Artificial Intelligence, Worker-Replacing Technological Progress and Income Distribution

Anton Korinek
Johns Hopkins and NBER

Joseph E. Stiglitz
Columbia and NBER

September 2017

(preliminary and incomplete – please do not circulate)

Abstract

Technologists predict that in coming decades, artificially intelligent machines will increasingly replace human labor and, once they surpass human levels of general intelligence, will make human labor redundant. We study the implications of this transition for wages and economic growth, and we examine what policy measures would ensure that technological progress generates a Pareto improvement. We find several results. First, once machines surpass humans, the economy will experience a singularity in which machines producing more machines generate an explosion in labor supply and lead to exponential economic growth. Secondly, however, growth will eventually be limited by irreproducible scarce factors, e.g. land and energy, which causes real wages to plummet, leading to increasing inequality among workers and other factor owners. Third, by imposing non-distortionary taxes on these irreproducible factors, we can generate a Pareto improvement, i.e. ensure everybody in the economy is better off. If such redistribution is not possible, then interventions to direct technological progress may act as a second-best device to limit the losses of workers.

1 Introduction

Artificial intelligence has advanced rapidly in recent years and is starting to make an increasing number of human professions redundant. Over the next decade or two, Frey and Osborne (2013) observe that 47% of US jobs could be automated. Some technologists predict that artificial intelligence will reach and then surpass human levels of general intelligence within the next several decades (see e.g. Kurzweil, 2005). If this milestone is passed, there will be no human tasks left that cannot be better performed by machines. This includes tasks that require what is frequently called emotional intelligence (i.e. the ability to read, process and appropriately respond to non-verbal human communication) and creative tasks (as academics, we wish to think that these will be the last domain for AI to conquer).

The objective of this paper is to study the implications of artificially intelligent machines that replace human labor for inequality. Starting from a benchmark model of economic growth, we introduce a technology to create machine labor that is a perfect substitute for human labor. In the first part of the paper, we study the implications of this new technology for economic growth, human wages, and other factor prices. In the second part of the paper, we examine what policy measures would ensure that technological progress generates a Pareto improvement.

Our first result is to show that once the cost of machine labor falls below a critical threshold, the economy will experience a singularity that leads to a sudden take-off in economic growth. In modern economies, labor is the most important factor of production and, as such, the limited availability of labor represents the main constraint on economic output. Sufficiently cheap machine labor lifts this constraint and ushers in an era of exponential growth since the two main factors of production – labor and capital – can both be accumulated without bounds. The resulting growth is driven purely by factor accumulation and occurs in addition to any growth effects from technological progress. Intuitively, once humans have created some machines, the increasing supply of machine labor allows them to create more and more machines. If complementary factors such as capital keep pace with the supply of labor, overall wages remain unchanged, although the share of human labor in the economy converges to zero. However, if the capital stock adjusts more slowly than the stock of machine labor grows, then the capital-labor ratio in the economy declines and wages fall.

Secondly, we show that economic growth may eventually be limited by the scarcity of irreproducible factors such as land and energy. As machine labor and traditional capital grow exponentially, it is likely that other factors of production that previously did not seem to matter much will become limiting factors once again, and economic growth starts to level off. Given that machine labor will be the largest factor of production in the economy, we expect factors complementary to the machines to be the main limiting factors, for example energy (used to run machine labor) and land (used both to house machines and/or to harvest solar energy). Given large growth in the supply of labor, the amount of each of the limiting factors per unit of labor will decline sharply, causing real wages to plummet and leading to increasing inequality between workers and the owners of the limiting factors. If workers require a minimum subsistence wage to be productive, then machine investment will first lead to falling wages until the subsistence level is reached, then declining employment until all human labor is phased out, and then further declining wages for machine labor.

Thirdly, we turn our attention to how economic policy can mitigate the adverse effects of machine labor on humans. In principle – if lump-sum transfers are available – it is always possible to redistribute the gains from machine technology so that all agents in the economy are better off, i.e. we achieve a Pareto improvement. Even if pure transfers are unavailable, it turns out that taxing the gains experienced by the owners of the fixed factors in the economy (e.g. land or energy) generates sufficient revenue to implement Pareto improvements, and since these factors are fixed, levying taxes on them is non-distortionary. Lastly, if such redistribution is not possible, interventions to slow down technological progress may act as a second-best device to limit the losses of workers. In other words, policymakers need to find the optimal trade-off between equity and efficiency.

2 Machine Labor and the Singularity

We start by extending two simple economic benchmark models with machine labor in order to illustrate that advances in machine technology may eventually lead to a singularity at which the growth behavior of the economy abruptly changes: pervasive robotization will make the most important factor of production in the economy, labor, easily reproducible. In traditional models of economic growth, labor is limited by the size of the population, which is taken as exogenous.[1] When the cost of replacing human labor by machine labor falls below a critical threshold, labor supply and economic output grow quickly and without bounds.

2.1 A Labor-Only Economy

For simplicity, our first model considers a closed economy in which there is a representative agent with preferences[2]

$$U = \sum \beta^t u\left(c_t\right) \qquad \text{where} \qquad u\left(c_t\right) = \left(c_t\right)^{1-\theta} / \left(1-\theta\right)$$

In the first version of our model, we assume that there is a single factor of production, labor. The representative agent is endowed with one unit of human labor $h_t \equiv 1$ each period, which she supplies inelastically. Furthermore, assume that there exists a technology in the economy to invest γm_t units of output at time $t-1$ in order to create m_t units of machine labor at time t. Both human

[1] Even if we accounted for growth in the human population, as is for example the practice in Malthusian growth models, the size of the human population is a very slow-moving variable – "producing" a human who is able to participate in the workforce takes an order of magnitude longer than producing a machine.

[2] Although time-separable preferences with constant elasticity of intertemporal substitution are widespread in macroeconomics, this specification imposes important restrictions on preferences that drive the long-run behavior of the economy. We will discuss this futher in the context of some of our results below.

labor and machine labor are used as perfect substitutes to produce output y_t using a constant-returns production function $AF(h_t + m_t) = A(h_t + m_t)$.

The optimization problem of the representative agent is

$$\max_{\{c_t, m_t \geq 0\}} U \qquad \text{s.t.} \qquad c_t + \gamma m_{t+1} = AF(h_t + m_t)$$

Assigning the shadow price μ_t to the non-negativity constraint of investment in machines, the optimal investment in machines is described by the optimality condition

$$\beta u'(c_{t+1}) AF_m(\cdot) + \mu_t = \gamma u'(c_t) \tag{1}$$

It is easy to see that there are two regimes in the economy, depending on the exogenous parameter values:

Proposition 1 (Economic Singularity)**.** *Consider an economy with $m_t = 0$ which experiences a one-time decline in the cost of investing in machine labor, γ, at time t. There is a critical threshold, $\gamma^* := \beta A$, such that:*
(i) as long as $\gamma \geq \gamma^*$, the economy remains in a steady state with no machines, fixed output and consumption $y_t = A$, constant human labor share $s_h = 1$, and a shadow interest rate $R = 1/\beta$, but
(ii) if $\gamma < \gamma^*$, the economy starts to invest in machine labor and experiences a temporary drop in consumption, $c_t < c_{t-1}$; it then experiences exponential long-run growth at rate $1 + g = (\beta A/\gamma)^{\frac{1}{\theta}}$, with machine labor, output and consumption m_t, y_t, $c_t, \to \infty$, shadow interest rate of $R = A/\gamma$, and human labor share converging to zero, $s_h \to 0$.

Proof. If $\gamma > \gamma^*$, it is easy to verify that the optimality condition (1) of the representative agent is satisfied for constant consumption $c_t = A$ and $\mu_t > 0$ so no machine investment takes place, $m_{t+1} = 0$. The cost of machine labor exceeds the discounted future benefit of machine investment so any $m_{t+1} > 0$ would be suboptimal. This proves part (i) of the Proposition.

If $\gamma < \gamma^*$, the steady state described in (i) violates the agent's optimality condition. The only steady state of the economy exhibits positive machine investment $m_{t+1} > 0$. Rearranging the optimality condition (1), we can see that the economy's steady state growth rate is given by $1+g = \frac{c_{t+1}}{c_t} = \left(\frac{\beta A}{\gamma}\right)^{\frac{1}{\theta}}$. For a given supply of machines m_t, we substitute $c_t = A\left(h_t + m_t\right)$ and $h_t = 1$ to obtain the difference equation

$$m_{t+1} = (1+g)(1+m_t) - 1$$

Starting from $m_t = 0$, machine investment at time t is given by $m_{t+s} = (1+g)^s - 1$ and output is $y_{t+s} = (1+g)^s$. This implies a drop in consumption at time t since $c_{t-1} = y_{t-1} = A$ but $c_t = y_t - \gamma m_{t+1} < A$. The labor share is given by $s_{h,t+s} = 1/(1+g)^s$ which converges to zero for $s \to \infty$. $\square$

The described threshold captures an element of what technologists call a "singularity" — the economy's behavior follows fundamentally different rules once the threshold is crossed.[3] Machine labor is not used as long as it is more expensive than human labor so the economy relies on human labor only, and output is pinned down by the productivity of labor, A. Once machine labor is cheaper than human labor, investment in machines and economic growth take off. Output is determined by the supply of machine labor in the economy, and the share of human labor converges towards zero.

Figure TK below provides an illustration in which we assume that an innovation occurs at time t^* that pushes the cost of machine investment below the threshold γ^*. Before the innovation, the economy is stagnant; after the innovation, exponential economic growth occurs.

[3] One difference is that the singularity described by technologists also involves dramatic progress in technology itself (Kurzweil, 2005). Our example described here only involves rapid proliferation of machines.

Graphic to be inserted

If the economy already is in the machine labor regime and the cost of machine investment declines further, say from γ to γ' at time t', then the growth rate of the economy speeds up from g to $g' = (\beta A/\gamma')^{\frac{1}{\theta}}$. Machine investment goes up, which requires a dip in consumption at time t' compared to its previous trend. Output and consumption grow at the faster pace g' subsequently.

2.2 Machine Labor and Other Factors of Production

This section introduces additional factors of production in our analysis to examine the effects of robotization. The model of the previous section illustrated the potential for a technological singularity in the simplest possible setting. We will see that when generalize our production technology to employ traditional capital, our results on the technological singularity continue to hold, and we can develop a number of further insights.

Let us assume a homothetic constant-returns-to-scale production technology that satisfies the Inada conditions

$$F\left(h_t + m_t, k_t\right)$$

Our analysis refers to k_t as capital, but more broadly, our results apply equally to any other reproducible factor of production. For example, k_t could stand for skilled labor since skill can be accumulated.

After including capital accumulation, the budget constraint of the representative agent is

$$c_t = AF(h_t + m_t, k_t) - \gamma m_{t+1} - k_{t+1}$$

For simplicity, we assume that capital fully depreciates. In addition to condition 1, the representative agent now faces the standard optimality condition for capital investment,

$$\frac{\beta u'(c_{t+1})}{u'(c_t)} \cdot AF_k(\cdot) = 1$$

Proposition 2 (Singularity with Reproducible Production Factors)**.** *Consider an economy in a no-machine steady-state with $m_t = 0$ which experiences a one-time decline in the cost of investing in machine labor, γ, at time t. There is a critical threshold for the cost of machines, γ^*, such that:*
(i) as long as $\gamma \geq \gamma^*$, the economy remains in the no-machine steady state, where $\gamma^* = \beta AF_m\left(1, \bar{k}\right)$ and steady-state capital $\bar{k}$ is pinned down by $\beta AF_k\left(1, \bar{k}\right) = 1$, output is $y_t = AF\left(1, \bar{k}\right)$, the shadow interest rate is $R = 1/\beta$, and the human labor share is strictly interior, $s_h \in (0, 1)$, but
(ii) if $\gamma < \gamma^*$, the economy invests in machine labor, experiences exponential long-run growth with $m_t \to \infty$, $k_t \to \infty$ and $y_t \to \infty$, and the human labor share $s_h \to 0$. If $F(\cdot)$ is homothetic, the capital share and machine labor share converge to constant values that satisfy $s_k + s_m = 1$, and the economy's interest rate converges to a constant $\bar{R} = AF_k/\beta = AF_m/\beta\gamma > 1/\beta$.

Proof. For part (i), if $\gamma > \gamma^*$, it is easy to verify that the optimality condition (1) of the representative agent is satisfied for constant consumption $c_t = A$ and $\mu_t > 0$ so no machine investment takes place. The remaining results follow directly from the optimality condition for capital investment.

For part (ii), observe that once $\gamma < \gamma^*$, $\mu_t = 0$ so machine investment starts taking place. $\square$

Rigidities in Factor Adjustment and Factor Compensation Let us next focus on the transitional dynamics of factor compensation. Up until now, we have assumed that the stock of capital was easily adjustable from one time period to the next. This implied that the representative agent invested into capital up to the point where the marginal return of capital equaled the inverse of the economy's interest rate, $AF_k(\cdot) = 1/R$, where $R = \beta u'(c_{t+1})/u'(c_t)$.

Let us now consider the opposite situation. For simplicity, we assume that capital investment needs to be set one period in advance so that the capital stock is pre-determined when machine labor is first introduced. (Similar results – but further drawn out over over time – can be achieved by introducing e.g. quadratic adjustment costs for capital investment.)

Proposition 3 (Short-Run Effects of Rigidities in Factor Adjustment)**.** *Consider an economy with $m_t = 0$ in which the cost of machines drops below the critical threshold γ^* but capital investment $k_{t+1} = \bar{k}$ is pre-determined for one period.*
(i) At time $t+1$, the economy's interest rate immediately rises. a) If capital and labor are Hicks complements, then the return to capital rises and the return to labor drops. b) If they are Hicks substitutes, the opposite conclusions apply.
(ii) At time $t+2$, all factors have adjusted and the economy's interest rate as well as the returns to capital and labor rise further to their steady state values.

Proof. See appendix. □

Our results so far have painted the picture of an economy in which the abundant availability of machine labor lifts the main constraint on production – the limited supply of human labor, the most important factor of production in our economies. Being freed of any constraints generated by factors in limited supply, this ushers in an era of unlimited growth fueled by the parallel accumulation of machines and traditional capital, even in the absence of traditional technological progress. Although Proposition 3 highlighted that there may be some adverse effects of machine labor on other factors that are substitutes, these adverse effects are purely temporary and only last as long as factor supplies have not fully adjusted (although "temporary" may be a rather long time[4]).

As we mentioned earlier, these results apply to any constant-returns economy in which all factors of production are reproducible. Other than traditional capital, examples for such factors include:

Skilled Labor is commonly considered as a complementary factor to raw human labor in models of technological progress. Since skill can be acquired over time, it is reproducible. Our results in (i.a) suggest that the returns on skilled labor will temporarily benefit from improvements in machine technology.

Entrepreneurial Capital captures the competences and relationships of entrepreneurs to organize factors of production so that output can be generated. It is typically considered complementary to other factors, and it is reproduced as firms grow and new firms are founded. Our results in (i.a) suggest that entrepreneurial capital will also temporarily benefit from improvements in machine technology.

3 Machine Labor and the Return of Scarcity

Our formal analysis so far followed the example of modern macroeconomics to analyze labor and capital as the only two factors of production. Other factors, for example land in a post-agricultural world, do not seem to impose binding constraints at the aggregate level. However, as machine labor and capital grow without bounds, it is likely that other factors of production that previously did not seem to matter much will become limiting factors once again. Important candidates for such limiting factors are energy (used to run machine labor) and land (used both to house machines and/or to harvest solar energy).

We extend our earlier analysis to include (for tangibility) land as a factor of production that is naturally in limited supply. We assume that land is owned

[4] For example, Berg et al. (2017) find that it may take decades for the economy's complementary capital stock to adjust to major revolutions in robot technology.

by the representative agent and denote the agent's holdings by ℓ_t which we normalize to $\bar{\ell} = 1$. We assume that the economy's production function is given by the homothetic CRS technology

$$AF\left(h_t + m_t, k_t, \ell_t\right)$$

We extend the representative agent's budget constraint to

$$c_t = AF\left(h_t + m_t, k_t, \ell_t\right) - \gamma m_{t+1} - k_{t+1} - q_t\left(\ell_{t+1} - \ell_t\right)$$

where q_t is the price of land. The agent's optimality condition for land holdings together with market clearing imply

$$q_t = AF_\ell\left(\cdot\right)/R_t$$

Proposition 4 (Scarcity of Irreproducible Production Factors)**.** *Consider an economy in a no-machine steady-state with $m_t = 0$ which experiences a one-time decline in the cost of investing in machine labor, γ, at time t. There is a critical threshold for the cost of machines, γ^*, such that:*
(i) as long as $\gamma \geq \gamma^*$, the economy remains in the no-machine steady state, in which capital $\bar{k}$ is pinned down by $\beta AF_k\left(1, \bar{k}, 1\right) = 1$, output, interest rates, wages, rents, and land prices remain constant, and the human labor share is strictly interior, $s_h \in (0, 1)$.
(ii) if $\gamma < \gamma^*$, the economy invests in machine labor and converges to a new steady state characterized by $\beta AF_m\left(1 + m^*, k^*, 1\right) = \gamma$ and $\beta AF_k\left(1 + m^*, k^*, 1\right) = 1$. Wages and the human labor share unambiguously go down throughout the transition, with wages declining to γ/β. The interest rate spikes and then converges back to $R = 1/\beta$. If land is a Hicks complement to labor, the price of land, $q_t = \beta AF_\ell\left(1 + m^*, k^*, 1\right)$, goes up.

Proof. See appendix. □

Our long-run results in Proposition 4 hold whenever there is an irreproducible factor that introduces decreasing returns for the remaining reproducible factors into the economy. Other than land, examples for such factors include:

Energy is commonly considered as a complementary factor to raw human labor in models of technological progress. Since skill can be acquired over time, it is reproducible. Our results in (i.a) suggest that the returns on skilled labor will temporarily benefit from improvements in machine technology.

Decreasing returns in aggregate Whenever the aggregate production technology of an economy is described by a decreasing-returns-to-scale production function, the technology can equivalently be viewed as exhibiting one further unstated factor *access to the production technology* that is irreproducible.

Efficiency Wages and Phasing Out of Human Labor An important characteristic of labor markets is that wages are not always set at their market-clearing level. In particular, most theories of efficiency wages recognize that there is a lower bound on wages that cannot be undercut, for reasons ranging from deterring shirking (Shapiro and Stiglitz, 1984) to nutritional necessity [reference tk]. Here we appeal to the latter set of theories and assume that there is a minimum wage $\underline{w}$ that workers need to receive in order to afford the nutirition necessary to perform their job. This implies the following:

Proposition 5 (Efficiency Wages and Unemployment)**.** *If there is a minimum wage* $F_h(1+m^*, k^*, 1) < \underline{w} < AF_h\left(1, \bar{k}, 1\right)$*, there is a critical threshold* $\gamma^{**} = \beta A / \underline{w}$ *for the cost of machines such that a decline* $\gamma < \gamma^{**}$ *will first lead to a drop in the wage to* $\underline{w}$*, followed by a complete phasing out of human labor to* $h = 0$*, and a subsequent further decline in the market wage of machine labor.*

Proof. See appendix. □

A sample path of wages, human and machine employment, and output is provided in the figure below.

Graphic to be inserted

4 Inequality and Public Policy

This section separates factor ownership into the hands of two agents, capitalists and hand-to-mouth workers, and studies the effects of redistributive policies in the set of models considered above.

- with perfect instruments (i.e. lump sum redistribution):
  - if utility functions depend on absolutes: Pareto-improvement is easily possible
    - since economy is getting more wealthy, all we have to do is compensate workers for their existing income
    - in the case when all factors are reproducible, the share of income that is actually redistributed goes to zero
  - with relative wealth concerns in utility function, transfers need to be scaled up proportionately
    - gains would have to be shared proportionately, with share depending on trade-off between level and relative effects. If just relative, then they would have to get constant share of gdp. If some level effect, then the share of income redistributed could go down.

- analyze constraints on set of instruments:
  - if taxes on irreproducible factors available = like lump-sum
  - taxes on reproducible factors introduce equity/efficiency trade-off
  - Intellectual property rules can be thought of as a tax on innovation

5 Machine Labor and Endogenous Technological Progress

To be completed.

References

Berg, A., Buffie, E., and Zanna, F. (2017). Robots, growth and inequality: Should we fear the robot revolution? *IMF Working Paper*.

Frey, C. B. and Osborne, M. (2013). The future of employment. *manuscript, Oxford Martin School*.

Kurzweil, R. (2005). *The Singularity Is Near*. Viking.

Shapiro, C. and Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review*, 74(3).

A Mathematical Proofs

To be completed.