Inefficient Automation*

Martin Beraja
MIT and NBER

Nathan Zorzi
Dartmouth

October 16, 2022

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 first show that these frictions result in inefficient automation. Firms fail to internalize that displaced workers have a limited ability to smooth consumption while they reallocate. We then analyze a second best problem where the government can tax automation but lacks redistributive tools to fully overcome borrowing frictions. The equilibrium is (constrained) inefficient. The government finds it optimal to slow down automation on efficiency grounds, even when it has no preference for redistribution. 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 aggregate efficiency and achieves welfare gains of 4%. Slowing down automation achieves important gains even when the government implements generous social insurance policies.

*Martin Beraja: maberaja@mit.edu. Nathan Zorzi: nathan.g.zorzi@dartmouth.edu. We thank 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., 2017 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, an extensive 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. There is a continuum of occupations that 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 this model for our quantitative analysis.

We have two main theoretical results. Our first result shows that the interaction between slow reallocation and borrowing constraints results in *inefficient automation*. 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. There is a conflict between how the firm and displaced workers value the gains from automation over time. Effectively, firms fail to internalize that displaced workers have a limited ability to smooth consumption while they reallocate. Private and social incentives to automate do not coincide.

Our second result characterizes optimal policy. In principle, the government could restore efficiency if it was able to fully relax borrowing constraints using redistributive transfers. This is unlikely in practice.[2] This motivates us to study second best interventions, where the government can tax automation and (poten-

[1] This is consistent with the empirical evidence. Displaced workers borrow to smooth consumption when they are able to (Sullivan, 2008). Many workers are constrained and are either unable to borrow or forced to delever their existing debt (Braxton et al., 2020).

[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.

tially) implement active labor market interventions but is unable to fully alleviate the borrowing constraints of displaced workers by redistributing income.[3]

We find that the equilibrium is generically (constrained) inefficient, as defined by Geanakoplos and Polemarchakis (1985). Automation and reallocation choices impose *pecuniary externalities* on workers. Firms do not internalize that automation displaces workers and lowers their earnings, and workers do not internalize how their reallocation affects the wage of their peers. The optimal policy addreses these pecuniary externalities. This policy reduces the present discounted value of output (net of resource costs) compared to the laissez-faire, but increases welfare through two channels (Bhandari et al., 2021): it improves *aggregate efficiency* by changing the flows of aggregate consumption over time, and it improves *redistribution* by changing how consumption is allocated across workers.

We show that the government should tax automation on efficiency grounds — even when it has no preference for redistribution. In particular, the government should *slow down* automation while labor reallocation takes place but should not intervene in the long-run. The logic is as follows. The output gains from automation build over time, since they materialize slowly as more workers reallocate. The government values future gains *less* than firms do. It recognizes that automated workers have steeper consumption profiles and are *effectively* more impatient than the average worker who prices the firms' equity. Slowing down automation lowers output but improves *aggregate efficiency* by flattening consumption profiles, raising consumption early on in the transition when displaced workers value it more.

We then suppose that 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 is reinforced, as borrowing constrained workers rely excessively on mobility as a source of self-insurance.

We conclude the paper with a quantitative exploration. Our goal is to evaluate the efficiency and welfare gains from slowing down automation, while allowing

[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 recently reduced its tax credits on investment in automated technologies, the canton of Geneva in Switzerland taxes automated cashiers, and Nevada imposed an excise tax on autonomous vehicles. See Kovacev (2020) for a detailed review of these cases.

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 idyosincratic 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 Lindé, 2001), which produces a realistic distribution of savings. We also allow for progressive income taxation (Heathcote et al., 2017) and unemployment benefits (Krueger et al., 2016) to account for existing insurance that helps workers.

The constrained efficient intervention slows down the speed of automation substantially compared to the laissez faire. A government that only values *efficiency* should tax automation so as to reduce its half-life by a factor of 2 at least. This policy achieves sizable welfare gains of about 4% in consumption equivalent terms. The gains are even larger (around 6%) for a utilitarian government that values redistribution since the policy improves not only efficiency but also equity.

We then consider two alternative calibrations and two alternative polices. First, our benchmark calibration implies that reallocation out of automated occupations is more rapid than in the data. This is conservative in that it allows automated workers to better self-insure. We thus increase the variance of mobility shocks so that automated workers reallocate less. The second best policy now slows down automation more and produces larger welfare gains. Second, unemployment spells could be longer for workers displaced by automation than for the average US worker (1 quarter in our benchmark calibration). We thus increase their average duration, which steepens the consumption profiles of automated workers. The government finds it optimal to slow down automation even more. Finally, we allow the government to insure automated workers by giving them a lump-sum transfer of $10k — the maximum amount allowed by the Reemployment Trade Adjustment Assistance program (RTAA) for instance. The transfers achieve smaller welfare gains than those from slowing down automation, especially when the government does not value redistribution. Put it differently, transfers of this magnitude are effective in improving equity but do not alleviate borrowing constraints much in the medium-run and address inefficient automation. Combining transfers and automation taxes achieves large welfare gains.

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., 2017; Costinot and Werning, 2022; Thuemmel, 2018; Korinek and Stiglitz, 2020). In this literature, taxing automation results in production inefficiency (Diamond and Mirrlees, 1971). Instead, the optimal policy preserves (or restores) production efficiency in our model. Finally, Costinot and Werning (2022) point to sufficient statistics for the optimal taxation of automation in static (efficient) economies. Our analysis uncovers empirical moments that determine how a government should slow down automation to improve efficiency.

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.

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 final goods 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 final goods aggregating the output of occupations. Occupations use labor as an input. Some occupations can be automated (e.g., routine-intensive occupations) whereas others cannot. At time $t = 0$, the firm chooses the degree of automation α in the automatable occupations.[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

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

with (weak) complementarity across occupations. 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] We formalize these assumptions below.

[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.

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.

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

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

where $\varphi > 0$ is the relative productivity of automation. Finally, given an elasticity of substitution $\nu < 1$ across occupations, the aggregate production function is

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

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

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

[6] An increase in α decreases the marginal product of labor *within* the automated occupation but can increase the *aggregate* marginal product of labor (see Appendix A.10). This is the case in our 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 optimal profits given equilibrium wages $\{w_t^h\}$ and the price of the final good which we normalize to be 1.

We impose a regularity condition to rule out corner solutions, so that there is positive but finite automation in equilibirum. This is needed for a meaningful discussion of automation.[7]

Assumption 2 (Interior solution)**.** *The direct effect of automation $G^\star(\mu, \mu'; \alpha)$ is concave in α and satisfies $\partial_\alpha G^\star(\mu, \mu'; \alpha)|_{\alpha=0} > 0$ and $\lim_{\alpha\to+\infty} \partial_\alpha G^\star(\mu, \mu'; \alpha) = -\infty$ for any $0 \leq \mu \leq \frac{1}{2}$ and $\mu' \geq \frac{1}{2}$.*

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

$$U_0 = \mathbb{E}_0\left[\int \exp(-\rho t)\, u(c_t)\, dt\right] \tag{2.6}$$

for some discount rate $\rho > 0$ and some isoelastic utility $u(c) \equiv \frac{c^{1-\sigma}-1}{1-\sigma}$ 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. They 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 ι. The latter exit this state at rate $\kappa > 0$, at which point they enter their new occupation too The non-employment state can be interpreted either as involuntary unemployment due to search frictions or as temporary

[7] Assumption 2 is satisfied if the cost of automation $\mathcal{C}(\alpha)$ is sufficiently convex.

exit from the labor force during which workers retrain for their new occupation.[8] Finally, we assume that workers incur 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.[9] 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 and states. We suppose that financial markets are incomplete: workers cannot trade contingent securities against the risk that their initial occupation is automated.[10] Workers trade riskless bonds available in zero net supply. In addition, each worker is endowed with one unit of the firm's equity.[11] For simplicity, we assume that only a vanishing mass of workers can trade equity.[12]

Budget constraint. Worker's flow budget constraint is

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

[8] 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) as well as an additional cost upon reallocation (a permanent productivity loss due to skill specificity).

[9] 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.

[10] 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.

[11] All our results carry through if we assumed that automated workers did not hold any firm equity. If anything, this assumption is conservative with respect to our mechanism of interest because the income from their equity holdings provides displaced workers with an additional liquidity buffer.

[12] In practice, the wealthiest 10% holds close to 90% of firm equity in the US (Survey of Consumer Finances, 2022). The typical displaced worker does not trade equity. Alternatively, we could have introduced Ricardian investors that hold and trade the equity. All of our results carry through in this case since there would be no arbitrage between equity and bonds too, i.e., equilibrium condition (2.14).

where $\hat{\mathcal{Y}}_t^h$ is labor income, Π_t is profits, and $r_t \geq 0$ is the return on savings. 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. To save on notation, expression (2.8) already uses the fact that, in equilibrium, workers never reallocate from the non-automated to the automated occupation. The expression also assumes that unemployed workers earn no income.[13]

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 sectors 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 become employed in the non-automated occupation $(\tilde{\mu}_t)$ are

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

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

with $u_0 = \tilde{\mu}_0 = 0$.

Next, we impose a regularity condition on reallocation frictions so that output

[13] In our quantitative model, we introduce gross flows across occupations and unemployment benefits.

does not decline over time in equilibrium.

Assumption 3 (Reallocation frictions)**.** *The productivity loss* θ *and the probability of unemployment* ι *are sufficiently small that* $1-(1-\theta)(1-\iota) < Z^{\star}$ *for some* $Z^{\star} > 0$ *defined in Appendix A.3.*

2.3 Equilibrium

Market clearing in the labor market requires

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

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 or equity, as some workers trade both. Therefore, 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

Firm optimization implies that wages equal the marginal productitivites of labor in each occupation

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

Automation is labor-displacing, decreasing the relative wage of workers in automated occupations. This induces them to move towards non-automated occupations. 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

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

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

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

3.2 Consumption and savings

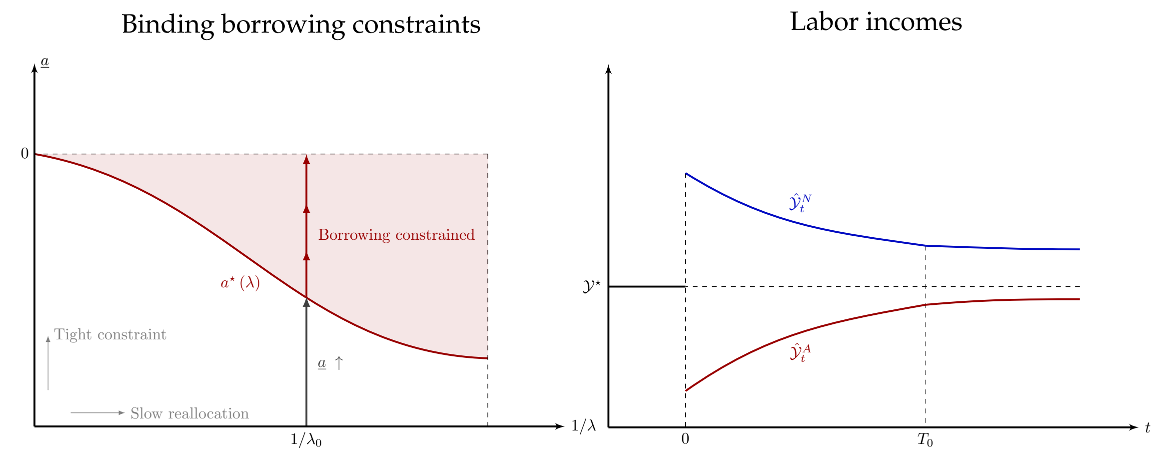
We now show that the labor displacement induced by automation creates a motive for borrowing and that workers become borrowing constrained when reallocation and borrowing frictions are sufficiently severe.

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. 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. □

The left panel of Figure 3.1 illustrates this result in the space of reallocation frictions $(1/\lambda)$ and borrowing frictions $(\underline{a})$ in the particular case where unemployment spells are short $(1/\kappa \to 0)$. This space is partitioned in two main regions. Borrowing constraints do not bind as long as the frictions fall in the white region where $\underline{a} \leq a^\star(\cdot)$. This occurs when either reallocation is sufficiently fast or borrowing constraints are sufficiently loose. In constrast, automated workers become borrowing constrained when the frictions fall in the colored region $\underline{a} > a^{\star}(\cdot)$.[14]

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



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

$$\hat{y}_t^h = \underbrace{\partial_h G^{\star}(\cdot)}_{\text{Initial wage}} + \mathbf{1}_{\{h=A\}} \times 2 \times \left[\underbrace{\left(\frac{1}{2} - \mu_t^A\right) \times \left((1-\theta)\,\partial_N G^{\star}(\cdot) - \partial_A G^{\star}(\cdot)\right)}_{\text{Reallocation gains}} \right. \tag{3.6}$$
$$\left. - \underbrace{\left(1 - \mu_t^A - \mu_t^N\right) \times (1-\theta)\,\partial_N G^{\star}(\cdot)}_{\text{Unemployment loss}} \right].$$

When reallocation is *slow*, automation decreases the income of workers displaced by automation, both directly by lowering the wage $w_t^h = \partial_h G^{\star}(\cdot)$ 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. The following remark states this insight.

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

[14] 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 existing workers never reallocate, which Assumption 3 rules out.

When reallocation and borrowing frictions are sufficiently mild, workers are never borrowing constrained, i.e., the white region in the left panel of Figure 3.1. As the frictions become more severe, borrowing constraints eventually bind $\underline{a} > a^{\star}(\cdot)$, i.e., the colored region in the figure.

To complete the characterization, we note that automation affects both the levels and growth rates of consumption. First, automation has distributional consequences: the marginal utility of consumption is larger for automated workers, i.e., $u'\left(c_t^A\right) > u'\left(c_t^N\right)$. Second, non-automated workers save and automated workers borrow in equilibrium. When borrowing constraints do not bind, the (intertemporal) *marginal rate of substitution* (MRS) of all workers coincides with the economy's interest rate, i.e., $u'\left(c_t^A\right)/u'\left(c_0^A\right) = u'\left(c_t^N\right)/u'\left(c_0^N\right) = \exp\left(-\int_0^t r_s ds\right)$. However, when automated workers are borrowing constrained, there is a wedge between the MRSs of automated and non-automated workers. Automated workers have steeper consumption profiles and are *effectively* more impatient than non-automated workers, i.e., $u'\left(c_t^A\right)/u'\left(c_0^A\right) < u'\left(c_t^N\right)/u'\left(c_0^N\right) = \exp\left(-\int_0^t r_s ds\right)$.

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

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

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 MRS of non-automated workers $u'\left(c_t^N\right)/u'\left(c_0^N\right)$ 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 $\Delta_t^\star$. Assumption 1 ensures that these gains build up over time in equilibrium. The reason is that labor is freed up from the automated occupation and reallocates to the non-automated occupation, and the two occupations are complements. The flows Δ_t^* are *back-loaded* in this case. Automation crowds out consumption early on but eventually increases output (and consumption) as labor reallocates.

3.4 When is Automation Inefficient?

Firms value the gains from automation using the equilibrium interest rate. This stochastic discount factor coincides with the MRS of non-automated workers who are not displaced (Lemma 3). Would the firm choose differently if it valued the gains from automation over time like displaced workers would? The answer is no in any efficient allocation, since the MRS must be the same for all workers. That is, the firm and displaced workers agree on how to value the gains from automation $\Delta_t^\star$ over time

$$\int_0^{+\infty} \exp\left(-\rho t\right)\left(\frac{u'\left(c_t^N\right)}{u'\left(c_0^N\right)} - \frac{u'\left(c_t^A\right)}{u'\left(c_0^A\right)}\right)\Delta_t^\star dt = 0. \tag{3.10}$$

Automation is thus efficient whenever borrowing constraints do not bind (Lemma 2). In particular, there are two limit cases where this occurs. 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. Second, suppose that there are no borrowing frictions $(\underline{a} \to -\infty)$ as in Guerreiro et al. (2017).[15] In this

[15] In Guerreiro et al. (2017), reallocation takes place (entirely) through new generations replacing older ones. In our model, this corresponds to reinterpreting the reallocation rate λ as the birth/death rate of generations, and letting new generations pick any occupation without experiencing unemployment $(1/\kappa \to 0)$.

case, automation still creates a motive for borrowing but there is no wedge between the MRS of automated and non-automated workers.

Suppose instead that borrowing constraints do bind, as in Lemma 2. Then, automated workers who are displaced are more impatient than the firm so that $u'\left(c_t^A\right)/u'\left(c_0^A\right) < u'\left(c_t^N\right)/u'\left(c_0^N\right)$.[16] As a result, the equilibrium degree of automation is inefficient. The following remark summarizes this insight.

Remark 2. *There is a conflict between how the firm and displaced workers value the gains from automation over time. Effectively, firms fail to internalize that displaced workers have a limited ability to smooth consumption while they reallocate. In turn, the laissez-faire degree of automation is inefficient.*

The mechanism in practice. Our mechanism relies on displaced workers becoming borrowing constrained while they slowly reallocate. Empirically, workers who loose their job indeed attempt to borrow to smooth consumption (Sullivan, 2008), but are often unable to do so or are even forced to delever their existing debt (Braxton et al., 2020). While we abstract from ex-ante heterogeneity across workers, our mechanism is more likely to be relevant when automation affects workers with small liquidity buffers. For example, industrial robots, automated cashiers, or autonomous vehicles would tend to displace low-to-middle income routine workers who are more likely to be hand-to-mouth. In contrast, artificial intelligence software for natural language processing tends to affect higher income skilled workers for which borrowing frictions are less severe.

Restoring efficiency. A government that has access to a sufficiently rich set of redistributive tools to fully undo borrowing frictions could, in theory, restore efficiency. Two interventions would close the wedge in the MRS across workers. First, the government could use *targeted* lump sum tranfers $\left\{T_t^h\right\}$ (indexed by worker and time) to help displaced workers. The literatures on optimal income taxation (Piketty and Saez, 2013) and the regulation of automation rule out such a rich set of transfers, in part because the informational requirements to implement them are

[16] The wedge in MRS would occur even if the sequence of interest rates was fixed (as in a small open economy) or in a model with an outside Ricardian household that invests in firm equity. Beyond this wedge, automation is distorted for two additional reasons in general equilibrium: (i) the MRS of unconstrained non-automated workers (or the interest rate that firms face) changes; and (ii) so do wages (and hence) profits.

too large.[17] Second, the government could use *symmetric* transfers $\{T_t\}$ to undo workers' borrowing constraints. Effectively, the government would borrow on behalf of workers in the short-term and repay its debt later on by taxing them. In practice, the future tax burden would tighten borrowing constraints (Aiyagari and McGrattan, 1998) and carry potentially large distorsions (Guner et al., 2021), limiting or entirely reversing the benefits of the transfers. The fiscal cost is also likely to be prohibitive.[18] Therefore, the next section recognizes that the government may not have access to such rich redistributive tools.

4 Optimal Policy Interventions

We now discuss second best policy interventions. In Section 4.1, we state the constrained Ramsey problem of a government that has a restricted set of instruments. In Section 4.2, we show that the equilibrium is generically *constrained inefficient*. In Section 4.3, we show our main result: the government should tax automation on efficiency grounds. Section 4.4 considers a number of extensions, such as restricting the set of instruments further, introducing equity concerns, or allowing for gradual investment in automation.

4.1 The Constrained Ramsey problem

We now assume that the government cannot *fully* alleviate borrowing frictions and restore (first best) efficiency. For tractability and to obtain more compact expressions, we also assume in the following that workers cannot borrow $\underline{a} \to 0$ and abstract from any direct redistributive tools altogether. We re-introduce these tools later in our quantitative model.

Instead, 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 in-

[17] That said, some existing policies partially insure displaced workers, e.g., Reemployment Trade Adjustment Assistance program (RTAA) in the US. However, this type of programs have shown low take-up rates (Schochet et al., 2012) and often have unintended consequences (Crépon and van den Berg, 2016). We allow for realistic direct transfers in our quantitative model (Section 5).

[18] The payments 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.

terventions (Card et al., 2018) that tax or subsidize labor reallocation $\{\varsigma_t\}$.[19] 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) the Swiss canton of Geneva's tax on retail stores installing automated cashiers, and (iii) Nevada's excise tax on transportation companies using autonomous vehicles that would displace human drivers.

The government effectively controls two choices with its instruments: the degree of automation α; and the reallocation of displaced workers, as governed by the stopping time T.[20] All other choices must be consistent with workers' and firms' optimality. The government's constrained Ramsey problem reduces to the following primal problem.

Lemma 4 (Primal problem)**.** *Given Pareto weights $\{\eta^A, \eta^N\}$, the government's primal problem is*

$$\max_{\left\{\alpha, T, \mu_t^A, \mu_t^N, c_t^A, c_t^N\right\}} \sum_h \eta^h \int_0^{+\infty} \exp\left(-\rho t\right) u\left(c_t^h\right) dt \tag{4.1}$$

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).

It is worth noting that the only difference between the constrained problem above and the unconstrained (first best) Ramsey problem lies in the set of implementable consumption allocations. In the second best problem, workers must consume their income, since we have assumed that borrowing is not possible ($\underline{a} \to 0$).

[19] 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).

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

In the first best problem, the planner can choose any consumption allocation that satisfies the resource constraint (2.13).

4.2 Constrained Inefficiency

We now show that the government should intervene regardless of the Pareto weights $\{\eta^A, \eta^N\}$ that it uses. Formally, Appendix A.4 establishes that the laissez-faire is *generically* constrained inefficient in the sense of Geanakoplos and Polemarchakis (1985). That is, there are no Pareto weights such that the second best automation α^{SB} and reallocation T^{SB} choices coincide with the laissez-faire.

To see this, compare the private and social incentives to automate and reallocate

$$\int_0^{+\infty} \exp(-\rho t) \frac{u'\left(c_t^N\right)}{u'\left(c_0^N\right)} \Delta_t^\star dt = -\Phi^\star\left(\alpha^{\text{SB}}, T^{\text{SB}}; \eta^A, \eta^N\right)$$

$$\underbrace{\int_{T^{\text{SB}}}^{+\infty} \exp(-\rho t) \frac{u'\left(c_t^A\right)}{u'\left(c_0^A\right)} \Delta_t dt}_{\text{laissez-faire}} = \underbrace{-\Phi\left(\alpha^{\text{SB}}, T^{\text{SB}}; \eta^A, \eta^N\right)}_{\text{pecuniary externalities}}$$

where the terms $\Phi^\star(\cdot)$ and $\Phi(\cdot)$ capture the *pecuniary externalities* that automation and reallocation impose on workers. These externalities can be decomposed into aggregate components $\Phi^{\text{aggreg},(\star)}$, and distributional components $\Phi^{\text{distrib},(\star)}$, which we define in Appendix A.4. The government takes into account that an increase in automation (α) reduces wages in automated occupations, but increases profits that benefit all workers (or *some* workers when profits are not claimed symmetrically).[21] Similarly, the government internalizes that an increase in reallocation (T) reduces wages in non-automated occupations, but lifts wages in automated occupations. Firms and workers do not internalize these effects. We show that these pecuniary externalities do not net out at the laissez-faire in presence of reallocation and borrowing frictions.

This finding echoes the constrained-inefficiency results in the incomplete markets literature (Lorenzoni, 2008; Dávila and Korinek, 2018). The nature of the inefficiency is different, however. Constrained inefficiency occurs in our economy

[21] Again, all our results carry through in the case where displaced workers do not claim profits. Assuming that profits are claimed symmetrically is conservative, if anything, since the increase in profits partly compensates for the decline in labor income experienced by displaced workers.

despite the absence of uncertainty and incomplete markets, or endogenous borrowing constraints. Instead, it occurs when firms and workers make *technological choices*, and borrowing constraints distort the (shadow) prices that these agents face.[22]

4.3 Taxing Automation on Efficiency Grounds

We now present the main result of the paper, which signs optimal policy interventions. We show that the government should tax automation on efficiency grounds, even when it does not have a preference for redistribution.

Efficiency vs. redistribution. Consider a change in automation $\delta\alpha$ starting from the laissez-faire. This induces a change in welfare $\delta\mathcal{U}\left(\alpha;\left\{\eta^{h}\right\}\right)$, where $\mathcal{U}$ is the value of the problem in Lemma 4.1 conditional on α. This change has two effects on welfare. The first effect is *aggregate*: the change $\delta\alpha$ affects the path of aggregate output, which is captured by $\Delta_t^{\star}$. The second effect is *distributional*: it affects the relative incomes of workers, which is captured by the distributional pecuniary externalities $\Phi^{\star,\text{distrib}}$. The two effects correspond to the *aggregate efficiency* and *redistribution* components of the decomposition in Bhandari et al. (2021).[23] The Pareto weights $\left\{\eta^{h}\right\}$ determine the relative contribution of these two effects to the change in welfare $\delta\mathcal{U}\left(\alpha;\left\{\eta^{h}\right\}\right)$.

To highlight the new rationale for policy intervention that we propose, we introduce *efficiency weights* such that the government only values aggregate efficiency and not redistribution.[24] We show in Appendix A.9 that our results carry through if we allow instead the government to offset the distributional effects of the intervention with some additional tools, as in Costinot and Werning (2022).

Definition 2 (Efficiency weights)**.** The efficiency weights $\left\{\eta^{h,\text{effic}}\right\}$ are such that

[22] Labor reallocation — just like automation — is isomorphic to a technological choice in the Arrow-Debreu construct. Each worker owns a firm that chooses the type of labor services to provide.

[23] Bhandari et al. (2021) decompose the welfare effects of policy changes into gains in: (i) aggregate efficiency from changes in total resources, (ii) redistribution from changes in ex ante consumption shares, and (iii) insurance from changes in consumption risk. In our baseline model, taxing automation affects welfare via (i) and (ii) alone. In our quantitative model, (iii) is also present.

[24] Absent borrowing constraints, these efficiency weights take the familiar form $1/\eta^{\text{effic},h} \propto u'\left(c_0^h\right)$. More generally, they imply that the government values displaced workers *less* compared to a utilitarian government that values equity. See Appendices A.5–A.6 for details.

$\Phi^{\star,\text{distrib}}\left(\alpha^{\text{LF}}, T\left(\alpha^{\text{LF}}\right); \left\{\eta^{h,\text{effic}}\right\}\right) = 0$. That is, the distributive pecuniary externalities of automation net out when evaluated at the laissez-faire.

We now show that the government finds it optimal to curb (or tax) automation on efficiency grounds alone. In the following, we assume that the government's objective $\mathcal{U}$ is concave in α when using the weights $\left\{\eta^{h,\text{effic}}\right\}$.[25] This ensures that there is a unique local maximum $\alpha^{\text{SB,effic}}$.

Proposition 1 (Taxing automation on efficiency grounds)**.** *Automation at the laissez-faire α^{LF} is excessive compared to the second best $\alpha^{SB,effic}$ associated with efficiency weights. Taxing automation is optimal on efficiency grounds.*

Proof. See Appendix A.5. □

To understand the result, consider the *private* incentives to automate starting at the laissez-faire

$$\text{(LF)} \qquad \int_0^{+\infty} \exp\left(-\rho t\right) \frac{u'\left(c_t^N\right)}{u'\left(c_0^N\right)} \Delta_t^{\star} dt \; = 0 \tag{4.2}$$

where $\Delta_t^{\star}$ is the response of output to automation and is given by (3.8). The gains from changing automation $\Delta_t^{\star}$ are zero in present discounted value.

Now, consider the *social* incentives to automate starting at the laissez-faire

$$\text{(SB)} \qquad \int_0^{+\infty} \exp\left(-\rho t\right) \left\{\sum_h \frac{1}{2} \eta^{h,\text{effic}} \frac{u'\left(c_t^h\right)}{u'\left(c_0^h\right)}\right\} \Delta_t^{\star} dt \;\; < 0. \tag{4.3}$$

To understand why this is the case, note that the government is effectively more impatient than the firm. It puts a positive weight on displaced workers who become borrowing constrained and have steeper consumption profiles than non-automated workers who price the firms' equity. Moreover, the gains from automation $\Delta_t^{\star}$ increase over time (Lemma 3). They are negative initially as the firm incurs a cost when investing in automation ($\Delta_t^{\star} < 0$ for small t), and they build over time as workers reallocate ($\Delta_t^{\star} > 0$ for large t). Thus, the government values earlier flows $\Delta_t^{\star}$ more compared to the firm, precisely when they are negative, and it values future flows less when they are positive. As a result, the government finds it

[25] Otherwise, our results only apply locally. That is, a decrease in automation $\delta\alpha < 0$ implies $\delta\mathcal{U}\left(\cdot\right) > 0$ starting from the laissez-faire.

optimal to tax automation on efficiency grounds. The following remark summarizes these insights.

Remark 3. *Firms (partly) overlook that the gains from automation take time to materialize. The optimal tax on automation improves aggregate efficiency. It raises consumption early on in the transition, precisely when displaced workers are borrowing-constrained.*

4.4 Extensions

We next consider a number of extensions to our analysis.

4.4.1 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 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.

Proposition 2 (Third best)**.** *Suppose that the government only controls automation — but the reallocation choice T must be consistent with workers' optimization. This strengthens the rationale for taxing automation on efficiency grounds.*

Proof. See Appendix A.6. □

Again, it is useful to inspect the social incentives to automate starting from the laissez-faire

$$\text{(TB)} \quad \int_0^{+\infty} \exp\left(-\rho t\right) \sum_h \frac{1}{2} \eta^{h,\text{effic}} \frac{u'\left(c_t^h\right)}{u'\left(c_0^h\right)} \times \left\{ \Delta_t^\star + \mathbf{1}_{\left\{t > T\left(\alpha^{\text{LF}}\right)\right\}} \frac{1}{2} \lambda \exp\left(-\lambda T\left(\alpha^{\text{LF}}\right)\right) T'\left(\alpha^{\text{LF}}\right) \Delta_t \right\} dt \; < 0. \quad (4.4)$$

To understand why this is the case, focus on the additional terms involving the response of output to reallocation Δ_t defined in (3.3). The government now also internalizes the indirect effect of automation on output Δ_t due to the reallocation it induces in equilibrium $T'(\cdot) > 0$. The flows Δ_t decrease over time, since labor reallocation closes the gap between the marginal products of labor across occupations. Compared to the government, displaced workers put an excessive weight on early, positive payoffs: binding constraints incentivize them to rely on mobility to self-insure. This indirect effect reinforces the government's desire to *tax* automation.

4.4.2 Equity Concerns

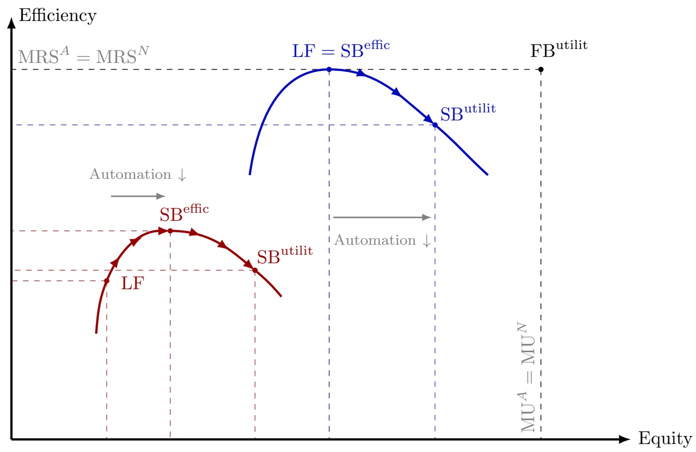
We now introduce equity concerns in our model. This allows us to clarify our contribution relative to the literature on the taxation of automation on equity grounds. Proposition 5 in Appendix A.7 shows that a government with a preference for redistribution curbs automation even if the economy is efficient. Figure 4.1 illustrates this result schematically and connects back to our previous results.

Automation has distributional effects: it reduces equity at the laissez-faire (LF in the figure) relative to the first best of a utilitarian planner ($\text{FB}^{\text{utilit}}$). Displaced workers are worse off and their marginal utility is (persistently) higher than other workers' $\text{MU}^A > \text{MU}^N$. In an efficient economy (blue line in the figure), the intertemporal marginal rates of substitution of displaced workers coincide with the equilibrium interest rate faced by firms who automate $\text{MRS}^A = \text{MRS}^N$. The government does not intervene ($\text{LF} = \text{SB}^{\text{effic}}$) unless it has a preference for redistribution ($\text{SB}^{\text{utilit}}$), in which case it taxes automation and sacrifices efficiency to improve equity. This is the canonical trade-off between equity and efficiency emphasized in the literature on the regulation of automation. In an inefficient economy, there is a wedge between the (intertemporal) marginal rate of substitutions of different workers $\text{MRS}^A < \text{MRS}^N$. Firms are effectively too patient: automation is inefficient. The government can improve *both* efficiency and equity by taxing automation, i.e., there is no trade-off.

4.4.3 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. These two rationales share two features: they rely on the presence on uninsured idiosyncratic risk and optimal policies affect investment in the *long-run*.

Figure 4.1: Second best with efficiency ($\underline{a} \to -\infty$) and inefficiency ($\underline{a} \to 0$)



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 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 that changes in automation are subject to a convex adjustment cost. 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 associated with Pareto weights $\eta_s^h = \exp(-(\rho+\chi)\, s)$ for all generations $s \in (-\infty, +\infty)$ and occupations $h = \{A, N\}$. 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.8. □

The government can neither improve 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. That said, some workers could remain borrowing constrained and have different marginal utilities in richer environments with uninsured income risk (as in our quantitative model). This creates a motive for policy intervention in the long-run too (see footnote 37).

4.4.4 The Direction of Investments

So far, firms 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 firms choose 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 forms of investment, e.g., through lower investment subsidies as in South Korea.

5 Quantitative Model

In the remaining of this paper, we quantitatively evaluate the efficiency rationale for slowing down automation — even when allowing for various redistributive instruments. To this end, we enrich our baseline model along several dimensions that are important for a credible normative analysis. In particular, we allow for gradual automation, overlapping generations of workers, gross flows across occupations, uninsurable idyosincratic earnings and mobility risks, and unemployment benefits and progressive taxation. Appendix B provides further details on the quantitative model.

5.1 Firms

Production. There is a continuum of occupations of mass 1. A share ϕ are automatable $(h = A)$ and a share $1-\phi$ are non-automatable $(h = N)$. Occupations operate the technology

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

for some elasticity $\eta \in (0,1)$, relative productivity of automation $\varphi > 0$, and productivities $A^h > 0$.[26] The aggregate 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. Automated occupations $h = A$ rent the stock of automation α_t on spot markets (Guerreiro et al., 2017) at rate $\{r_t^\star\}$ from a mutual fund.

Investment. A competitive mutual fund invests workers' savings in government's bonds and automation. The law of motion of automation is

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

where δ is the rate of depreciation, and x_t is the investment rate. Investment is subject to quadratic adjustment cost $\Omega\left(x_t; \alpha_t\right) = \omega\left(x_t/\alpha_t - \delta\right)^2 \alpha_t$.[27] In particular, the effective price of investment x_t falls as automation α_t increases.[28] The government taxes automation linearly at rate $\{\tau_t^x\}$ and rebates the revenue to the mutual fund.

[26] The example in Section 2.1 micro-founds this technology using the task-based model of Acemoglu and Restrepo (2018).

[27] This specification provides a micro-foundation for the cost of automation in our baseline model. The production function net of investment is $\hat{F}(\mu;\alpha) \equiv A\left(\varphi^A\alpha + \mu\right)^{1-\eta} - \delta\alpha$ at the steady state.

[28] This captures the price decline of automation technologies over time (Graetz and Michaels, 2018).

5.2 Workers

There are overlapping generations of workers that are replaced at rate χ.[29] 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. Workers invest in the mutual fund with return $\{r_t\}$. In addition, they 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

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

where $\mathcal{Y}_t^{\text{net}}(\mathbf{x})$ denotes net income and r_t is the return on the mutual fund. 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 variance $\gamma > 0$, 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

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

non-automated $(h = N)$ occupations, and the parameter γ governs the elasticity of labor supply. Workers who reallocate go through unemployment / retraining spells which they exit at rate κ. Upon entering their new occupation, workers experience a permanent productivity loss θ. We assume that workers experience this loss only the first time they reallocate.

Income. Employed workers $(e = E)$ earn a 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 employment status switches to $e_t = U$ upon reallocation and reverts to $e_t = E$ upon exiting unemployment. All workers are born with $e_t = E$. 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 $\sigma_z > 0$. Following Krueger et al. (2016), we suppose that unemployed workers $(e = U)$ receive benefits that are proportional to the gross labor income they would have earned if they had remained employed in their previous occupation. The replacement rate is $b \in [0,1]$, and we assume that these earnings take the form of home production (Alvarez and Shimer, 2011).[30] We suppose that workers claim profits in proportion to their idiosyncratic (mean-reverting) productivity, as in Auclert et al. (2018).[31] Workers net income is

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

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

[30] This last assumption is mostly innocuous. Its only purpose is to avoid introducing an additional motive for distorsionary taxation to finance unemployment insurance.

[31] 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.

5.3 Policy and Equilibrium

The government's flow budget constraint is

$$dB_t = (T_t + r_t B_t - G_t)\, dt \tag{5.8}$$

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

$$\int a_t(\mathbf{x})\, d\pi_t = -B_t \tag{5.9}$$

The wages are still given by (3.1). The rental rate of automation adjusts so that the firm's demand for automation α_t^A equals the supply α_t from the mutual fund. We normalize the final good price to 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, Appendix C provides 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 1970. Table 6.1 shows the parameterization.

External calibration. External parameters are set to standard values in the literature. The initial labor share $1 - \eta$ is 0.64 based on BLS data. The depreciation rate δ is 10%, as in Graetz and Michaels (2018). The elasticity of substitution across occupations ν is 0.75, in between the values in Buera and Kaboski (2009) and Buera et al. (2011).[32] The inverse elasticity of intertemporal substitution σ is 2. We set the

[32] We interpret automated occupations as routine-intensive ones which are well represented in manufacturing. Accordingly, we set the elasticity of substitution between automated and non-

replacement rate χ to obtain an average active life of 50 years. 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 mean reversion ρ_z 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 0.75 at the initial and final steady states, which lies between the values used by Kaplan et al. (2018) and McKay et al. (2016). During the transition, we let the supply of liquidity converge exponentially to its long run level with a half-life of roughly 15 years, following Guerrieri and Lorenzoni (2017).

Internal calibration. We calibrate eight parameters internally: the discount rate (ρ); the mobility hazard (λ); the Fréchet parameter (γ); the occupations' productivities $\left(A^h\right)$; the share of automated occupations (ϕ); the productivity of automation (φ); and the adjustment cost for automation (ω). We pick these to jointly match eight moments. The discount rate targets an annualized real interest rate of 4 percent. We adjust the mobility hazard to match an occupational mobility rate of 10% per year at the initial steady state, which corresponds to the U.S. level in 1970 in Kambourov and Manovskii (2008). The Fréchet parameter targets an elasticity of labor supply of 2 for the stock of workers (i.e., all generations) following Hsieh et al. (2019).[33] The occupations' productivity $\left\{A^h\right\}$ are such that output is 1 and wages are identical across occupations at the initial stationary equilibrium. The mass of automated occupations ϕ targets an employment share of 56% in routine occupations in 1970 (Bharadwaj and Dvorkin, 2019). We choose the productivity of automation φ to match a labor share of 56% in the final steady state, which is the

automated occupations to that between manufacturing and other sectors. The structural change literature strongly suggests that the these occupations are gross complements.

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

Table 6.1: Calibration

Parameter	Description	Calibration	Target / Source
Workers			
ρ	Discount rate	0.102	4% real interest rate
σ	EIS (inverse)	2	-
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	$(0.938, 1.157)$	Initial output (1) and symmetric wages
$1-\eta$	Initial labor share	0.64	1970 labor share (BLS)
δ	Depreciation rate	0.1	Graetz and Michaels (2018)
ϕ	Share of automated occupations	0.546	Routine occs. employment share in 1970
φ	Productivity of automation	0.43	Final labor share
ω	Adjustment cost	4	Half-life of automation
ν	Elasticity of subst. across occs.	0.75	(Buera and Kaboski, 2009; Buera et al., 2011)
Mobility frictions			
λ	Mobility hazard	0.312	Occupational mobility rate in 1970
$1/\kappa$	Average unemployment duration	1/3.2	Alvarez and Shimer (2011)
θ	Productivity loss from relocation	0.18	Kambourov and Manovskii (2009)
γ	Fréchet parameter	0.052	Elasticity of labor supply
Government			
ψ_0	Tax intercept	0.35	BEA
ψ_1	Tax elasticity	0.181	Heathcote et al. (2017)
$-B/Y$	Liquidity / GDP	0.75	Liquid assets / GDP (Kaplan et al., 2018)
Income process			
ρ_z	Mean reversion	0.0228	Floden and Lindé (2001)
σ_z	Volatility	0.1025	Floden and Lindé (2001)
b	Replacement rate	0.4	Ganong et al. (2020)

empirical value in 2020 (Bergholt et al., 2022). Finally, we choose the investment adjustment cost ω so that automation converges to its long-run level with a half-life of 20 years.[34]

Untargeted moments. The model matches well several untargeted moments (see Appendix D.1 for details). First, the share of hand-to-mouth workers is roughly 17% at the initial steady state, which lies between the estimates of Kaplan et al. (2014) and Aguiar et al. (2020). Second, we obtain that 72% of output in occupation $h = A$ is produced by automation at the final steady state. For comparison, the McKinsey (2017) report finds that in occupations most susceptible to automation — 51% of employment compared to 56% in our model — roughly 70% of output previously produced by labor could be automated. Third, the (partial equilibrium) effects of automation on employment and labor productivity in our model are comparable to the firm-level estimates in Bonfiglioli et al. (2022). They find that the causal effect of adopting automation reduces employment by 54% at the firm level and increases value added *per worker* by 174%, compared to 62% and 163% in partial equilibrium in our model or 48% and 184% across steady states..

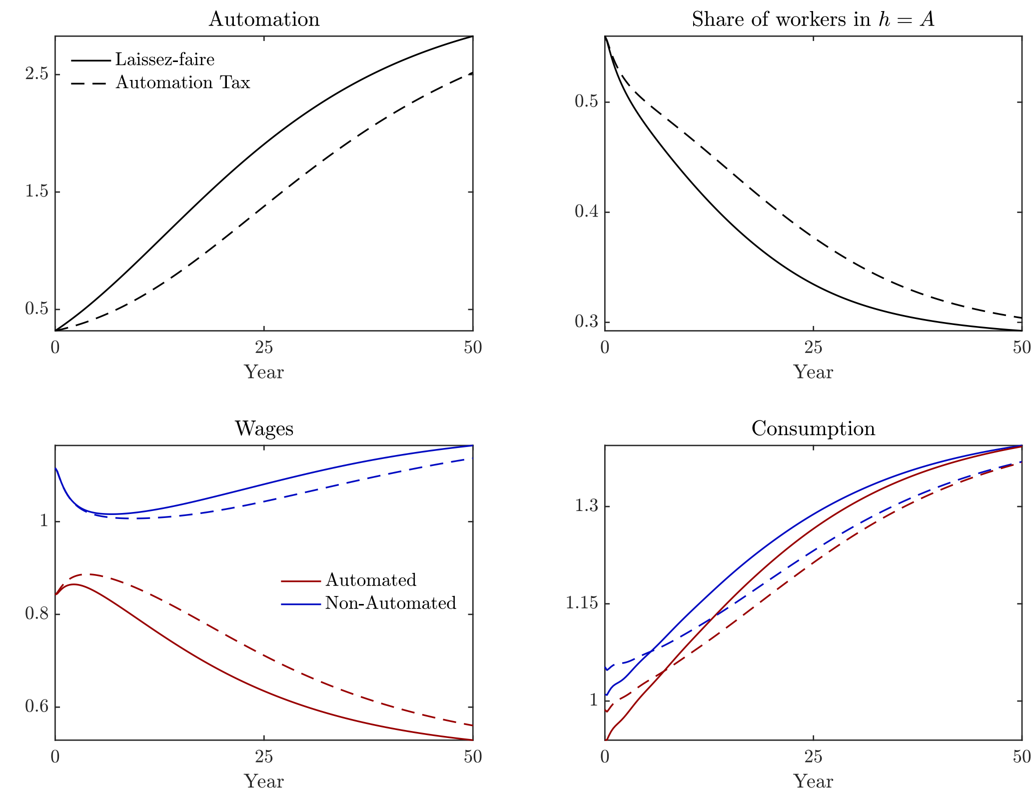
6.2 Automation, Reallocation and Inequality

We start by simulating the transition of our economy to its long-run steady state with automation. The economy is initially at its steady state with $\varphi = 0$ and no automation takes place ($\alpha = 0$). In period $t = 0$, automation becomes possible ($\varphi > 0$). The initial equilibrium without automation is now an unstable steady state. We endow the firm with a small initial stock of automation $\alpha_0 > 0$, which moves the economy away from this unstable steady state and initiates the convergence to the new long-run (stable) steady state with automation. We choose the initial stock α_0 to be $1/10$ of its long-run level.[35]

[34] This is a typical convergence rate in neoclassical growth models. For comparison, the data of Acemoglu and Restrepo (2020) shows that the stock of industrial robots doubles roughly every 10 years, i.e., a half-life of 10 years. This is also consistent with the evidence in Bharadwaj and Dvorkin (2019). Assuming a longer half-life (20 years) dampens our mechanism by limiting the consequences of automation early on during the transition.

[35] For comparison, the earliest reliable data on industrial robots puts their stock in the early 1990s at roughly 1/5 of its 2020 counterpart (Acemoglu and Restrepo, 2020). If anything, choosing a lower initial stock dampens our mechanism by limiting the impact of automation early on during the transition.

Figure 6.1: Allocations



Notes: Solid curves correspond to the laissez-faire and dashed curves to the equilibrium with the automation tax. Red curves denote workers initially in automated occupations and blue curves in non-automated ones. Wages and consumptions are normalized by their initial steady state levels.

Figure 6.1 illustrates the transition at the laissez-faire (solid lines). Automation converges to its steady state with a half-life of 20 years (a targeted moment). 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. Finally, automated workers consume less and have steeper consumption profiles — their MRS is lower — as they are more likely to become borrowing constrained.

In terms of magnitudes, the wage gap between occupations is 50% after 30 years, which almost exactly matches the (composition-adjusted) wage gap measured by Cortes (2016) for the U.S. in 2007. Our baseline calibration allows automated workers to reallocate too rapidly compared to the data. For instance, it takes only 15 years in our model for the employment share in automated occupations to reach 40%, compared to roughly 45 years in the data (Bharadwaj and Dvorkin, 2019).[36] However, note that this rapid reallocation is conservative with respect to our mechanism, since workers can self-insure by moving away from automated occupations. We explore later on the sensitivity of our results to key parameters governing the speed of reallocation.

The same figure illustrates the effect of slowing down automation (dashed lines). The sequence of distortionary taxes on automation $\{\tau_t^x\}$ that we feed in are such that the half-life of automation increases to roughly 25 years. As expected, labor reallocation slows down and so does the fall in wages and consumption in automated occupations. Finally, consumption profiles become flatter and the wedge between MRSs closes faster, as the share of automated workers who are constrained is much less persistent.

6.3 Second Best Policies and Welfare

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

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

where $V_t^{\text{birth}}(\mathbf{x})$ is the value of a worker born in period t that draws a state $\mathbf{x}$, and $\eta_t(\mathbf{x})$ are Pareto weights. The government maximizes this objective by choosing taxes on investment $\{\tau_t^x\}$ and rebating the proceedings to the mutual fund.

As in our tractable model, we work with the primal problem. Solving for the exact sequence of $\{\alpha_t\}$ is computationally challenging and beyond the scope of this paper. Instead, we restrict our attention to simpler (or arguably more realistic) parametric perturbations of this sequence. Details are provided in Appendix B.3.[37] For each of these perturbations, we compute the transition dynamics and evaluate welfare (6.1). We then find the second best sequence of automation $\{\alpha_t^{\text{SB}}\}$ and calculate welfare gains $\Delta\mathcal{W}$ (in consumption equivalent terms) relative to the laissez-faire. We repeat this exercise using *efficiency* and *utilitarian* weights $\eta_t(\mathbf{x})$

[36] This is the case despite automation taking place more slowly than in the data (footnote 34).

[37] In particular, we do not constrain automation to converge to its laissez-faire level in the long-run. The reason is that our quantitative model also features uninsured idiosyncratic risk which introduces an additional motive for intervention. It is well known that a long-run tax (or subsidy) on *capital* can be optimal when markets are incomplete (Section 4.4.3). However, we find that long-run interventions produce modest improvements in the government's objective (6.1).

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

		Alternative calibrations		Alternative policies	
	Benchmark	Low $1/\gamma$	High $1/\kappa$	Transfers	Joint
Efficiency	3.8%	4.5%	3.5%	0.3%	3.9%
Utilititarian	5.9%	6.6%	5.8%	3.0%	8.7%

Note: 'Benchmark' corresponds to the gains from optimal automation taxes under the calibration described in Section 6.1. 'Low $1/\gamma$' and 'High $1/\kappa$' denote alternative calibrations with γ chosen to match a lower labor supply elasticity of 1 and κ chosen to match a longer average unemployment duration of 2 years. 'Transfers' corresponds to the gains from an alternative policy that transfers \$10k to automated workers at time $t = 0$ financed with government debt. 'Joint' combines both optimal automation taxes and targeted transfers. 'Efficiency' and 'Utilitarian' compute the gains and optimal automation taxes using the two Pareto weights (Appendix B.3).

(Appendix B.3). Efficiency weights are inversely related to workers' marginal utility at birth, whereas utilitarian weights are symmetric within generations.

Table 6.2 reports our findings. In our benchmark calibration, the government finds it optimal to slow down automation substantially *on efficiency grounds.* The optimal half-life is about 46.5 years — more than double the half-life at the laissez faire — and this policy achieves sizable welfare gains of roughly 4%. The gains are even larger with utilitarian weights (roughly 6%) since slowing down automation improves not only efficiency but also equity.[38] For example, the welfare gap (in consumption equivalent terms) between the average automated and non-automated worker decreases by 2.2%. As anticipated in Remark 3, we find in our simulations that these welfare gains are achieved by flattening consumption profiles and raising consumption early on during the transition when displaced workers value it more.

Alternative calibrations. We consider two alternative calibrations of our model.[39] The goal is to explore the sensitivity of results to two important features that affect workers' reallocation. First, we increase the Fréchet parameter γ to match a

[38] The optimal speed of automation turns out to be similar in both cases because the additional incentives to redistribute are small when automation takes place sufficiently slowly.

[39] For each of these two alternative calibrations, we re-calibrate the rest of the parameters to match the same moments as in our benchmark.

lower elasticity of labor supply of 1. As shown in Figure D.1 in the Appendix, the share of workers in automated occupations declines more slowly and less overall. It reaches 40% after 25 years and converges to 37% in the long-run. This figure is closer to the 41% share in the data in 2020, which our benchmark calibration missed. However, the model now overpredicts the wage gap across occupations since less labor reallocation takes place. As a result, the consumption profiles of displaced workers are steeper, and the second best intervention produces larger welfare gains. The optimal half-life of automation is now larger too (48 years with efficiency weights).

Our second alternative calibration increases the average duration of unemployment spells ($1/\kappa$) to 2 years with the idea that displaced workers could take more than the 3 months needed by the typical U.S. worker to exit unemployment.[40] We find that the optimal half-life of automation increases to 52 years as a larger share of displaced workers become borrowing constrained. The welfare gains from slowing down automation are comparable to our benchmark. Note that our benchmark calibration with shorter unemployment can also be interpreted as one where retraining is shorter, e.g., due to government active labor market interventions. This suggests that slowing down automation is desirable even in this case.

Finally, we consider additional calibrations where we vary technological parameters that govern the overall gains from automation and its speed. Specifically, we vary the elasticity of substitution across occupations ν, the productivity of automation φ, the adjustment cost ω, and the initial stock of automation α_0. We also explore the role of liquidity to GDP $(-B/Y)$. All results are presented and discussed in Appendix D.3.

Targeted transfers. Government transfers that target automated workers could in principle be an effective tool to respond to automation. In particular, we argued in Section 3.4 that a government could implement a first best allocation without taxing automation if the transfers fully alleviate the borrowing constraints. We allow for realistic targeted transfers in the following, and compare the welfare gains that they produce to those from the optimal tax on automation. Specifically, at time

[40] We suppose that unemployment benefits last for the entire duration of the reallocation spells. This policy is similar to Trade Readjustment Allowances which extend benefits to workers negatively affected by foreign imports while they retrain.

$t = 0$, the government gives a transfer of \$10k to workers initially employed in automated occupations. The transfers are financed via an increase in debt. These transfers are rather generous: they correspond to the maximum amount allowed by the Reemployment Trade Adjustment Assistance program (RTAA).[41]

The fourth column in the table shows that targeted transfers of this magnitude are not sufficient to improve aggregate efficiency. The welfare gains are 0.3% under the efficiency weights, which is smaller than the 3.8% gains from the automation tax (first column). The reason is twofold. First, transfers of this magnitude do not relax borrowing constraints as much as the optimal tax on automation. Second, they need to be financed with taxes on non-automated workers and new generations, some of whom are also borrowing constrained.

Together, these results imply that targeted transfers of a realistic magnitude are an effective tool for redistributing towards automated workers but do little to alleviate borrowing constraints in the medium-run and address inefficient automation. Finally, we combine the optimal automation tax with targeted transfers which delivers substantially higher welfare gains when the government is utilitarian.

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. Firms fail to internalize that workers displaced by automation have a limited ability to smooth consumption while they reallocate. 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 improves aggregate efficiency, raising consumption early on in the transition precisely when displaced workers value it more. Quantitatively, we found that slowing down automation achieves substantial efficiency and welfare gains, even when the government can implement generous transfers to displaced workers.

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

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

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 for automated cashiers in the Swiss canton of Geneva or automonous vehicles used by transportation companies in Nevada. 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 is crucially determined by 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.

——— (2020): "Robots and jobs: Evidence from US labor markets," *Journal of Political Economy*, 128, 2188–2244.

——— (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.

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.

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

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

BERGHOLT, D., F. FURLANETTO, AND N. MAFFEI-FACCIOLI (2022): "The decline of the labor share: new empirical evidence," *American Economic Journal: Macroeconomics*, 14, 163–98.

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., K. F. HERKENHOFF, AND G. M. PHILLIPS (2020): "Can the Unemployed Borrow? Implications for Public Insurance," Working paper.

BUERA, F. J. AND J. P. KABOSKI (2009): "Can Traditional Theories of Structural Change Fit The Data?" *Journal of the European Economic Association*, 7, 469–477.

BUERA, F. J., J. P. KABOSKI, AND Y. SHIN (2011): "Finance and Development: A Tale of Two Sectors," *American Economic Review*, 101, 1964–2002.

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, 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.

DIAMOND, P. A. AND J. A. MIRRLEES (1971): "Optimal Taxation and Public Production: I–Production Efficiency," *American Economic Review*, 61, 8–27.

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 (2017): "Should Robots be Taxed?" NBER Working Papers 23806, National Bureau of Economic Research, Inc.

GUERRIERI, V. AND G. LORENZONI (2017): "Credit Crises, Precautionary Savings, and the Liquidity Trap," *The Quarterly Journal of Economics*, 132, 1427–1467.

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*.

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

JACOBSON, L. S., R. J. LALONDE, AND D. G. SULLIVAN (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.

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*.

LORENZONI, G. (2008): "Inefficient Credit Booms," *Review of Economic Studies*, 75, 809–833.

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.

MCKAY, A., E. NAKAMURA, AND J. STEINSSON (2016): "The Power of Forward Guidance Revisited," *American Economic Review*, 106, 3133–3158.

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

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

POTERBA, J. M. (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.

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 all workers who are able to $(m_t = 1)$ if and only if the present discounted value of the 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.

We next show that there exists some $T^{\text{LF}} > 0$ such that all workers who can reallocate do so $(m_t = 1)$ for all $t \in \left[0, T^{\text{LF}}\right)$.

A sufficient condition is that

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

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 $\alpha = \alpha^{\text{LF}}$. Consumption and wages are constant over time when $T = 0$, so inequality (A.3) holds if and only if

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

where $\alpha = \alpha^{\text{LF}}$. This necessarily holds by Assumption 3.

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 $a^{\star}(\lambda, \kappa)$.

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 such that $a_S^A > 0$ and $a_S^N < 0$ but all workers are still unconstrained $a_S^N > \underline{a}$. 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

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

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

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

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

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

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),

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

with weights

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

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. Furthermore, $a_S^N < a_S^A$ under our postulate. Therefore, $d\left(a_t^N - a_t^A\right) > 0$. 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 below, 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 (A.8) over time and using (A.5) gives the assets of automated workers

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

if they were never to become borrowing constrained. The sequence $\left\{a_t^A\right\}$ depends on reallocation frictions (λ, κ) but not the borrowing limit $\underline{a}$. Let $a^\star(\lambda, \kappa) \equiv$

$\inf a_t^A$ be the lowest value attained by this sequence. We have shown above that $a^\star(\lambda,\kappa) < 0$. 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 $\inf a_t^A$. 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 $a^\star(\lambda,\kappa) \equiv \inf a_t^A \to 0$. 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 $1/\kappa > 0$. In this case, there is always a (small) borrowing limit $\underline{a} > 0$ such that automated workers become borrowing constrained (for example $\underline{a} = 0$).

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 sector, 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 probability of becoming unemployed ι is sufficiently small that output still increases over time despite some workers becoming unemployed.

Output increases over time when

$$\partial_t G^\star\left(\mu_t^A, \mu_t^N; \alpha\right) = G_A^\star(\cdot)\,\partial_t \mu_t^A + (1-\theta)\,G_N^\star(\cdot)\,\partial_t \mu_t^N > 0, \tag{A.13}$$

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 probability of unemployment and the productivity loss of reallocation are sufficiently small $1-(1-\theta)(1-\iota) \to 0$. Note that μ_t^A, μ_t^N and α are continuous in (θ,ι) at the laissez-faire. Therefore, there exists some threshold $Z^\star > 0$ such that (A.13) still holds for all (θ,ι) such that $1-(1-\theta)(1-\iota) < Z^\star$.

It remains to show that the sequence $\{z_t\}$ decreases over time when $1-(1-\theta)\times$

$(1-\iota) < Z^\star$. 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 $1-(1-\theta)(1-\iota) < Z^\star$. Using labor incomes (2.8) and the effective labor supplies (3.4)–(3.5),

$$\begin{aligned}\partial_t \left(\hat{\mathcal{Y}}_t^N - \hat{\mathcal{Y}}_t^A\right) = & - \left\{ G_{AA}(\cdot)\,\partial_t \mu_t^A + G_{AN}(\cdot)\,\partial_t \mu_t^N \right\} \mu_t^A - G_A(\cdot)\,\partial_t \mu_t^A \\ & + \left(\frac{1}{2} - (1-\theta)\,\tilde{\mu}_t\right) \left\{ G_{NA}(\cdot)\,\partial_t \mu_t^A + G_{NN}(\cdot)\,\partial_t \mu_t^N \right\} \\ & - (1-\theta)\, G_N(\cdot)\,\partial_t \tilde{\mu}_t \end{aligned} \tag{A.14}$$

Therefore,

$$\partial_t \left(\hat{\mathcal{Y}}_t^N - \hat{\mathcal{Y}}_t^A\right) < - \left\{ G_A(\cdot)\,\partial_t \mu_t^A + (1-\theta)\, G_N(\cdot)\,\partial_t \tilde{\mu}_t \right\} < 0 \tag{A.15}$$

using $G_{AA}(\cdot) < 0$ and $G_{NN}(\cdot) < 0$ since G is neoclassical, $G_{AN}(\cdot) > 0$ by Assumption 1, and $\partial_t \mu_t^A < 0$ and $\partial_t \mu_t^N > 0$ in equilibrium. It follows that $\partial_t \left(\hat{\mathcal{Y}}_t^N - \hat{\mathcal{Y}}_t^A\right) < 0$ when $1-(1-\theta)(1-\iota) < Z^\star$. Taken together, the inequalities (A.13) and (A.15) 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 (a vanishing mass) of workers can trade both. In Appendix A.2, we have shown that non-automated workers are not borrowing constrained, i.e., they are on their Euler equation. Therefore, $\exp(-\rho t)\, u'\left(c_t^N\right)/u'\left(c_0^N\right) = \exp\left(-\int_0^t r_s ds\right)$. Next, we show that the solution to the firm's automation problem is interior and unique. Using a standard variational argument, a necessary condition for an interior optimum is

$$\int_0^{+\infty} \exp\left(-\int_0^t r_s ds\right) \frac{\partial}{\partial \alpha} \Pi_t(\alpha)\, dt = 0. \tag{A.16}$$

Furthermore, the following envelope condition applies

$$\frac{d}{d\alpha} \Pi_t(\alpha) = \frac{\partial}{\partial \alpha} G^\star(\cdot). \tag{A.17}$$

Therefore, the condition

$$\int_0^{+\infty} \exp\left(-\int_0^t r_s ds\right) \frac{\partial}{\partial \alpha} G^\star(\cdot) = 0 \tag{A.18}$$

is necessary. By Assumption 2, this condition is also sufficient, and the solution is unique and interior.

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

$$\Delta_t^\star \equiv \frac{\partial}{\partial \alpha} G^\star(\cdot) \equiv G_A\left(F\left(\mu_t^A;\alpha\right), F\left(\mu_t^N;0\right)\right) F_\alpha\left(\mu_t^A;\alpha\right) - \mathcal{C}'(\alpha) \tag{A.19}$$

Therefore,

$$\partial_t \Delta_t^\star \geq G_{AN}\left(F\left(\mu_t^A;\alpha\right), F\left(\mu_t^N;0\right)\right) F_\mu\left(\mu_t^N\right) F_\alpha\left(\mu_t^A;\alpha\right) \partial_t \mu_t^N, \tag{A.20}$$

using $G_{AA}(\cdot) < 0$ since G is neoclassical, $F_{\alpha\mu}(\cdot) < 0$ by Assumption 1, and $\partial_t \mu_t^A < 0$ in equilibrium. It follows that $\partial_t \Delta_t^\star > 0$ since $G_{AN}(\cdot) > 0$ by Assumption 1.

A.4 Constrained Inefficiency

Proposition 4 (Constrained inefficiency)**.** *Fix the production function $G^\star$. Suppose that the* laissez-faire *is constrained efficient for some Pareto weights $\{\eta^A, \eta^N\}$. 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 second best and* laissez-faire *for this alternative economy do not coincide for all $0 < \epsilon \leq \bar{\epsilon}$.*

The government's optimality conditions to reallocate and automate are

$$\int_{T^{\rm SB}}^{+\infty} \exp(-\rho t) \frac{u'\left(c_t^A\right)}{u'\left(c_0^A\right)} \Delta_t dt = \Phi\left(\alpha^{\rm SB}, T^{\rm SB}; \eta^A, \eta^N\right) \tag{A.21}$$

and

$$\int_0^{+\infty} \exp(-\rho t) \frac{u'\left(c_t^N\right)}{u'\left(c_0^N\right)} \Delta_t^\star dt = \Phi^\star\left(\alpha^{\rm SB}, T^{\rm SB}; \eta^A, \eta^N\right) \tag{A.22}$$

respectively. The terms on the left-hand side of (A.21)–(A.22) correspond to the private incentives to automate and reallocate, respectively. The terms on the right-hand capture pecuniary externalities that affect workers through labor incomes and profits — which firms and workers do not internalize. These pecuniary externalities are given by[42]

$$\Phi\left(\alpha^{SB},T^{SB};\eta^{A},\eta^{N}\right)\equiv\int_{T^{SB}}^{+\infty}\exp\left(-\rho t\right)\hat{\Phi}_{t}\left(\cdot\right)dt \tag{A.23}$$

$$\Phi^{\star}\left(\alpha^{SB},T^{SB};\eta^{A},\eta^{N}\right)\equiv\int_{0}^{+\infty}\exp\left(-\rho t\right)\hat{\Phi}_{t}^{\star}\left(\cdot\right)dt \tag{A.24}$$

where[43]

$$\begin{aligned}\hat{\Phi}_{t}\left(\cdot\right)\equiv&-\frac{\eta^{N}}{\eta^{A}}\frac{u'\left(c_{t}^{N}\right)}{u'\left(c_{0}^{N}\right)}\left\{\hat{w}_{t}^{N}-\sum_{h}\mu_{t}^{h}\hat{w}_{t}^{h}\right\}\\&-\frac{u'\left(c_{t}^{A}\right)}{u'\left(c_{0}^{A}\right)}\left\{\hat{w}_{t}^{A}+2\times\left[\left(\frac{1}{2}-\mu_{t}^{A}\right)\times\left(\left(1-\theta\right)\hat{w}_{t}^{N}-\hat{w}_{t}^{A}\right)\right.\right.\\&\left.\left.\qquad-\left(1-\mu_{t}^{A}-\mu_{t}^{N}\right)\times\left(1-\theta\right)\hat{w}_{t}^{N}\right]-\sum_{h}\mu_{t}^{h}\hat{w}_{t}^{h}\right\}\end{aligned} \tag{A.25}$$

$$\begin{aligned}\hat{\Phi}_{t}^{\star}\left(\cdot\right)\equiv&-\frac{u'\left(c_{t}^{N}\right)}{u'\left(c_{0}^{N}\right)}\left\{\hat{w}_{t}^{N,\star}-\sum_{h}\mu_{t}^{h}\hat{w}_{t}^{h,\star}\right\}\\&-\frac{\eta^{A}}{\eta^{N}}\frac{u'\left(c_{t}^{A}\right)}{u'\left(c_{0}^{A}\right)}\left\{\hat{w}_{t}^{A,\star}+2\times\left[\left(\frac{1}{2}-\mu_{t}^{A}\right)\times\left(\left(1-\theta\right)\hat{w}_{t}^{N,\star}-\hat{w}_{t}^{A,\star}\right)\right.\right.\\&\left.\qquad-\left(1-\mu_{t}^{A}-\mu_{t}^{N}\right)\times\left(1-\theta\right)\hat{w}_{t}^{N,\star}\right]\\&\left.\qquad+\Delta_{t}^{\star}-\sum_{h}\mu_{t}^{h}\hat{w}_{t}^{h,\star}\right\}\end{aligned} \tag{A.26}$$

with $\{\mu_t^A,\mu_t^N\}$ given by (3.4)-(3.5) evaluated at $T=T^{SB}$. The sequences $\{\hat{w}_t^h\}$ and $\{\hat{w}_t^{h,\star}\}$ denote the perturbation of equilibrium wages $w_t^h\equiv\partial_h G(\cdot)$ with respect to a change in T and α. The brackets in (A.25)–(A.26) capture the change in labor incomes $\hat{\mathcal{Y}}_t^h$ in (3.6), and profits Π_t (the last terms in the brackets). These pecuniary externalities can be decomposed into aggregate and distributional terms.

[42] The expressions below are obtained by rearranging the optimality conditions from the government's problem (Lemma 4). The derivation of these expressions uses the fact that the stopping time T_0 is chosen optimally, i.e., an envelope condition applies.

[43] The effective Pareto weights are normalized so $\eta^h/u'\left(c_0^h\right)$ are the weights in the government's objective (4.1).

For instance, the pecuniary externality (A.26) can be written as

$$\hat{\Phi}_t^{\star}(\cdot) = \underbrace{-\frac{\eta^A}{\eta^N}\frac{u'\left(c_t^A\right)}{u'\left(c_0^A\right)}\Delta_t^{\star}}_{\equiv \hat{\Phi}_t^{\text{aggreg},\star}(\cdot)} \underbrace{-\sum_h \frac{\eta^h}{\eta^N}\frac{u'\left(c_t^h\right)}{u'\left(c_0^h\right)}\Sigma_t^{h,\star}}_{\equiv \hat{\Phi}_t^{\text{distrib},\star}(\cdot)} \tag{A.27}$$

for some distributional terms $\{\Sigma_t^{h,\star}\}$ that sum up to zero in every period $\sum_h \Sigma_t^{h,\star} = 0$. The term $\hat{\Phi}_t^{\text{aggreg},\star}(\cdot)$ captures the aggregate effect of the intervention $(\Delta_t^{\star})$, while the term $\hat{\Phi}_t^{\text{distrib},\star}(\cdot)$ captures the distributional effect that it has across different workers. We define the same objects for reallocation, and denote them without a $\star$. Note that $\hat{\Phi}_t^{\text{aggreg}}(\cdot) = 0$ for reallocation.

The equilibrium is *constrained efficient* if and only if

$$\Phi\left(\alpha^{\text{LF}}, T^{\text{LF}}; \eta^A, \eta^N\right) = \Phi^{\star}\left(\alpha^{\text{LF}}, T^{\text{LF}}; \eta^A, \eta^N\right) = 0 \tag{A.28}$$

for *some* weights $\{\eta^A, \eta^N\}$. We now show that if these conditions hold, there is a small perturbation of the production function such that (A.28) does not hold with these weights. In particular, consider the perturbed production function

$$\mathcal{G}(G^{\star}, \epsilon) = G^{\star} + \epsilon g\left(\mu^A, \mu^N; \alpha\right) \tag{A.29}$$

where g is any function that satisfies that following conditions when evaluated at the laissez-faire. First,

$$g\left(\mu_t^{A,\text{LF}}, \mu_t^{N,\text{LF}}; \alpha^{\text{LF}}\right) = 0 \tag{A.30}$$

for all $t \geq 0$. Second,

$$\int_{T_0^{\text{LF}}}^{+\infty} \exp(-\rho t)\frac{u'\left(c_t^A\right)}{u'\left(c_0^A\right)}\partial_h g(\cdot)\, dt \quad = 0 \tag{A.31}$$

for each occupation $h \in \{A, N\}$. Finally,

$$\int_0^{+\infty} \exp(-\rho t)\frac{u'\left(c_t^N\right)}{u'\left(c_0^N\right)}\partial_\alpha g(\cdot)\, dt \quad = 0 \tag{A.32}$$

along the initial equilibrium. For instance,

$$g\left(\mu_t^A, \mu_t^N; \alpha\right) \equiv \left\{\mu_t^A + \varrho \mu_t^N\right\}\left(\alpha^{\mathrm{LF}} - \alpha\right) \tag{A.33}$$

satisfies (A.30)–(A.32) when choosing $\varrho < 0$ appropriately.

Then, the allocation $(\mu_t^{A,\mathrm{LF}}, \mu_t^{N,\mathrm{LF}}, \alpha^{\mathrm{LF}})$ still satisfies all equilibrium conditions after an infinitesimal variation $\varepsilon > 0$. That is, the laissez-faire is unchanged. It follows that the pecuniary externality that concerns reallocation (A.25) still nets out too, i.e., $\Phi\left(\alpha^{\mathrm{LF}}, T^{\mathrm{LF}}; \eta^A, \eta^N\right) = 0$. The reason is that this pecuniary externality only involves terms in $D^2 G^\star$ with respect to labor (μ^A, μ^N), while the perturbation (A.33) is linear in (μ^A, μ^N), and cannot affect these terms.

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

$$\omega_t^h \equiv \frac{\exp\left(-\rho t\right)\lambda_t^h}{\int_0^{+\infty} \exp\left(-\rho s\right)\lambda_t^h} \tag{A.34}$$

for each $h = A, N$. Note that the sequences $\{\omega_t^h\}$ integrate to 1, and $\{\omega_t^A - \omega_t^N\}$ decreases over time. The reasons 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). Furthermore, note that $\partial_\alpha g\left(\mu_t^{A,\mathrm{LF}}, \mu_t^{N,\mathrm{LF}}; \alpha\right) = -\left(\mu_t^A + \varrho \mu_t^N\right)$ *increases* over time when evaluated at the laissez-faire, since $\partial_t \mu_t^A < 0$ and $\partial_t \mu_t^N > 0$ in equilibrium. Put it differently, automated workers put a relatively higher weight on earlier (smaller) flows compared to non-automated workers. It follows that

$$\begin{aligned} &\int_0^{+\infty} \omega_t^A \partial_\alpha g\left(\mu_t^{A,\mathrm{LF}}, \mu_t^{N,\mathrm{LF}}; \alpha^{\mathrm{LF}}\right) dt \\ &< \int_0^{+\infty} \omega_t^N \partial_\alpha g\left(\mu_t^{A,\mathrm{LF}}, \mu_t^{N,\mathrm{LF}}; \alpha^{\mathrm{LF}}\right) dt = 0 \end{aligned} \tag{A.35}$$

given (A.32). Thus, we constructed a variation $\mathcal{G}\left(G^\star, \epsilon\right)$ such that the pecuniary externality $\Phi^\star\left(\alpha^{\mathrm{SB}}, T^{\mathrm{SB}}; \eta^A, \eta^N\right) \neq 0$. That is, the second best and the laissez-faire do not coincide after the perturbation $\varepsilon > 0$. Finally, $G^{\star,\prime}\left(G^\star, \epsilon\right) \to G^\star$ uniformly as $\epsilon \to 0$ given (A.33), as claimed.

A.5 Proof of Proposition 1

The equilibrium degree of automation α^{LF} satisfies

$$\int_0^{+\infty} \exp(-\rho t) \times \frac{u'\left(c_t^N\right)}{u'\left(c_0^N\right)} \Delta_t^\star dt = 0, \tag{A.36}$$

where $\Delta_t^\star$ is defined by (3.8) and denotes the response of aggregate output to automation. Let

$$\Psi(\alpha) \equiv \int_0^{+\infty} \exp(-\rho t) \times \sum_h \frac{1}{2} \eta^{h,\text{effic}} \frac{u'\left(c_t^h\right)}{u'\left(c_0^h\right)} \left(\Delta_t^\star + \Sigma_t^{h,\star}\right) dt, \tag{A.37}$$

where the equilibrium objects on the RHS are implicitly indexed by α, $\{\Sigma_t^{h,\star}\}$ capture distributional pecuniary externalities with $\sum_h \frac{1}{2}\Sigma_t^{h,\star} \equiv 0$ for all periods t, and $\left\{\eta^{A,\text{effic}}, \eta^{N,\text{effic}}\right\}$ are the efficiency weights.

The second-best level of automation associated with efficiency weights $\alpha^{\text{SB,effic}}$ satisfies

$$\Psi\left(\alpha^{\text{SB,effic}}\right) = 0 \tag{A.38}$$

By definition, the efficiency weights ensure that the distributional terms net out for the automation choice when evaluated at the laissez-faire. Therefore,

$$\Psi\left(\alpha^{\text{LF}}\right) = \int_0^{+\infty} \exp(-\rho t) \times \sum_h \frac{1}{2} \eta^{h,\text{effic}} \frac{u'\left(c_t^h\right)}{u'\left(c_0^h\right)} \Delta_t^\star dt, \tag{A.39}$$

where the RHS is evaluated at the laissez-faire α^{LF}. Define $\lambda_t^h \equiv u'\left(c_t^h\right)/u'\left(c_0^h\right)$ as in Appendix A.4. Note that

$$\Psi\left(\alpha^{\text{LF}}\right) \propto \frac{1}{2} \int_0^{+\infty} \underbrace{\frac{\exp(-\rho t) \times \sum_h \eta^{h,\text{effic}} \lambda_t^h}{\int \exp(-\rho s) \times \sum_h \eta^{h,\text{effic}} \lambda_s^h ds}}_{\equiv \omega_t^\star} \Delta_t^\star dt, \tag{A.40}$$

up to some positive constant. Similarly,

$$\int_0^{+\infty} \underbrace{\frac{\exp\left(-\rho t\right) \times \lambda_t^N}{\int \exp\left(-\rho s\right) \times \lambda_s^N ds}}_{\equiv \omega_t^N} \Delta_t^\star dt = 0 \tag{A.41}$$

using the the optimality condition (A.36). Furthermore, note that the sequences $\{\omega_t^\star, \omega_t^N\}$ both integrate to 1, and $\{\omega_t^\star - \omega_t^N\}$ decreases over time. The reason is the same as in Appendix A.4. Moreover, note that the sequence $\{\Delta_t^\star\}$ increases over time (Lemma 3). It follows that

$$\Psi\left(\alpha^{\mathrm{LF}}\right) < 0. \tag{A.42}$$

Therefore, a decrease in automation $\delta\alpha < 0$ results in a welfare increase $\delta\mathcal{U}\left(\cdot\right) > 0$ starting from the laissez-faire. Given our assumption that the government's objective $\mathcal{U}$ is concave in α when using the weights $\{\eta^{h,\mathrm{effic}}\}$, this also shows that $\alpha^{\mathrm{LF}} > \alpha^{\mathrm{SB,effic}}$ and the government finds it optimal to *tax* automation.

A.6 Proof of Proposition 2

We proceed as in Appendix A.5. The equilibrium stopping time T^{LF} satisfies

$$\int_{T^{\mathrm{LF}}}^{+\infty} \exp\left(-\rho t\right) \times \frac{u'\left(c_t^A\right)}{u'\left(c_0^A\right)} \Delta_t = 0, \tag{A.43}$$

where Δ_t is defined by (3.3) and denotes the response of aggregate output to labor reallocation. The government's incentives to automate when using efficiency weights is

$$\hat{\Psi}\left(\alpha^{\mathrm{LF}}\right) = \int_0^{+\infty} \exp\left(-\rho t\right) \sum_h \frac{1}{2} \eta^{h,\mathrm{effic}} \frac{u'\left(c_t^h\right)}{u'\left(c_0^h\right)} \times \left\{\Delta_t^\star + \mathbf{1}_{\{t > T(\alpha^{\mathrm{LF}})\}} \frac{1}{2} \lambda \exp\left(-\lambda T\left(\alpha^{\mathrm{LF}}\right)\right) T'\left(\alpha^{\mathrm{LF}}\right) \Delta_t\right\} dt, \tag{A.44}$$

when evaluated at the laissez-faire, where $T\left(\cdot\right) > 0$ is the equilibrium stopping time for a given level of automation.

Let $\{\omega_t^\star, \omega_t^N\}$ be the weights defined in Appendix A.5. We have shown in that appendix that the terms in $\Delta_t^\star$ contribute negatively to (A.44). We now turn our attention to the terms in Δ_t. The sequences $\{\omega_t^\star, \omega_t^N\}$ still both integrate to 1, and $\{\omega_t^\star - \omega_t^N\}$ decreases over time. Moreover, note that $\Delta_0 > 0$ (Assumption 3) and $\Delta_t < 0$ as $t \to +\infty$ (Lemma 1). Then, the sequence $\{\Delta_t\}$ decreases over time, assuming monotonicity for all t. Therefore, we have

$$\int_0^{+\infty} \omega_t^\star \Delta_t dt < \int_0^{+\infty} \omega_t^A \Delta_t d = 0$$

using the optimality condition (A.43). Finally, note that $T'(\cdot) > 0$ as more labor reallocates when automation increases. It follows that the terms in Δ_t also contribute negatively to (A.44). This provides an additional motive to tax automation on efficiency grounds.

A.7 Second best with equity concerns

Proposition 5 (Second best with equity concerns)**.** *Consider the special case of our model with no borrowing frictions — so that the laissez-faire is first best. Suppose that the government is utilitarian. The government should curb automation.*

Proof. Suppose that there are no borrowing frictions ($\underline{a} \to -\infty$) and so the MRS of all workers coincide $\left(\frac{u'(c_t^N)}{u'(c_0^N)} = \frac{u'(c_t^A)}{u'(c_0^A)} = \frac{u'(C_t)}{u'(C_0)}\right)$. The second best level of automation with utilitarian weights $\left(\eta^{h,\text{utilit}} = 1/2\right)$ satisfies the optimality condition

$$\int_0^{+\infty} \exp(-\rho t) \frac{u'(C_t)}{u'(C_0)} \sum_h \frac{1}{2} u'\left(c_0^h\right) \times \left\{\Delta_t^\star + \Sigma_t^{h,\star}\right\} = 0 \tag{A.45}$$

If we evaluate the left-hand side at the laissez-faire allocation, we obtain

$$\tilde{\Psi}\left(\alpha^{\text{LF}}\right) \equiv \int_0^{+\infty} \exp(-\rho t) \frac{u'\left(C_t^{\text{LF}}\right)}{u'\left(C_0^{\text{LF}}\right)} \frac{1}{2} \sum_h u'\left(c_0^h\right) \times \Sigma_t^{h,\star} dt \tag{A.46}$$

using (3.7). Note that $u'\left(c_0^A\right) > u'\left(c_0^N\right)$ since automated workers are worse off. Furthermore, note that

$$\Sigma_t^{A,\star} < 0 \quad \text{and} \quad \Sigma_t^{N,\star} > 0 \tag{A.47}$$

for all t, since these terms capture the distributional effects of automation in equilibrium. Workers initially employed in these occupations are worse off. Therefore,

$$\tilde{\Psi}\left(\alpha^{\text{LF}}\right) < 0 \tag{A.48}$$

It follows that the laissez-faire degree of automation is excessive compared to the second best $\left(\alpha^{\text{LF}} > \alpha^{\text{SB,utilit}}\right)$ associated with utilitarian weights and the government finds it optimal to *tax* automation. □

A.8 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

$$Y_t = G\left(\mu_t^A, \mu_t^N; \alpha_t\right) - x_t - \Omega\left(x_t; \alpha_t\right), \tag{A.49}$$

where x_t is the gross investment rate in automation and $\Omega(\cdot)$ is a convex function with $\Omega(\delta\alpha; \alpha) = 0$. 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.

Laissez-faire. We guess (and verify) that the economy converges 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 $t \to +\infty$, and these levels satisfy

$$\frac{1}{\rho+\delta}\left(G_\alpha\left(\mu^{A,\text{LF}}, \mu^{N,\text{LF}}; \alpha^{\text{LF}}\right) - \Omega_\alpha\left(\delta\alpha^{\text{LF}}; \alpha^{\text{LF}}\right)\right) = \left(1 + \Omega_x\left(\delta\alpha^{\text{LF}}; \alpha^{\text{LF}}\right)\right), \tag{A.50}$$

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

$$w^{A,\text{LF}} = G_1\left(\mu^{\text{LF}}, 1-\mu^{\text{LF}}; \alpha^{\text{LF}}\right) = G_2\left(\mu^{\text{LF}}, 1-\mu^{\text{LF}}; \alpha^{\text{LF}}\right) = w^{N,\text{LF}}, \tag{A.51}$$

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.50)-(A.51) pin down the long-run labor allocation $\{\mu^{A,\mathrm{LF}}, \mu^{N,\mathrm{LF}}\} = \{\mu^{\mathrm{LF}}, 1-\mu^{\mathrm{LF}}\}$, automation α^{LF}, and aggregate consumption

$$C^{\mathrm{LF}} = G\left(\mu^{\mathrm{LF}}, 1-\mu^{\mathrm{LF}}; \alpha^{\mathrm{LF}}\right) - \delta\alpha^{\mathrm{LF}}. \tag{A.52}$$

Finally, all workers are hand-to-mouth in the long-run since the borrowing limit is $\underline{a} \to 0$. Therefore, $c_s^{h,\mathrm{LF}} \to C^{\mathrm{LF}}$ as $t \to +\infty$ for all generations s and each $h = A, N$. Therefore, $u'\left(c_{s,t+\tau}^{h,\mathrm{LF}}\right) / u'\left(c_{s,t}^{h,\mathrm{LF}}\right) \to 1$ 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. Production efficiency requires that the marginal products of labor must be equalized in a long-run in any first best allocation, so equation (A.51) holds. Moreover, the planner values the gains from automation over time with the discount rate ρ, so (A.50) holds too in the long-run. These two restrictions define the same long-run allocation for labor $\{\mu^{A,\mathrm{LF}}, \mu^{N,\mathrm{LF}}\} = \{\mu^{\mathrm{LF}}, 1-\mu^{\mathrm{LF}}\}$, automation α^{LF}, and aggregate consumption (A.52) as the one that the laissez-faire converges to. 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

$$\frac{\eta_s^h \exp\left(-\left(\rho+\chi\right)\left(t-s\right)\right) u'\left(c_{s,t}^{h,\mathrm{FB}}\right)}{\eta_\tau^j \exp\left(-\left(\rho+\chi\right)\left(t-\tau\right)\right) u'\left(c_{\tau,t}^{j,\mathrm{FB}}\right)} = 1 \tag{A.53}$$

for generations $s, \tau \leq t$ and each occupations $h, j = A, N$. Thus, consumption is equalized across workers when using the weights

$$\eta_s^h = \exp\left(-\left(\rho+\chi\right)s\right), \tag{A.54}$$

Therefore, $c_{s,t}^{h,\mathrm{FB}} \to C^{\mathrm{FB}} = G\left(\mu^{\mathrm{FB}}, 1-\mu^{\mathrm{FB}}; \alpha^{\mathrm{FB}}\right) - \delta\alpha^{\mathrm{FB}}$ for all s,t and each h. This implies that the laissez-faire allocation and the first best allocation with weights (A.54) coincide asymptotically.

A.9 An Alternative Definition of Efficiency Improvements

Following Costinot and Werning (2022), suppose that the government has access to limited redistributive instruments that allow it to undo the distributional effects of its policy interventions. In this case, the government adjusts these instruments to offset the distributional effects $\Sigma_t^{h,(\star)} = 0$ for each $h = A, N$ as defined in Appendix A.4. Therefore, the distributional component of the pecuniary externalities is $\hat{\Phi}_t^{\text{distrib},(\star)}(\cdot) = 0$. The government's optimality conditions become

$$\int_0^{+\infty} \exp(-\rho t) \sum_h \frac{1}{2} \eta^h u'\left(c_t^N\right) \Delta_t^\star dt = 0 \tag{A.55}$$

$$\int_{T^{\text{SB}}}^{+\infty} \exp(-\rho t) \sum_h \frac{1}{2} \eta^h u'\left(c_t^N\right) \Delta_t dt = 0 \tag{A.56}$$

at the second best. This holds for *any* sets of Pareto weights. It follows that the optimality condition (A.55) cannot hold at the laissez-faire and that the government finds it optimal to curb automation. This part of the proof is identical to the last part of Appendix A.5.

A.10 Task-Based Example

Using our task-based 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.

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

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

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

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

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

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

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

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.[44] The workers' problem can be formulated recursively

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

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

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

[44] 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.

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[45]

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

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

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

In turn, the continuation value for unemployed workers $(e = U)$ is

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

where $\mathcal{S}(\cdot)$ is the mobility hazard, and $\xi'(\cdot) = (1-\theta)\xi$ when the reallocation spell is complete. New generations 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 chooses to enter her new occupation. Finally, the mean-reverting component income (z) evolves as

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

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

B.2 Firms' Problem

We solve the mutual fund's and the firm's problem in continuous time. The mutual fund invests in automation subject to convex adjustment costs and rents its stock to the firm. The mutual fund's problem can be formulated recursively

$$r_t W_t(\alpha) = \max_{\{x,\alpha'\}} r_t^\star \alpha - (1+\tau_t^x)\, x - \omega \left(\frac{x}{\alpha} - \delta\right)^2 \alpha + (x - \delta\alpha)\, W_t'(\alpha) + \frac{\partial}{\partial t} W_t(\alpha) \tag{B.8}$$

$$\text{s.t. } x \geq 0$$

where α is the stock of automation, x is gross investment, i.e. $d\alpha_t = (x_t - \delta\alpha_t)\, dt$, $r_t^\star$ is the rental rate of automation, and τ_t^x is a potential distorsionary tax on investment. The optimal supply of automation satisfies

$$(r_t + \delta)\left((1+\tau_t^x) + 2\omega\left(x_t^\star - \delta\right)\right) = \left\{ r_t^\star + \omega\left[\left(x_t^\star\right)^2 - \delta^2\right]\right\} + \partial_t \tau_t^x + 2\omega \partial_t x_t^\star, \tag{B.9}$$

with $x_t^\star \equiv x_t/\alpha_t$, together with the law of motion

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

the initial α_0 and a standard transversality condition. The firm's problem is

$$\max_{\{\alpha_t^h, \mu_t^h\}} G^\star\left(\left\{\alpha_t^h, \mu_t^h\right\}\right) - \phi^A r_t^\star \alpha_t^A - \sum_h \phi^h w_t^h \mu_t^h \quad \text{s.t.} \quad \alpha_t^N = 0$$

where α_t^h and μ_t^h denote automation and labor rented by the firm in $h = A, N$, and

$$G\left(\left\{\alpha_t^h, \mu_t^h\right\}\right) = \left(\sum_h \phi^h \left\{A^h\left(\varphi\alpha_t^h + \mu_t^h\right)^{1-\eta}\right\}^{\frac{\nu-1}{\nu}}\right)^{\frac{\nu}{\nu-1}}$$

is the aggregate production function. The rental rate is $r_t^\star \equiv \varphi w_t^A$ and wages are

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

Finally, market clearing for inputs requires that the firm rents the stock of automation supplied by the mutual fund

$$\alpha_t^A = \alpha_t/\phi \quad \text{and} \quad \tilde{\alpha}_t^N = 0,$$

and that the firm hires the (effective) labor supplied in each occupation[46]

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

B.3 Second Best

In this appendix, we state the second best problem we consider in our numerical exercise and discuss our choice of Pareto weights.

Objective. The government's objective is

$$\begin{aligned} \mathcal{W} \equiv & \chi \int_{-\infty}^{0} \int \eta_s (\mathbf{x}) \exp\left(\left(\rho+\chi\right) s\right) V_0^{\text{old}} (\mathbf{x})\, \pi_{s,0}^{\text{old}} (d\mathbf{x})\, ds \\ & + \chi \int_0^{+\infty} \eta_s V_s^{\text{new}} ds, \end{aligned} \tag{B.11}$$

for some Pareto weights $\boldsymbol{\eta}$. The first and second terms capture the contributions of existing $(s<0)$ and new generations $(s \geq 0)$, respectively. Following Calvo and Obstfeld (1988), these (continuation) values are evaluated at birth.[47] The value $\exp\left(\left(\rho+\chi\right)s\right) V_0^{\text{old}}$ is the continuation utility of existing generations over periods $t \geq 0$. The measure $\pi_{s,0}^{\text{old}}$ is the distribution of idiosyncratic states in period $t=0$ for existing generations born in $s<0$ (conditional on survival). In turn, the value

$$V_t^{\text{new}} \equiv \int \gamma \log\left(\sum_h \phi^h \exp\left(\frac{V_t^h(0,1,0,z)}{\gamma}\right)\right) P^\star(dz) \tag{B.12}$$

is the continuation utility for new generations born in period $t=s \geq 0$, which reflects their occupational choice.[48] Here, $P^\star$ denotes the ergodic distribution of

[46] Labor supply in *each* occupation is the *total* mass of workers employed in occupations of type h, i.e., $\int \mathbf{1}_{\{e=1,h'=h\}} \xi d\pi_t$, divided by the mass of such occupations ϕ^h.

[47] This explains the presence of the additional discounting $\exp\left(\left(\rho+\chi\right)s\right)$ for existing generation $s<0$.

[48] Members of a new generation are born with no assets $a=0$, are employed $e=1$, and have not

the income process $z'|z \sim P(z)$, i.e., the distribution of productivities at birth.

Pareto weights. We choose two sets of weights: weights that capture the *efficiency* motive for policy intervention, and utilitary weights. We now describe these efficiency weights. Our approach is similar to the one we adopted in our tractable model (Section 4.3). The weights that the government puts on a given worker are inversely related to this worker's marginal utility at birth (evaluated at the laissez-faire transition). This ensures that the government has no incentive to redistribute resources (at birth) to improve equity. In particular, the government weights constrained workers (with a higher marginal utilitary) *less* compared to a utilitarian government. We also assume the the government discounts generations at rate ρ over time, which ensures that the planner does not discriminate across generations at a first best — see equation (A.53). Therefore, the weights assigned to old generations satisfy

$$\eta_s(\mathbf{x}) = \exp(-\rho s) \times 1/\partial_a V_0^{\text{old,LF}}(\mathbf{x}), \tag{B.13}$$

where $1/\partial_a V_0^{\text{old,LF}}(\mathbf{x})$ is the marginal utility of financial wealth at the laissez-faire. In turn, the weights assigned to new generations satisfy

$$\exp(-\rho s)/\eta_s(z) = \sum_h \mathcal{S}_s^h(0,1,0,z)\,\partial_a V_t^{h,\text{LF}}(a,1,0,z)\Big|_{a=0} \tag{B.14}$$

for all $s \geq 0$ since new generations start with no financial assets $a = 0$ and have not reallocated yet $\xi = 0$.

Summarizing, the government's objective becomes

$$\begin{aligned}\mathcal{W} \equiv & \int \frac{V_0(\mathbf{x})}{\partial_a V_0^{\text{old,LF}}(\mathbf{x})} \pi_0(d\mathbf{x})\, ds \\ & + \chi \int_0^{+\infty} \exp(-\rho s) \frac{V_s^{\text{new}}}{\int \sum_h \mathcal{S}_s^h(0,1,0,z')\,\partial_a V_s^{h,\text{LF}}(a,1,0,z')\Big|_{a=0} P^\star(dz')} ds,\end{aligned} \tag{B.15}$$

incurred the productivity cost associated to switching occupations $\xi = 0$.

where

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

is the unconditional (initial) distribution of idiosyncratic states. When solving for the constrained efficient steady state, we maximize the contribution of generations $s \to +\infty$ to the objective (B.15), i.e., $\lim_{s\to+\infty} V_s^{\text{new}}$.

Policy tools and implementability. The government maximizes the objective (B.11) by choosing an appropriate sequence of distortionary taxes on investment $\{\tau_t^x\}$ and rebating the proceedings back to the mutual fund or the workers. The implementability constraints consist of workers' reallocation and consumption choices.

C Numerical Implementation

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

Workers' problem. We solve the problem worker's (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. The firm's optimal choice of investment and automation is characterized by the non-linear system of differential equations (B.9)–(B.10). We solve this system using a standard shooting algorithm. Fixing an initial value for investment x_0, we iterate the system forward. We then adjust this initial value until automation converges to its long-run level.

Policy. For numerical reasons, we restrict our attention to simple perturbations of $\{\alpha_t\}$ from the sequence that prevails at the laissez-faire. We do so by repeatedly feeding sequences of taxes $\{\tau_t^x\}$ in the mutual fund's problem (B.8).[49] These taxes

$$\tau_t^x = \exp\left(-\beta t\right)\hat{\tau} + \bar{\tau} \tag{C.1}$$

consist of a persistent component $\hat{\tau}$ and a permanent one $\bar{\tau}$. The taxes converge to monotonically to their permanent level. The persistent component allows to *slow down* automation early on during the transition. In turn, the permanent component controls the long-run level of automation. It is well-known that a long-run tax (or subsidy) on *capital* can be optimal when markets are incomplete — it can improve insurance and / or prevent dynamic inefficiency (Section 5.4). We choose a subsidy $\bar{\tau} = -39.9\%$ so that the economy converges to its constrained efficient steady state. We set the mean-reversion speed β so that the half-life of τ_t^x is the same as the one of automation at the laissez-faire (20 years). Finally, we optimize over $\hat{\tau}$ on a fine grid to find the second best intervention. The Pareto weights (Section B.3) are evaluated at the allocation with the permanent subsidy $\bar{\tau}$ (but no persistent tax $\hat{\tau}$). This ensures that $\bar{\tau}$ is the optimal long-run policy.

D Additional Numerical Results

We present additional numerical results that were omitted from the main text.

D.1 Employment and Value Added

We argued that our model matches well the output produced by automation in McKinsey (2017), as well as the firm-level effects of automation on employment and value added per worker in 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 results from 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

[49] The differential equation (B.9) can become stiff when prices are sufficiently persistent. We thus evaluate prices at the *laissez-faire* to avoid stability issues. Re-optimizing for a given sequence of taxes $\{\tau_t^x\}$ yields a new sequence $\{\alpha_t\}$ which was feasible in the original government's problem.

model, the output in occupation $h = A$ that is produced by automation is $\varphi\alpha / \left(\varphi\alpha + \mu^A\right)$, which is 72% 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 is

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

where w is wage and p is the price of the intermediate good. Next, consider the following partial equilibrium exercise. 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

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

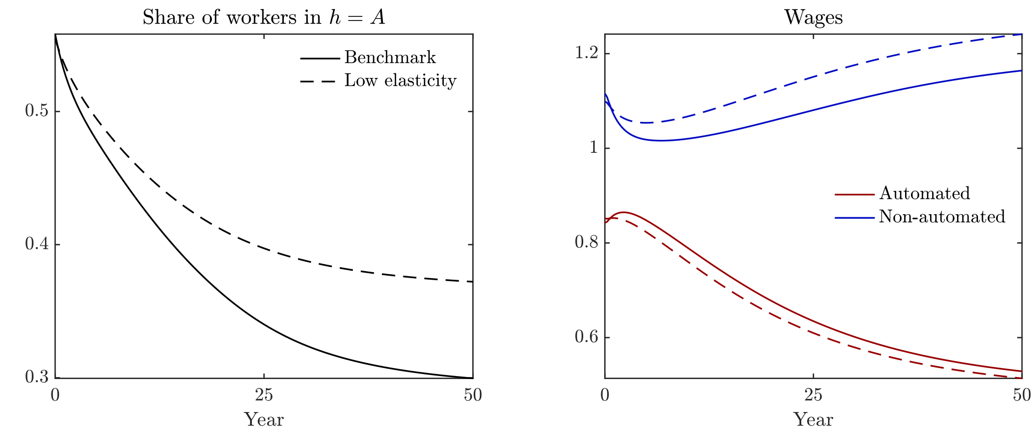
So, the percent difference in employment is

$$\frac{\mu_1 - \mu_0}{\mu_0} = -\frac{\varphi\alpha_1}{\mu_0} = -62\% \tag{D.3}$$

using our calibration, when μ_0 is the initial steady state employment in automated occupations and α_1 equals the stock of automation 10 years out in the transition, which is roughly half the sample period in Table 2 of Bonfiglioli et al. (2022). Alternatively, we could also have considered a general equilibrium exercise where we compare the decline in employment in the firm across steady states. This gives $\frac{\mu_1-\mu_0}{\mu_0} = -48\%$ instead. To sum up, our quantitative model gives declines in employment from automation that are roughly between 50% and 60%, which are comparable in magnitude to the 54% estimated by Bonfiglioli et al. (2022).

Productivity. The percent change in value added per worker of a firm (i.e., labor productivity) that adopted automation, relative to a firm that did not, can be computed by the ratio of the coefficients in column (3) to column (5) in the first line of Table 2 of Bonfiglioli et al. (2022). This gives -0.302/0.174 $= 175\%$. In our model,

Figure D.1: Allocations



Notes: Solid curves correspond to the baseline calibration and dashed curves to the alternative calibration with a lower elasticity of labor supply (1). Wages are normalized by their initial steady state levels.

the percent difference in labor productivity across the firms is

$$\frac{(\varphi\alpha_1 + \mu_1)^{1-\eta} / \mu_1}{(\mu_0)^{1-\eta} / \mu_0} - 1 = \frac{\mu_0 - \mu_1}{\mu_1} = 163\% \tag{D.4}$$

using our calibration, where μ_1 and μ_0 are as above and the first equality in the previous expression uses (D.2). Alternatively, when comparing across steady states, we obtain a percent difference in labor productivity of 184%. To sum up, our quantitative model gives labor productivity increases from automation that are comparable in magnitude to the 175% estimated by Bonfiglioli et al. (2022).

D.2 Elasticity of Labor Supply

In Section 6.3, we discussed an alternative calibration with a Fréchet parameter chosen to match an elasticity of labor supply of 1 instead of 2. Figure D.1 illustrates the transition dynamics in this case. Labor reallocation is slower since labor supply is less responsive to wage changes, and fewer workers reallocate overall. In turn the wage gap widens across occupations.

D.3 Alternative Calibrations

We explore the role of various parameters, in addition to those discussed in Section 6.3. Table D.1 reports the welfare gains for these alternative calibrations.

The first parameter of interest is the elasticity of substitution between occupations. We choose the value $\nu = 0.75$ in our benchmark calibration (Section 6.1). This value is almost identical to the elasticity across tasks estimated by Gregory et al. (2021). It is slightly lower than the estimate of 0.9 in Goos et al. (2014). For this reason, we recalibrate our model with $\nu = 0.9$. The welfare gains (first column in the table) are slightly higher than in our benchmark calibration.

Second, we decrease the productivity of automation φ by roughly 10% compared to our benchmark. The welfare gains are slightly lower in this case. Third, we double the adjustment cost ω to 8 from 4 in our benchmark. This decreases the speed of automation at the laissez-faire, which reduces the welfare gains from the policy intervention. Fourth, we increase the ratio of liquidity to GDP $(-B/Y)$ from our benchmark 0.75 to 1.4 (as in McKay et al., 2016). This level of liquidity is several times larger than effective liquid asset holdings by the average US household (Kaplan et al., 2018), which substantially alleviates borrowing constraints. We find much smaller welfare gains, especially under efficiency weights as anticipated in Section 3.4. Finally, we decrease the initial stock of automation to 1/20 of its final steady state from 1/10 in our benchmark. This decreases the welfare gains.

Table D.1: Welfare Gains $\Delta\mathcal{W}$ from Second Best Interventions

	Alternative calibrations				
	High ν	Low φ	High ω	High $-B/Y$	Low α_0
Efficiency	4.0%	3.6%	2.4%	0.6%	3.1%
Utilititarian	6.3%	5.6%	3.7%	2.3%	4.7%

Note: 'High ν' and 'Low φ' denote calibrations with $\nu = 0.9$ and $\varphi = 0.38$, respectively. 'High ω' and 'High $-B/Y$' denote calibrations with $\omega = 8$ and $-B/Y = 1.4$, respectively. 'Low α_0' uses an initial condition α_0 that is 1/20 of its final steady state level. 'Efficiency' and 'Utilitarian' compute the gains and optimal automation taxes using the two Pareto weights (Appendix B.3).