N - a_S^A}{C_t} \right) \right) dt \tag{A.9}$$

when $a_t^N = a_t^A = 0$, with flows $z_t \equiv \left(\hat{\mathcal{Y}}_t^N - \hat{\mathcal{Y}}_t^A\right)/C_t$. Using (A.6),

$$\frac{d\left(a_t^N - a_t^A\right)}{C_t} = \left(z_t - \int_t^S \psi_{t,s} z_s ds - \Gamma_{t,S}^\star \left(\frac{a_S^N - a_S^A}{C_t} \right) \right) dt, \tag{A.10}$$

with weights

$$\psi_{t,s} \equiv \frac{\exp\left(-\rho\left(s-t\right)\right)\left(\frac{C_s}{C_t}\right)^{1-\sigma}}{\int_t^S \exp\left(-\rho\left(\tau-t\right)\right)\left(\frac{C_\tau}{C_t}\right)^{1-\sigma} d\tau} > 0 \tag{A.11}$$

that integrate to $\int_t^S \psi_{t,s} ds = 1$. As we will establish at the end of this appendix, $\{z_s\}$ is positive and decreases over time. The reason is twofold. First, the labor income of automated workers is lower than that of non-automated workers, and the former increases over time while the latter decreases. Second, aggregate consumption grows over time too. Therefore, $z_t - \int_t^S \psi_{t,s} z_s ds > 0$. Furthermore, $a_S^N < a_S^A$ under our postulate. It follows that $d\left(a_t^N - a_t^A\right) > 0$. This contradicts our postulate that $da_t^N < da_t^A$. This establishes that $da_t^N \geq da_t^A$ when $a_t^N = a_t^A = 0$. Repeating the steps above, the inequality is strict $da_t^N > da_t^A$ after the shock $t = 0$. This shows that automated workers borrow in equilibrium.

Threshold $a^{\star}(\lambda, \kappa)$. Integrating the budget constraint (2.7) over time and using (A.5) gives the assets of automated workers if they were never to become borrowing constrained

$$a_t^A = \int_0^t \exp\left(\int_s^t r_\tau d\tau\right)\left[\hat{\mathcal{Y}}_s^A + \Pi_s - c_0^A \exp\left(\frac{1}{\sigma}\int_0^s (r_\tau - \rho)\, d\tau\right)\right] ds. \quad \text{(A.12)}$$

The sequence $\{a_t^A\}$ depends on reallocation frictions (λ, κ) but not the borrowing limit $\underline{a}$. Let $a^{\star}(\lambda, \kappa) \equiv \inf_t a_t^A$ be the lowest value attained by this sequence. If the borrowing limit is sufficiently tight that $\underline{a} > a^{\star}(\lambda, \kappa)$, then automated workers would become borrowing constrained in equilibrium. This shows that $\underline{a} > a^{\star}(\lambda, \kappa)$ is a sufficient condition for borrowing constraints to bind. It is also a necessary condition because, if borrowing constraints bind, then it must be that the borrowing limit $\underline{a}$ is above $\inf_t a_t^A$. Non-automated workers never become borrowing constrained since they save in equilibrium.

Finally, we show that $a^{\star}(\lambda, \kappa) < 0$ (i.e., borrowing constraints can bind) if and only if reallocation is slow $(1/\lambda > 0 \text{ or } 1/\kappa > 0)$. To prove sufficiency, note that the model is static when reallocation is instantaneous $(1/\lambda \to 0 \text{ and } 1/\kappa \to 0)$. Then, all labor income and profit changes are permanent, automated workers do not borrow, and therefore $a^{\star}(\lambda, \kappa) \equiv \inf_t a_t^A \to 0$. To prove necessity, note that automated workers borrow $a^{\star}(\lambda, \kappa) \equiv \inf_t a_t^A < 0$ when reallocation is slow $1/\lambda > 0$ or $1/\kappa > 0$. The reason is that $\{z_s\}$ is strictly positive and strictly decreasing over time so that automated workers borrow by (A.10). In this case, there is always a (small) borrowing limit $\underline{a} > 0$ such that automated workers become borrowing constrained.

Assumption 3. We have supposed so far that the sequence $z_t \equiv \left(\hat{\mathcal{Y}}_t^N - \hat{\mathcal{Y}}_t^A\right)/C_t$ is positive and decreases over time. The fact that $z_t > 0$ follows directly from Assumption 2 and Lemma 1. That is, automation drives a wedge between the marginal productivities of labor across occupations, and reallocation stops before the wages are fully equalized. As we show below, a sufficient condition for z_t to decrease over time is that the expected unemployment duration $1/\kappa$ is sufficiently short that output still increases over time despite workers becoming unemployed.

Output increases over time when

$$\partial_t G^\star\left(\mu_t^A, \mu_t^N; \alpha\right) = \partial_A G\left(\cdot\right)\partial_t\mu_t^A + \partial_N G\left(\cdot\right)\partial_t\mu_t^N > 0, \tag{A.13}$$

with $\partial_t\mu_t^h$ given by the effective labor supplies (3.4)–(3.5). The condition (A.13) holds in the limit where the productivity loss of reallocation and the duration of unemployment spells are sufficiently small $1-(1-\theta)(1-1/\kappa)\to 0$ since $\partial_t\mu_t^N = -\partial_t\mu_t^A > 0$ in this case. Note that μ_t^A, μ_t^N and α are continuous in $(\theta, 1/\kappa)$ at the laissez-faire. Therefore, there exists some threshold $Z^\star > 0$ such that (A.13) still holds for all $(\theta, 1/\kappa)$ such that $1-(1-\theta)(1-1/\kappa) < Z^\star$.

It remains to show that the sequence $\{z_t\}$ decreases over time when $1-(1-\theta)\times(1-1/\kappa) < Z^\star$. It suffices to show that $\partial_t\left(\hat{\mathcal{Y}}_t^N - \hat{\mathcal{Y}}_t^A\right) < 0$, as output and consumption C_t increase over time when this condition holds. Using labor incomes (2.8) and the effective labor supplies (3.4)–(3.5),

$$\frac{1}{2}\left(\hat{\mathcal{Y}}_t^N - \hat{\mathcal{Y}}_t^A\right) = -\,\partial_A G\left(\cdot\right)\mu_t^A + \left(1-\mu_t^N\right)\partial_N G\left(\cdot\right). \tag{A.14}$$

Therefore,

$$\begin{aligned}\frac{1}{2}\partial_t\left(\hat{\mathcal{Y}}_t^N - \hat{\mathcal{Y}}_t^A\right) = &-\left\{\partial^2_{AA}G\left(\cdot\right)\partial_t\mu_t^A + \partial^2_{AN}G\left(\cdot\right)\partial_t\mu_t^N\right\}\mu_t^A - \partial_A G\left(\cdot\right)\partial_t\mu_t^A \\ &+\left(1-\mu_t^N\right)\left\{\partial^2_{NA}G\left(\cdot\right)\partial_t\mu_t^A + \partial^2_{NN}G\left(\cdot\right)\partial_t\mu_t^N\right\} \\ &-\partial_N G\left(\cdot\right)\partial_t\mu_t^N.\end{aligned} \tag{A.15}$$

And so,

$$\frac{1}{2}\partial_t\left(\hat{\mathcal{Y}}_t^N - \hat{\mathcal{Y}}_t^A\right) < -\left\{\partial_A G\left(\cdot\right)\partial_t\mu_t^A + \partial_N G\left(\cdot\right)\partial_t\mu_t^N\right\} \tag{A.16}$$

using $\partial^2_{AA}G(\cdot) < 0$ and $\partial^2_{NN}G(\cdot) < 0$ since G is neoclassical, $\partial^2_{AN}G(\cdot) > 0$ by Assumption 1, and $\partial_t\mu_t^A < 0$ and $\partial_t\mu_t^N > 0$ in equilibrium. Thus, $\partial_t\left(\hat{\mathcal{Y}}_t^N - \hat{\mathcal{Y}}_t^A\right) < 0$ in the limit $1-(1-\theta)(1-1/\kappa)\to 0$ since $\partial_t\mu_t^N = -\partial_t\mu_t^A > 0$ in this case, using the fact that $\partial_A G(\cdot) < \partial_N G(\cdot)$ in equilibrium. By continuity of the equilibrium, we still have $\partial_t\left(\hat{\mathcal{Y}}_t^N - \hat{\mathcal{Y}}_t^A\right) < 0$ when $1-(1-\theta)\times(1-1/\kappa) < Z^\star$ for $Z^\star$ small enough. Taken together, the inequalities (A.13) and (A.16) imply that $z_t = \left(\hat{\mathcal{Y}}_t^N - \hat{\mathcal{Y}}_t^A\right)/C_t$ decreases over time, which completes the proof.

A.3 Proof of Lemma 3

In equilibrium, there is no arbitrage between bonds and equity $Q_t = \exp\left(-\int_0^t r_s ds\right)$, since workers can trade both. Appendix A.2 has shown that non-automated workers are on their Euler equation $\exp(-\rho t)\, u'\left(c_t^N\right)/u'\left(c_0^N\right) = \exp\left(-\int_0^t r_s ds\right)$. Next, we show that the firm's automation choice is interior and unique. Using a standard variational argument, a necessary condition for an interior optimum is

$$\int_0^{+\infty} \exp\left(-\int_0^t r_s ds\right) \frac{\partial}{\partial \alpha} \Pi_t(\alpha)\, dt = 0. \tag{A.17}$$

Furthermore, the following envelope condition applies

$$\frac{d}{d\alpha} \Pi_t(\alpha) = \partial_\alpha G^\star(\cdot). \tag{A.18}$$

Therefore, the following condition is necessary

$$\int_0^{+\infty} \exp\left(-\int_0^t r_s ds\right) \partial_\alpha G^\star(\cdot) = 0. \tag{A.19}$$

It is also sufficient by Assumption 2, and the solution is unique and interior.

Finally, we show that $\Delta_t^\star$ increases over time in equilibrium. By definition,

$$\Delta_t^\star \equiv \partial_\alpha G^\star(\cdot) \equiv \partial_A G\left(F\left(\mu_t^A; \alpha\right), F\left(\mu_t^N; 0\right)\right) \partial_\alpha F\left(\mu_t^A; \alpha\right) - \mathcal{C}'(\alpha). \tag{A.20}$$

Therefore,

$$\partial_t \Delta_t^\star \geq \partial_{AN}^2 G\left(F\left(\mu_t^A; \alpha\right), F\left(\mu_t^N; 0\right)\right) \partial_\mu F\left(\mu_t^N; 0\right) \partial_\alpha F\left(\mu_t^A; \alpha\right) \partial_t \mu_t^N, \tag{A.21}$$

using $\partial_{AA}^2 G(\cdot) < 0$ as G is neoclassical, $\partial_{\alpha\mu}^2 F(\cdot) < 0$ by Assumption 1, and $\partial_t \mu_t^A < 0$. It follows that $\partial_t \Delta_t^\star > 0$ as $\partial_{AN}^2 G(\cdot) > 0$ by Assumption 1 and $\partial_t \mu_t^N > 0$.

A.4 Proof of Proposition 1

The result consists of two parts. First, we prove that the equilibrium is generically constrained inefficient by showing that there is a Pareto improvement. Second, we show that the Pareto improvement involves taxing automation ($\delta\alpha < 0$) when unemployment spells are sufficiently short (Assumption 3).

Part I (generic constrained inefficiency). The changes in welfare starting from the laissez-faire are

$$\begin{aligned}\delta U^h =& \delta\alpha \times \int_0^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^h\right)\left(\Delta_t^\star + \Sigma_t^{h,\star}\right) dt \\ &+ \delta T \times \int_T^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^h\right)\left(\Delta_t + \Sigma_t^h\right) dt,\end{aligned}$$

for automated and non-automated workers $h = A, N$. Using the definition of earnings (2.8) and profits (2.5), the distributional terms are[36]

$$\Delta_t^\star + \Sigma_t^{N,\star} = \hat{w}_t^{N,\star} + \Delta_t^\star - \sum_h \mu_t^h \hat{w}_t^{h,\star} \tag{A.22}$$

$$\Delta_t + \Sigma_t^N = \hat{w}_t^N - \sum_h \mu_t^h \hat{w}_t^h \ , \tag{A.23}$$

for non-automated workers, where the sequences $\{ \hat{w}_t^h \}$ and $\{ \hat{w}_t^{h,\star} \}$ are the variation of wages $w_t^h \equiv \partial_h G\left(\cdot\right)$ after a variation in T and α, respectively. The distributional terms for automated workers follow from $\Sigma_t^{A,(\star)} + \Sigma_t^{N,(\star)} = 0$.

There exists a variation $(\delta\alpha, \delta T)$ with $\delta\alpha \neq 0$ that results in $\delta U^N = 0$ and $\delta U^A > 0$ if and only if

$$\begin{aligned}&\int_0^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^A\right)\left(\Delta_t^\star + \Sigma_t^{A,\star}\right) dt \\ &\neq \underbrace{\frac{\int_0^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^N\right)\left(\Delta_t^\star + \Sigma_t^{N,\star}\right) dt}{\int_T^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^N\right)\left(\Delta_t + \Sigma_t^N\right) dt}}_{\delta T/\delta\alpha \text{ that leaves } N \text{ worker indifferent}} \int_T^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^A\right)\left(\Delta_t + \Sigma_t^A\right).\end{aligned} \tag{A.24}$$

Equivalently,

$$\begin{aligned}&\int_0^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^A\right)\left(\Delta_t^\star + \Sigma_t^{A,\star}\right) dt \\ &\neq \underbrace{\frac{\int_T^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^A\right)\left(\Delta_t + \Sigma_t^A\right) dt}{\int_T^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^N\right)\left(\Delta_t + \Sigma_t^N\right) dt}}_{\equiv\Omega} \int_0^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^N\right)\left(\Delta_t^\star + \Sigma_t^{N,\star}\right) dt.\end{aligned} \tag{A.25}$$

[36] These expressions already use the fact that the firm chooses labor demand optimally.

We now show that (A.25) holds with inequality *generically*. Suppose that the expression does hold with equality. Then, there exists a perturbation of the production function $G^{\star,\prime} = \mathcal{G}(G^{\star}, \epsilon)$ (with $\mathcal{G}(G^{\star}, \epsilon) \to G^{\star}$ uniformly as $\epsilon \to 0$) and a threshold $\bar{\epsilon} > 0$ such that the expression does not hold with equality in this alternative economy, for all $0 < \epsilon \leq \bar{\epsilon}$. One such perturbation is

$$\mathcal{G}(G^{\star}, \epsilon) = G^{\star} + \epsilon g\left(\mu^{A}, \mu^{N}; \alpha\right) \tag{A.26}$$

with

$$g\left(\mu_t^{A}, \mu_t^{N}; \alpha\right) \equiv \left\{\mu_t^{A} - z\right\}\left(\alpha - \alpha^{\mathrm{LF}}\right), \tag{A.27}$$

where z is chosen so that

$$\int_0^{+\infty} \exp(-\rho t)\, u'\left(c_t^{N}\right)\left(\mu_t^{A} - z\right) dt = 0. \tag{A.28}$$

One can easily verify that all equilibrium conditions (Lemmas 1 and 3 and the resource constraint) are still satisfied after a perturbation $\epsilon > 0$ when evaluated at the laissez-faire. Therefore, the laissez-faire allocation is unchanged. Moreover, this perturbation ensures that (A.25) holds with inequality. To see that, note that Ω in (A.25) is unchanged after the perturbation $\epsilon > 0$. The reason is that the terms in $\Delta_t + \Sigma_t^{h}$ are unaffected since (A.23) depends on the second order derivatives of $G^{\star}$ with respect to labor $\left(\mu^{A}, \mu^{N}\right)$, while the perturbation (A.26)–(A.27) is linear in these variables.

Regarding the other terms on the left-hand and right-hand sides of (A.25), they change differently after the perturbation. Let

$$\Gamma^{h,\star} \equiv \frac{d}{d\epsilon}\int_0^{+\infty} \exp(-\rho t)\, u'\left(c_t^{h}\right)\left(\Delta_t^{\star} + \Sigma_t^{h,\star}\right) dt. \tag{A.29}$$

We have

$$\Gamma^{A,\star} = \int_0^{+\infty} \exp(-\rho t)\, u'\left(c_t^{A}\right)\left\{2\mu_t^{A} - z\right\} dt > 0 \tag{A.30}$$

$$\Gamma^{N,\star} = -\int_0^{+\infty} \exp(-\rho t)\, u'\left(c_t^{N}\right) z dt < 0, \tag{A.31}$$

using (A.22), the fact that $\Sigma_t^{A,\star} + \Sigma_t^{N,\star} = 0$, and (A.28). To see why the inequality

(A.30) holds, first note that

$$\Gamma^{A,\star} > \int_0^{+\infty} \exp(-\rho t)\, u'\left(c_t^A\right) \left\{\mu_t^A - z\right\} dt. \tag{A.32}$$

Then, let $\lambda_t^h \equiv u'\left(c_t^h\right) / u'\left(c_0^h\right)$ and

$$\omega_t^h \equiv \frac{\exp(-\rho t)\, \lambda_t^h}{\int_0^{+\infty} \exp(-\rho s)\, \lambda_s^h ds} \tag{A.33}$$

for each $h = A, N$. Note that the sequence $\{\omega_t^A - \omega_t^N\}$ integrates to zero and decreases over time. The reason is that the income (and thus consumption) of automated workers $\hat{\mathcal{Y}}_t^A$ grows faster over time than that of non-automated workers $\hat{\mathcal{Y}}_t^N$ (Appendix A.2). Therefore,

$$\int_0^{+\infty} \omega_t^A \left\{\mu_t^A - z\right\} dt > \int_0^{+\infty} \omega_t^N \left\{\mu_t^A - z\right\} dt = 0, \tag{A.34}$$

since μ_t^A decreases over time by (3.4). This shows why the inequality (A.30) holds.

Taken together, the previous steps show that (A.24) holds with inequality for virtually any economy, so that there exists a variation that improves the welfare of automated workers $\delta U^A > 0$ and leaves non-automated workers indifferent $\delta U^N = 0$. That is, the equilibrium is generically constrained inefficient.

Part II (taxing automation). We now prove that the Pareto improvement requires taxing automation. The variation $(\delta\alpha, \delta T)$ with $\delta\alpha < 0$ results in $\delta U^N = 0$ and $\delta U^A > 0$ if and only if

$$\begin{aligned} &\int_0^{+\infty} \exp(-\rho t)\, u'\left(c_t^A\right) \left(\Delta_t^\star + \Sigma_t^{A,\star}\right) dt \\ &\quad < \left(1 + \tilde{\Omega}\right) \times \int_0^{+\infty} \exp(-\rho t)\, u'\left(c_t^N\right) \left(\Delta_t^\star + \Sigma_t^{A,\star}\right) dt, \end{aligned} \tag{A.35}$$

where

$$\tilde{\Omega} \equiv \frac{\int_T^{+\infty} \exp(-\rho t)\, u'\left(c_t^A\right) \left(\Sigma_t^A - \Delta_t\right) dt}{\int_T^{+\infty} \exp(-\rho t)\, u'\left(c_t^N\right) \left(\Sigma_t^A - \Delta_t\right) dt} - 1, \tag{A.36}$$

using Lemmas 1 and 3 and the fact that $\Sigma^{A,(\star)} + \Sigma^{N,(\star)} = 0$. First, note that

$$\int_0^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^A\right)\left(\Delta_t^\star + \Sigma_t^{A,\star}\right) dt \\ < \int_0^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^N\right)\left(\Delta_t^\star + \Sigma_t^{A,\star}\right) dt \tag{A.37}$$

since

$$\int_0^{+\infty} \omega_t^A \Delta_t^\star dt < \int_0^{+\infty} \omega_t^N \Delta_t^\star dt = 0, \tag{A.38}$$

as $\Delta_t^\star$ increases over time (Lemma 3) and $\left\{\omega_t^h\right\}$ are given by (A.33). Moreover, $\Sigma_t^{A,\star} < 0$ while $u'\left(c_t^A\right) > u'\left(c_t^N\right)$. Putting this together, the inequality (A.35) holds when $\tilde{\Omega}$ is sufficiently small. This is the case by Assumption 3 as we discuss at the end of this proof. As a result, the Pareto improvement requires taxing automation $\delta\alpha < 0$.

Assumption 3. It remains to show that there exists a $Z^\star > 0$ such that $\tilde{\Omega}$ is sufficiently small. Note that $\tilde{\Omega} \to 0$ in the limit where $\theta \to 0$ and $1/\kappa \to 0$ since $u'\left(c_t^A\right) = u'\left(c_t^N\right)$ once reallocation is over for $t \geq T$. By continuity of the equilibrium in θ and $1/\kappa$, there exists a $Z^\star > 0$ such that $\tilde{\Omega}$ is sufficiently small for the inequality (A.35) to hold when Assumption 3 is satisfied. Finally, the threshold $Z^\star > 0$ in Assumption 3 is the minimum between this one and the ones identified in Appendices A.1 and A.2.

A.5 Proof of Proposition 2

The first order condition of the constrained Ramsey problem with respect to automation is

$$\partial_\alpha U = \sum_h \eta^{\text{effic,h}} \int_0^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^h\right) \times \left(\Delta_t^\star + \hat{\Sigma}_t^{h,\star}\right) dt \tag{A.39}$$

when using efficiency weights $\eta^{\text{effic,h}} = 1/\int_0^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^h\right) dt$. Note that

$$\sum_h \int_0^{+\infty} \exp\left(-\rho t\right) \eta^{\text{effic,h}} u'\left(c_t^h\right) \times \Delta_t^\star dt < 0 \tag{A.40}$$

using (A.38) and the definition of efficiency weights. Furthermore,

$$\sum_h \eta^{\text{effic,h}} \int_0^{+\infty} \exp\left(-\rho t\right) u'\left(c_t^h\right) \times \hat{\Sigma}_t^{h,\star} dt$$
$$= \sum_h \int_0^{+\infty} \left(\omega_t^A - \omega_t^N\right) \times \hat{\Sigma}_t^{A,\star} dt \tag{A.41}$$

using the definition of efficiency weights, the fact that $\Sigma_t^{N,\star} = -\Sigma_t^{A,\star}$, and that $\{\omega_t^h\}$ are given by (A.33). Proceeding as in Appendix A.4, the expression (A.41) is negative as soon as $\hat{\Sigma}_t^{A,\star}$ increases over time; or, equivalently, as soon as $\Sigma_t^{A,\star}$ does. This is the case by Assumptions 3 and 4 as we discuss at the end of this proof.

Taken together, the previous steps imply that $\partial_\alpha U < 0$ so the government finds it optimal to tax automation $\delta\alpha < 0$ locally starting from the laissez-faire α^{LF}. When the government's problem (4.1) is convex, this also implies that the laissez-faire automation is excessive compared to its second best counterpart $\alpha^{\text{SB,effic}} < \alpha^{\text{LF}}$.

Assumptions 3 and 4. It remains to show that $\Sigma_t^{A,\star}$ increases over time for some $Z^\star > 0$ in Assumption 3. Using (A.22) and the fact that $w_t^h = \partial_h G^\star\left(\mu_t^A, \mu_t^N; \alpha\right)$, we have

$$\Sigma_t^{A,\star} = \mu_t^A \left(\partial^2_{A\alpha} G^\star\left(\mu_t^A, \mu_t^N; \alpha\right) - \partial^2_{N\alpha} G^\star\left(\mu_t^A, \mu_t^N; \alpha\right)\right)$$
$$+ \left(\mu_t^N + \mu_t^A - 1\right) \partial^2_{N\alpha} G^\star\left(\mu_t^A, \mu_t^N; \alpha\right).$$

Absent productivity loss ($\theta \to 0$) or unemployment spells ($1/\kappa \to 0$), we have $\mu_t^N + \mu_t^A \to 1$. Therefore,

$$\begin{aligned}\partial_t \Sigma_t^{A,\star} \to \quad & \partial_t \mu_t^A \left(\partial^2_{A\alpha} G^\star(.) - \partial^2_{N\alpha} G^\star(.)\right) \\ & + \mu_t^A \times \partial_t \mu_t^A \left(\partial^3_{AA\alpha} G^\star(.) + \partial^3_{NN\alpha} G^\star(.) - 2\partial^3_{AN\alpha} G^\star(.)\right) \\ & = \underbrace{\partial_t \mu_t^A}_{<0} \underbrace{\partial^2_{\mu\alpha} \left\{\mu \partial_\mu G^\star\left(\mu, 1-\mu; \alpha\right)\right\} |_{\mu=\mu_t^A}}_{<0} > 0,\end{aligned}$$

using Assumption 4 and the fact that $\partial_t \mu_t^A < 0$ from (3.4). By continuity of the equilibrium in θ and $1/\kappa$, there exists a $Z^\star > 0$ such that $\Sigma_t^{A,\star}$ increases over time. Finally, the threshold $Z^\star > 0$ in Assumption 3 is the minimum between this one and the ones identified in Appendices A.1, A.2, and A.4.

A.6 Proof of Proposition 3

The law of motion of automation is $d\alpha_t = (x_t - \delta\alpha_t)\,dt$ for depreciation rate $\delta > 0$, and output (net of investment costs) is

$$Y_t = G\left(\mu_t^A, \mu_t^N; \alpha_t\right) - q_t x_t, \tag{A.42}$$

where x_t is gross investment. The cost of automation q_t decreases over time and converges to $q_t \to \bar{q}$ as $t \to +\infty$. Generations are indexed by s, and are born and die at rate χ. We show below that the equilibrium converges to a first best in the long-run. We refer the interested reader to the working paper version Beraja and Zorzi (2022) for a full description of the equilibrium with overlapping generations and the first best planning problem. We will show that the laissez-faire coincides with a particular first best allocation in the long-run.

Laissez-faire. We guess (and verify) that automation, the labor allocation and the interest rate all converge to a long-run steady state with $r_t^{\text{LF}} \to \rho$ as $t \to +\infty$. We omit the time indices at the final steady state. If the labor allocation converges to a steady state, i.e., $\mu_t^{h,\text{LF}} \to \mu_t^{\text{LF}}$ as $t \to +\infty$ in each $h = A, N$, then investment and automation also converge to steady state levels, i.e., $\alpha_t^{\text{LF}} \to \alpha^{\text{LF}}$ and $x_t^{\text{LF}} \to x^{\text{LF}}$ as $t \to +\infty$ with

$$G_\alpha\left(\mu^{A,\text{LF}}, \mu^{N,\text{LF}}; \alpha^{\text{LF}}\right) = (\rho + \delta)\,\underline{q} \tag{A.43}$$

and $x^{\text{LF}} = \delta\alpha^{\text{LF}}$. Conversely, if automation converges to a steady state level, so does the labor allocation and wages converge to

$$w^{A,\text{LF}} = G_1\left(\mu^{\text{LF}}, 1 - \mu^{\text{LF}}; \alpha^{\text{LF}}\right) = G_2\left(\mu^{\text{LF}}, 1 - \mu^{\text{LF}}; \alpha^{\text{LF}}\right) = w^{N,\text{LF}}, \tag{A.44}$$

as the entry of new generations implies that the marginal products of labor (and so wages) must be equal across occupations in the long-run. Note that equations (A.43)-(A.44) pin down the long-run labor allocation $\left\{\mu^{A,\text{LF}}, \mu^{N,\text{LF}}\right\} = \left\{\mu^{\text{LF}}, 1 - \mu^{\text{LF}}\right\}$, automation α^{LF}, and aggregate consumption

$$C^{\mathrm{LF}} = G\left(\mu^{\mathrm{LF}}, 1-\mu^{\mathrm{LF}}; \alpha^{\mathrm{LF}}\right) - \delta\alpha^{\mathrm{LF}}. \tag{A.45}$$

Finally, all workers are hand-to-mouth ($\underline{a} \to 0$) so $c_s^{\mathrm{LF}} \to C^{\mathrm{LF}}$ as $t \to +\infty$ for all generations s.[37] Therefore, $u'\left(c_{s,t+\tau}^{\mathrm{LF}}\right)/u'\left(c_{s,t}^{\mathrm{LF}}\right) \to 1$ as $t \to +\infty$ for all workers and horizons $\tau \geq 0$. This confirms that the interest rate $r_t^{\mathrm{LF}} \to \rho$ as $t \to +\infty$, and the guess is verified.

First best. Proceeding as above, we can show that any first best allocation also converges to a steady state. We will show that this first best allocation is the same as the one that prevails at the laissez-faire when the planner discounts generations with the subjective discount rate ρ, i.e., it uses weights $\eta_s^h = \exp(-\rho s)$.[38] The automation choice satisfies (A.43) in the long-run when the planner discounts generations at rate ρ. Production efficiency requires that the marginal products of labor must be equalized in a long-run, so equation (A.44) holds too. Therefore, the *aggregate* allocation coincides with the laissez-faire in the long-run. It remains to show that *individual* consumptions are equal at this allocation. Note that the planner equalizes weighted marginal utilities across workers in each period t, so

$$\frac{\eta_s^h \exp\left(-\rho\left(t-s\right)\right) u'\left(c_{s,t}^{\mathrm{FB}}\right)}{\eta_\tau^j \exp\left(-\rho\left(t-\tau\right)\right) u'\left(c_{\tau,t}^{\mathrm{FB}}\right)} = 1 \tag{A.46}$$

for generations $s, \tau \leq t$. Thus, consumption is equalized across workers when the planner discounts generations at rate ρ. Therefore, $c_{s,t}^{\mathrm{FB}} \to C^{\mathrm{FB}} = C^{\mathrm{LF}}$ for all s,t. This completes the proof.

A.7 Example

Using our example from Section 2.1, we show that an increase in the degree automation α decreases the marginal productivity of labor (MPL) *within* the automated occupation, while potentially raising the *aggregate* MPL.

[37] We do not index consumption by the worker's initial industry of occupation. The reason is that mass of surviving members of old generations (born in $s < 0$) vanishes asymptotically, and new generations (born in $s \geq 0$) can choose their initial occupation of employment.

[38] All continuation values are evaluated at birth as in Calvo and Obstfeld (1988).

The log-change in the MPL in the automated occupation is

$$\frac{d}{d\alpha}\log\left(\mathrm{MPL^A}\right)=-\frac{1}{\nu}\frac{1}{y^A}\frac{(1-\phi)\left(y^N\right)^{\frac{\nu-1}{\nu}}}{\phi\left(y^A\right)^{\frac{\nu-1}{\nu}}+(1-\phi)\left(y^N\right)^{\frac{\nu-1}{\nu}}}<0$$

since $\phi\in[0,1]$. Moreover,

$$\frac{d}{d\alpha}\log\left(\mathrm{MPL^N}\right)=\frac{1}{\nu}\frac{1}{y^A}\frac{\phi\left(y^A\right)^{\frac{\nu-1}{\nu}}}{\phi\left(y^A\right)^{\frac{\nu-1}{\nu}}+(1-\phi)\left(y^N\right)^{\frac{\nu-1}{\nu}}}\geq 0.$$

That is, the MPL declines in the automatable occupation but increases in non-automatable occupation. The marginal productivity of labor at the aggregate level, i.e., workers' average wage rate, is

$$\mathrm{MPL}\equiv\frac{\phi\mu^A}{\phi\mu^A+(1-\phi)\,\mu^N}\mathrm{MPL^A}+\frac{(1-\phi)\,\mu^N}{\phi\mu^A+(1-\phi)\,\mu^N}\mathrm{MPL^N}$$

can increase or decrease, depending on $\left(\mu^A,\mu^N,\phi,\nu\right)$.

B Quantitative Model

In this appendix, we describe our quantitative model in more detail. Section B.1 provides a recursive formulation of the workers' problem. Section B.2 states and characterizes the solution to the occupations' problem. Section B.3 discusses the second best.

B.1 Workers' Problem

We discretize time into periods of constant length $\Delta\equiv 1/N>0$, and solve the workers' problem in discrete time.[39] The workers' problem can be formulated

[39] Alternatively, we could have formulated the workers' problem in continuous time and solved the associated partial differential equation using standard finite difference methods. However, (semi-)implicit schemes are non-linear in our setting due to the discrete occupational choice. This requires iterating on (B.1)–(B.5) to compute policy functions which limits the efficiency of these schemes. We found that explicit schemes were unstable unless we use a particularly small time step Δ which again proves relatively inefficient. Formulating and solving the workers' problem in discrete time proves to be relatively fast.

recursively:

$$V_t^h(a,e,\xi,z) = \max_{c,a'} u(c)\Delta + \exp(-(\rho+\chi)\Delta) V_{t+\Delta}^{h,\star}(a',e,\xi,z) \tag{B.1}$$
$$\text{s.t. } a' = (\mathcal{Y}_t(\mathbf{x}) - c)\Delta + \frac{1}{1-\chi\Delta}(1+r_t\Delta)a$$
$$a' \geq 0$$

for employed workers ($e = E$) and unemployed workers ($e = U$). The continuation value $V^\star$ before workers observe the mean-reverting component of their income is given by

$$V_t^{h,\star}(a,e,\xi,z) = \int \hat{V}_t^h(a,e,\xi,z')\, P(dz',z), \tag{B.2}$$

where $\hat{V}_t(\cdot)$ is the continuation value associated to the discrete occupational choice. The continuation value for employed workers ($e = E$) associated to this discrete chocie problem is[40]

$$\hat{V}_t^h(a,e,\xi,z) = (1-\lambda\Delta) V_t^h(a,e,\xi,z) + \lambda\Delta\gamma \log\left(\sum_{h'} \phi^{h'} \exp\left(\frac{V_t^{h'}(a,e'(h',\mathbf{x}),\xi,z)}{\gamma}\right)\right), \tag{B.3}$$

with $e'(\cdot) = E$ if $h' = h$ and $e'(\cdot) = U$ otherwise. The associated mobility hazard across occupations is

$$\mathcal{S}_t(h';\mathbf{x}) = \frac{\phi^{h'} \exp\left(\frac{V_t^{h'}(\mathbf{x}'(h';\mathbf{x}))}{\gamma}\right)}{\sum_{h''} \phi^{h''} \exp\left(\frac{V_t^{h''}(\mathbf{x}'(h'';\mathbf{x}))}{\gamma}\right)}, \tag{B.4}$$

where $\mathbf{x}'(h';\mathbf{x})$ is short for $(a, e'(h',\mathbf{x}), \xi, z)$. In turn, the continuation value for unemployed workers ($e = U$) is

$$\hat{V}^h(a,e,\xi,z) = (1-\kappa\Delta) V^h(a,e,\xi,z) + \kappa\Delta V^h(a,1,\xi'(h',\mathbf{x}),z), \tag{B.5}$$

where $\mathcal{S}(\cdot)\,\xi'(\cdot) = (1-\theta)\xi$ when the reallocation spell is complete. New gen-

[40] See Artuç et al. (2010) for the derivation.

erations who enter the labor market draw a random productivity z from its stationary distribution and then choose their occupation with a hazard similar to the employed workers'. The only difference is that they experience neither an unemployment spell nor a productivity loss. Worker's labor income is

$$\mathcal{Y}_t(\mathbf{x}) = \begin{cases} \xi \exp(z)\, w_t^h & \text{if } e = E \\ b\mathcal{Y}_t^{h'}(a, E, \xi, z) & \text{otherwise} \end{cases}, \tag{B.6}$$

with $h' \neq h$ denoting the previous occupation of employment. The permanent component of workers' income (ξ) is reduced by a factor $(1-\theta)$ whenever a worker who exits unemployment enters her new occupation. Workers experience this productivity loss at most once during their lifetime. Finally, the mean-reverting component income (z) evolves as

$$z' = (1 + (\rho_z - 1)\Delta)\, z + \sigma_z \sqrt{\Delta} W' \quad \text{with} \quad W' \sim \text{i.i.d.}\mathcal{N}(0,1). \tag{B.7}$$

B.2 Firm's Problem

We solve the firm's problem in continuous time. It's problem is

$$\max_{\{x_t, \alpha_t, \mu_t^A \mu_t^N\}} \int_0^{+\infty} \exp\left(-\int_0^t r_s ds\right) \left\{ G\left(\mu_t^A, \mu_t^N; \alpha_t\right) - \phi q_t x_t - \sum_h \phi^h w_t^h \mu_t^h \right\} dt$$

$$\text{s.t. } d\alpha_t = (x_t - \delta\alpha_t)\, dt\,,\ \alpha_0 = 0\,,\ x_t \geq 0 \tag{B.8}$$

where α_t is the stock of automation, x_t is gross investment, μ_t^A and μ_t^N are (effective) labor demands, and q_t is the resource cost per unit of investment.[41] The optimal degree of automation satisfies

$$(r_t + \delta)\, \phi q_t = \partial_\alpha G\left(\mu_t^A, \mu_t^N; \alpha_t\right) + \phi \partial_t q_t, \tag{B.9}$$

[41] For concision, we omit any distortionary tax on investment τ_t^x since it is isomorphic to the cost of investment q_t.

together with the law of motion

$$d\alpha_t = (x_t - \delta\alpha_t)\, dt, \tag{B.10}$$

and the initial condition $\alpha_0 = 0$. Finally, the firm's labor demands satisfy

$$w_t^h = (1-\eta)\frac{1}{\alpha_t^h + \mu_t^h}\frac{\left\{A^h\left(\alpha_t^h + \mu_t^h\right)^{(1-\eta)}\right\}^{\frac{\nu-1}{\nu}}}{\sum_g \phi^g \left\{A^g\left(\alpha_t^g + \mu_t^g\right)^{(1-\eta)}\right\}^{\frac{\nu-1}{\nu}}} G\left(\left\{\alpha_t^h, \mu_t^h\right\}\right), \tag{B.11}$$

where

$$\mu_t^h = \frac{1}{\phi^h}\int \mathbf{1}_{\{e=1,h'=h\}}\xi d\pi_t \tag{B.12}$$

is the (effective) labor supplied in *each* occupation.

B.3 Second Best

In this appendix, we state the second best problem that we solve numerically and discuss our choice of Pareto weights.

Objective. The government's objective is

$$\begin{aligned}\mathcal{W} \equiv & \chi\int_{-\infty}^{0}\int \eta_s(\mathbf{x})\exp\left(-(\rho+\chi)(0-s)\right)V_0^{\text{old}}(\mathbf{x})\,\pi_{s,0}^{\text{old}}(d\mathbf{x})\,ds \\ & + \chi\int_0^{+\infty}\eta_s V_s^{\text{new}} ds,\end{aligned} \tag{B.13}$$

for some Pareto weights $\boldsymbol{\eta}$, where $\pi_{s,0}^{\text{old}}$ is the initial distribution of idiosyncratic states for existing generations born in $s < 0$ (conditional on survival). Following Calvo and Obstfeld (1988), all continuation values are evaluated at birth. The value $\exp\left(-(\rho+\chi)(0-s)\right)V_0^{\text{old}}$ is the continuation utility of existing generations over periods $t \geq 0$. In turn, the value

$$V_t^{\text{new}} \equiv \int \gamma \log\left(\sum_h \phi^h \exp\left(\frac{V_t^h(0,1,0,z)}{\gamma}\right)\right) P^\star(dz) \tag{B.14}$$

is the continuation utility for new generations born in period $t = s \geq 0$, which reflects their occupational choice.[42] Here, $P^\star$ denotes the ergodic distribution of the income process $z'|z \sim P(z)$, i.e., the distribution of productivities at birth.

Pareto weights. We use *efficiency* weights that capture the efficiency motive for policy intervention and ensure that the government has no incentive to redistribute resources. These weights are the ones we described in our baseline model (Section 4.5). The government discounts generations using the subjective discount rate ρ as in Itskhoki and Moll (2019) and Guerreiro et al. (2022). Therefore, the weights are

$$\eta_0(\mathbf{x}) = \exp(-\rho s) \times \frac{1}{\int_0^{+\infty} \exp(-(\rho+\chi)t)\, u'\left(C_t^{\text{old},h}\right) dt}$$

for old generations employed in automated $(h = A)$ and non-automated $(h = N)$ occupations, where $C_t^{\text{old},h}$ denotes average consumption over time for each of these groups at the laissez-faire.[43] The weights for new generations (indexed by $s > 0$) are similar, except that they are not indexed by h since new generations are able to choose their initial occupation of employment. They depend on the consumption streams $C_{s,t}^{\text{new}}$ for all $t \geq s$.

Summarizing, the government's objective becomes

$$\begin{aligned} \mathcal{W} \equiv & \int \frac{V_0(\mathbf{x})}{\int_0^{+\infty} \exp(-(\rho+\chi)t)\, u'\left(C_t^{\text{old},h}\right) dt} \pi_0(d\mathbf{x})\, ds \\ & + \chi \int_0^{+\infty} \exp(-\rho s) \frac{V_s^{\text{new}}}{\int_s^{+\infty} \exp(-(\rho+\chi)(t-s))\, u'\left(C_{s,t}^{\text{new}}\right) dt} ds, \end{aligned} \tag{B.15}$$

where

$$\pi_0(d\mathbf{x}) \equiv \int_{-\infty}^{0} \chi \exp(\chi s)\, \pi_{s,0}^{\text{old}}(d\mathbf{x})\, ds \tag{B.16}$$

is the initial distribution of idiosyncratic states.

Policy tools and implementability. The government maximizes the objective (B.15) by

[42] Members of a new generation are born with no assets $a = 0$, are employed $e = 1$, and have not incurred the productivity cost associated to switching occupations $\xi = 1$.

[43] An alternative approach would be to compute the expected marginal utility over time separately for each initial state $(\mathbf{x})$. This is computationally infeasible since our state space is too large.

choosing an appropriate sequence of distortionary taxes on investment $\{\tau_t^x\}$ and rebating the proceedings to the firm. The implementability constraints consist of workers' reallocation and consumption choices.

C Numerical Implementation

We discuss how we solve numerically for the equilibrium and the optimal policy.

Workers' problem. We solve the worker's problem (B.1) using the standard endogenous grid method (Carroll, 2006). In theory, this problem could be non-convex since it involves a discrete choice across occupations. However, we find that this is not the case in our calibration. The variance of the taste shocks γ is sufficiently large that the value function remains concave. We use Young (2010)'s non-stochastic simulation method to iterate on the distribution. Finally, we discretize the income process on a 7-point grid using the method of Rouwenhorst (1995).

Firm's problem. Given a sequence for the cost of investment $\{q_t\}$ and the interest rate $\{r_t\}$, the optimal sequence of automation and investment can be solved using (B.9)–(B.10) with initial condition $\alpha_0 = 0$. The initial cost of investment ensures that automation is continuous in $t = 0$ at the laissez-faire

$$q_0 = \frac{1}{\phi} \frac{\partial_\alpha G\left(\mu_0^A, \mu_0^N; \alpha_0\right) + \varphi\phi q_\infty}{r_0 + \delta + \varphi}, \tag{C.1}$$

using $q_t = q_\infty + \exp\left(-\varphi t\right)\left(q_0 - q_\infty\right)$ where $q_\infty \equiv \lim_{t\to+\infty} q_t$ is the long-run cost of investment.

Policy. We adopt a primal approach as in our benchmark model. We assume that the government can directly choose the wage gap across occupations $\hat{w}_t \equiv \log\left(w_t^A\right) - \log\left(w_t^N\right)$. For numerical reasons, we restrict our attention to parametric perturbations. Specifically, we consider policies of the form

$$\hat{w}_t = S\left(t; \Theta\right) \hat{w}_t^{\text{LF}} \tag{C.2}$$

where $\hat{w}_t^{\mathrm{LF}}$ is the wage gap that prevails at the laissez-faire, and

$$S\left(t;\Theta\right) \equiv \min\left\{\max\left\{3\left(\frac{t}{\Theta}\right)^2 - 2\left(\frac{t}{\Theta}\right)^3, 0\right\}, 1\right\} \tag{C.3}$$

is a smoothstep function with argument t and scale parameter Θ.[44] We search for the optimal Θ over a grid, computing welfare (B.15) for each point. The second best intervention is the one that delivers the highest welfare.

We proceed as follows to recover the taxes on investment $\{\tau_t^x\}$ that implement the second best. We define $q_t^{\mathrm{SB}} \equiv (1+\tau_t^x)\, q_t$, where q_t is the laissez-faire cost. Using (B.9),

$$(r_t+\delta)\,\phi q_t^{\mathrm{SB}} = \partial_\alpha G\left(\mu_t^{A,\mathrm{SB}}, \mu_t^{N,\mathrm{SB}}; \alpha_t^{\mathrm{SB}}\right) + \phi \partial_t q_t^{\mathrm{SB}}, \tag{C.4}$$

where the allocations are evaluated at the second best. This expression defines a differential equation for $\left\{q_t^{\mathrm{SB}}\right\}$ with terminal condition $\lim_{t\to+\infty} q_t^{\mathrm{SB}}/q_t = 1$ since the second best allocation converges to its laissez-faire level. We solve this differential equation using a standard shooting algorithm, and we recover the taxes $\tau_t^x = q_t^{\mathrm{SB}}/q_t - 1$.

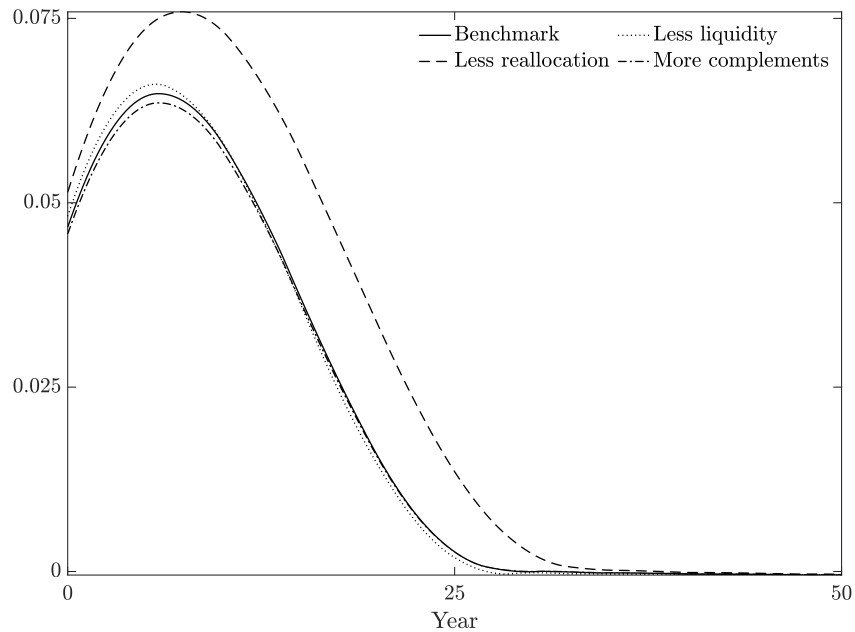
Welfare gains. We compute the welfare gains as the ratio between the certainty equivalent consumptions that produce the same welfare as in the second best $\left(\mathcal{W}^{\mathrm{SB}}\right)$ and the laissez-faire $\left(\mathcal{W}^{\mathrm{LF}}\right)$, respectively. These welfare gains are thus given by $\left(\mathcal{W}^{\mathrm{SB}}/\mathcal{W}^{\mathrm{LF}}\right)^{\frac{1}{1-\sigma}}$, where $\sigma > 0$ is the inverse elasticity of intertemporal substitution.

D Taxes on Automation

Figure D.1 plots the sequence of taxes on investment $\{\tau_t^x\}$ that implement the second best allocation, for each of the four calibrations in Table 6.2. We discuss how we solve for these taxes in Appendix C. By construction, taxes converge to zero in the long-run (footnote 44).

[44] Note that these policies constrain the government to intervene only along the transition. In theory, the government might also want to interevene in the long-run due to uninsured idiosyncratic risk (Dávila et al., 2012). By construction, allowing for such, more flexible, policy would produce even higher welfare gains compared to Table 6.2. We chose to abstract from long-run taxation to focus on the new motive for intervention that we highlight (Section 4.6.2).

Figure D.1: Investment taxes at the second best



Notes: The four curves correspond to the calibrations in Table 6.2.

E Output Share and Employment

We argued that our model matches well the share of output produced by automation forecasted by McKinsey (2017), as well as the firm-level effects of automation on employment estimated by Bonfiglioli et al. (2022). We now explain how we compute the model analogs of these (untargeted) moments.

Output share. Exhibit E3 in McKinsey (2017) shows that 71% of the output previously produced by labor could be automated. This figure is obtained by taking the weighted average of the time spent on automatable activities in the three most susceptible activities $0.71 = (17 \times 64 + 16 \times 69 + 18 \times 81) / (17 + 16 + 18)$. In our model, the share of output in occupation $h = A$ that is produced by automation is $\alpha / \left(\alpha + \mu^A\right)$, which is 67% when evaluated at the final steady state.

Employment. The percent change in employment of a firm that adopted automation, relative to a firm that did not, can be computed by the ratio of the coefficients in column (2) to column (5) in the first line of Table 2 of Bonfiglioli et al. (2022). This gives -0.094/0.174 $= -54\%$. In our model, labor demand from a "firm" producing the output of an occupation as an intermediate good satisfies

$$A\left(1-\eta\right)\left(\alpha+\mu\right)^{-\eta}=\frac{w}{p}, \tag{E.1}$$

where w is wage and p is the price of the intermediate good produced in automated occupations. Next, we consider the following partial equilibrium exercise. Let us compare two intermediate goods firms facing the same wage and price. One has automation $\alpha_1 > 0$ and the other has no automation $\alpha_0 = 0$. Then, it must be that

$$\alpha_1+\mu_1=\mu_0. \tag{E.2}$$

So, the log-change in employment is

$$\log\left(\mu_1\right)-\log\left(\mu_0\right)=\log\left(1-\frac{\alpha_1}{\mu_0}\right)=-0.66 \tag{E.3}$$

using our calibration, where μ_0 is the initial steady state employment in automated occupations and α_1 equals the stock of automation 20 years out in the transition, which is the sample period in Table 2 of Bonfiglioli et al. (2022). Next, we consider employment changes in general equilibrium across steady states. We find that employment in automated occupations changes by -0.37 log points. Overall, our partial and general equilibrium exercises deliver predictions that are comparable but slightly smaller than the -0.54 log change that Bonfiglioli et al. (2022) estimate.