# Economic Growth Given Machine Intelligence

**Robin Hanson*** HANSON@ECON.BERKELEY.EDU
*School of Public Health, 140 Warren Hall,*
*University of California, Berkeley, CA 94720-7360 USA*†

## Abstract

A simple exogenous growth model gives conservative estimates of the economic implications of machine intelligence. Machines complement human labor when they become more productive at the jobs they perform, but machines also substitute for human labor by taking over human jobs. At first, expensive hardware and software does only the few jobs where computers have the strongest advantage over humans. Eventually, computers do most jobs. At first, complementary effects dominate, and human wages rise with computer productivity. But eventually substitution can dominate, making wages fall as fast as computer prices now do. An intelligence population explosion makes per-intelligence consumption fall this fast, while economic growth rates rise by an order of magnitude or more. These results are robust to automating incrementally, and to distinguishing hardware, software, and human capital from other forms of capital.

## 1. Introduction

Since at least 1821 economists have recognized that dramatic consequences for wages and population can result from machines which can directly substitute for human labor (Ricardo, 1821; Samuelson, 1988). Yet since then machines seems to have mostly complemented human labor; new machine technology has seemed to raise, not lower, the demand for skilled labor. If anything, computers seem to have intensified this trend. The standard economic view seems to be that this trend will continue into the indefinite future (Simon, 1977).

Some economists, however, including a few famous ones (Keynes, 1933; Leontief, 1982), have forecast that machines will eventually substitute for most human labor. Popular fears of automation do not seem to have abated much, and many computer and robotics researchers have made strong predictions that just as technology first complemented horses in transportation (e.g., carriages), but later substituted for them (e.g., cars), computers which now complement human labor will eventually be intelligent enough to substitute for most such labor (Nilsson, 1985).

Today, Artificial Intelligence and Robotics are long-standing fields of engineering whose explicit goals are to develop principles of design to eventually enable machines to accomplish all human cognitive and physical tasks. These fields have made impressive progress over recent decades, though they also clearly have a long way to go. Another possible route to advanced machine intelligence is "uploading." An upload is a computer simulation of the particular neuronal connections observed in a particular human brain. If the observations

*. I thank Doug Bailey and Michael Wellman for helpful discussions, and the Robert Wood Johnson Foundation for financial support.
†. http://hanson.berkeley.edu 510-643-1884 FAX: 510-643-8614

and simulation are accurate enough, the simulated brain should be able to accomplish any cognitive task that the original brain could (Hanson, 1994). With a well-engineered android body, an upload could also accomplish all human physical tasks.

What are likely to be the long-term economic consequences of continued progress in developing machine intelligence? Beyond disputes over when machines have complemented vs. substituted for human labor, very little research seems to have been done on this question.

This paper examines the economic consequences of machine intelligence by using standard economic theory to construct a series of simple but increasingly more realistic formal models of economies where machines can both complement and substitute for human labor. Before describing these models in detail, we first summarize qualitatively the nature of our models and the conclusions they suggest.

## 2. Overview of Model and Results

As with models in any field, a good economic model includes the important relevant features of a system, while abstracting away from minor features whose inclusion would mainly obscure relationships between important features.

We are here interested in the wage, population, and economic growth consequences of machine intelligences. We thus focus here on how the amounts and prices of the major inputs to modern economic production change with time, paying special attention to the inputs provided by machine intelligences. Specifically, we consider human labor, human capital (e.g., education and training), computer hardware, computer software, and other forms of capital (e.g., factories, roads, and improvements to land).

To allow machines to both complement and substitute for human labor, we consider a continuum of jobs or roles in our economy, roles which either humans or machines can fill. Each job strongly complements all the others, in the sense that having one job done better increases the value (i.e, marginal product) of having the other jobs done well. Thus when computers do some jobs and humans do others, computers strongly complement humans.

In filling a particular role, however, a computer can also directly substitute for a human. If the relative advantage of humans over computers varies from job to job, and if computer hardware and software gets cheaper faster than comparable human hardware and software, then computers slowly take over more and more jobs.

We consider both an overall technology level, which can allow more total production from the same inputs, and computer-specific technology levels, which allows computer inputs to be produced more cheaply from other inputs. These technology levels vary slowly, and so there is no special date when machine intelligence is "invented" or achieved. Rather, computers become more widely used as they slowly get "better," i.e., can accomplish the same tasks more cheaply.

While we will usually assume that these technology levels change "exogenously,", i.e., independently of the rest of our model parameters, we also give an example of an "endogenous" growth where technology improves with production experience. Growth implications are much more dramatic with endogenous growth, as is typical for such models. To avoid suspicions that our dramatic implications are due to peculiar model features, this paper focuses on producing conservative estimates via exogenous technological growth.

We assume that the product of the economy can be either consumed or used to produce more of any kind of capital (i.e., human, hardware, software, or other). We assume this new capital is immediately available for further production, and we ignore capital depreciation or decay. We assume that investor arbitrage ensures that the price of capital reflects its marginal rate of return, and we consider competitive wages, set by the marginal product of human labor.

As steady exponential growth seems to describe the world economy over the last half century, economists typically seek steady growth solutions to such models. To allow this, we assume diminishing returns to all inputs (to avoid super-exponential growth), and assume investor preferences are such that interest rates remain constant under steady growth. (Since our models have no uncertainty, there is no need to distinguish risk-less from market rates of return.)

We assume a fixed rate of human population growth. We typically also assume that humans devote a fixed fraction of their hours to work, though at one point we show that a fixed work fraction is a plausible consequence of other assumptions under steady growth.

Given all of these assumptions and other modeling choices, these models can reconcile our historical experience with machines which have mostly complemented human labor with projections of dramatic consequences from mature machine intelligence. This is because the behavior of our models depend on whether machines do a small or large fraction of jobs.

When computers are very expensive, they do few jobs. Human labor is very important then, and so human labor wages rise as the economy grows. During this phase, computers are mostly irrelevant to growth, which is mainly determined by improving general technology magnified by increases in ordinary forms of capital.

When computers do most jobs, human labor is relatively unimportant, and whether human wages rise or fall depends on whether owners of capital place a strong special value on services that only humans can provide. If they do, human wages can rise with the economy again, but if not, then human wages fall faster than computer prices now do.

During this phase, economic growth is much faster. Some conservative parameter choices suggest it could be an order of magnitude faster or more. Faster growth is due in part to the fact that computer technology is now more important, and computer technology improves faster than general technology.

Faster growth is also due to the fact that, in contrast with a slowly growing human population, we assume that the population of machine intelligences can increase as fast as needed to keep up with the demand for labor. The population of machine intelligences grows very fast in all periods, growing faster than the rate at which computer prices fall *plus* the rate at which the economy grows. A Malthusian analysis of population is appropriate for this era, and per-intelligence consumption levels fall rapidly. Per-human consumption levels can rise rapidly, however, if humans own a fixed fraction of capital.

We now derive these results explicitly in a series of three simple but increasingly realistic models of an economy containing machine intelligences. The first model is extremely simple, containing just enough structure to say something about how machine intelligence might effect growth rates and wages. We show how this simple model is effected by choices of leisure time, by a reside of exclusively human jobs, and by learning by doing. The next two models show that the results of this first model are robust to two extentions: to smoothing the transition via a continuum of job types, and to including "human" capital in the analysis.

## 3. The Simplest Model

Let us now take a standard neo-classical (Solow-Swan) growth model with diminishing returns, and modify it minimally to include ordinary computers and machine intelligence.

First we have a Cobb-Douglas production equation,

$$Y = Y(A, L, K, M) = AL^{\alpha}K^{\beta}M^{\gamma}, \tag{1}$$

where $Y$ is the rate at which stuff can be produced, $A$ is a general technology level, $L$ is labor, $M$ is ordinary computer capital, and $K$ includes all other forms of capital (including "human" capital, e.g., education and training). Note that the marginal products for these inputs, $Y_L, Y_K, Y_M$ (the partial derivative of $Y$ with respect to these inputs), satisfy $Y_L = \alpha Y/L$, $Y_K = \beta Y/K$, and $Y_M = \alpha Y/M$.

In a competitive economy with $\alpha + \beta + \gamma = 1$, the parameters $\alpha, \beta, \gamma$ describe the share of total product $Y$ paid to each input to production, $L, K, M$. Competition would ensure that each unit of an input received its marginal product, and with $\alpha + \beta + \gamma = 1$, total marginal product equals total average product. Thus, for example, $\alpha$ would describe the fraction of total product paid to labor $L$ as wages, and $\beta$ would describe the fraction paid as interest on capital $K$ invested.

Let us assume instead that $\alpha + \beta + \gamma < 1$, giving diminishing returns to the set of all inputs. In this case, total marginal product is less than total product, and all we can say is that competitively priced units should be paid their marginal product. Thus the competitive wage for a unit of labor is $Y_L$ and $\alpha, \beta, \gamma$ give lower bounds on the share of total product paid to each input when all units of an input are paid at least their marginal product.

To allow for the possibility of machine intelligence, let labor $L = H + U$, where $H$ is human labor, and $U$ is labor by machines intelligent enough to substitute for human labor. (Ordinary computers complement human labor.) Let us also assume that the spending equation is

$$Y = C + K' + P(M' + U'),$$

where $C$ is the rate of consumption, $K', M', U'$ are time rates of change of these forms of capital, and $P$ is the "price" (really cost) of computers.

Note that the main reason for distinguishing computer capital from other capital in our model is because computer hardware prices have been falling much faster than other prices for a long time. Thus whether some computer-related capital, such as software, is considered part of $M$ or part of $K$ depends on whether its price has been falling more like computer hardware prices, or more like other forms of capital.

Let us assume that exogenous functions determine the growth in population $H(t)$ and improvements in technology $A(t), P(t)$. Note that we do not model any variation in leisure time or fertility. We have also assumed zero capital depreciation, and that producing more of each form of capital requires the same relative shares $\alpha, \beta, \gamma$ of other forms of capital. We have further assumed that all forms of capital, including machine intelligence, can be instantly produced and are then instantly available to aid production.

Let us also assume that the interest rate $I$ that investors demand remains constant when total product $Y$ grows exponentially, i.e., when $\ln' Y$ is constant, where $\ln' F \equiv (\ln F)' =$

$F'/F$ and $F'$ is the time derivative of $F$ for any $F$. The interest rate should be constant, for example, when each investor consumes some constant fraction of total product, and when all investors have the same inter-temporal elasticity of substitution (e.g., the same discount rate and logarithmic utility on their consumption rate; see (Barro, 1995) p.65). Finally, let us assume that investment arbitrage eliminates opportunities to obtain returns greater than $I$ by varying $K, M, U$.

Arbitrage sets the marginal product of capital. Consider a small increase in general capital production, financed by a small reduction in consumption, followed a time $\Delta t$ later by an equal general capital decrease. If the extra product from this extra capital is consumed instead of invested, then capital amounts for the economy before and after this $\Delta t$ period remain unchanged. If interest $I$ is constant over this period, investors are indifferent to this plan if

$$\int_0^{\Delta t} e^{-It} Y_K(t) dt = 1 - e^{-I\Delta t}.$$

This implies $Y_K = I$ both for $I$ constant *and* for $I$ varying, since in the latter case we can consider the limit as $\Delta t$ goes to zero.

Arbitrage also sets the marginal product of computer capital. A small increase in computer capital followed by a matching decrease $\Delta t$ later makes investors indifferent if

$$\int_0^{\Delta t} e^{-It} Y_M(t) dt = P(0) - P(\Delta t) e^{-I\Delta t} = P(0)(1 - e^{\Delta \ln P - I\Delta t})$$

This is solved by

$$Y_M = (I - \ln' P)P. \qquad (2)$$

Note how declining computer prices make computer investments less attractive; investors would rather wait till prices fall further.

If computer prices $P$ are initially very high, then very few will be bought. If computer prices then fall, more ordinary computers will appear. At first machine intelligence would be prohibitively expensive, and so none would be produced, with $U = 0$. As the computer price $P$ falls, investments in machine intelligence $U$ would remain unattractive as long as $Y_L < Y_M$. Since $Y_L = \alpha Y/H$ in this period, competitive wages are always proportional to per-capita production. Thus when production rises, wages rise, as is our current experience.

In contrast, when investments in $U$ become attractive, so that machine intelligences are produced, with $U > 0$, arbitrage should ensure that spending on $M'$ vs. $U'$ gives equal profits. This requires that $Y_L = Y_M$. Such arbitrage requires that those who help pay to create new machine intelligences be repaid from the wages such machines will earn. This is possible via either autonomous machines who repay their debts, or via directly owned machines.

By our computer marginal product equation (2), $Y_L = Y_M$ implies that wages will be falling with computer price $P$ unless interest rates rise very rapidly. And for steady consumption growth, interest rates must be constant. Since competitive wages are proportional to per-intelligence production, rapidly declining wages imply rapidly declining per-intelligence production, due to an intelligence population growing much faster than total production. Wages might well fall below human subsistence levels, if machine subsistence levels were lower.

This model seems to confirm the intuition that machine intelligence has Malthusian implications for population and wages. Note, however, that these results may be consistent with a rapidly rising per-capita income for humans, if humans retain a constant fraction of capital, perhaps including the wages of machine intelligences, either directly via ownership or indirectly via debt.

Since our results have been expressed in terms of marginal products, let us express the production equation (1) in these terms, as in

$$Y^{1-\alpha-\beta-\gamma} = \alpha^\alpha \beta^\beta \gamma^\gamma A Y_L^{-\alpha} Y_K^{-\beta} Y_M^{-\gamma}.$$

Expressed in terms of growth rates, this is

$$(1-\alpha-\beta-\gamma)\ln' Y = \ln' A - \alpha \ln' Y_L - \beta \ln' Y_K - \gamma \ln' Y_M. \tag{3}$$

To see what effect a transition to machine intelligence may have on steady state growth rates, let us assume a steady decline in computer prices, with $\ln' P$ a constant. Since we earlier assumed constant interest rates under constant product growth, we have $\ln' Y_K = 0$ and $\ln' Y_M = \ln' P$. We also know that when $U = 0$, $\ln' Y_L = \ln' Y - \ln' H$, and when $U > 0$, $\ln' Y_L = \ln' Y_M$.

Putting this all into the growth rate equation (3), we get

$$\ln' Y = \frac{\ln' A + \tilde{\alpha} \ln' H - \tilde{\gamma} \ln' P}{1 - \beta - \tilde{\gamma}},$$

which for $\tilde{\alpha} = \alpha$ and $\tilde{\gamma} = \gamma$ is valid when $U = 0$, and for $\tilde{\alpha} = 0$ and $\tilde{\gamma} = \alpha + \gamma$ is valid for $U > 0$. Thus steady state growth rates *are* different with machine intelligence, and the difference can be thought of as a change in the product shares.

To see how different the growth rates are, consider some conservative parameter values. Let the product shares be $\alpha = .25$, $\beta = .5$, and $\gamma = .02$, let annual growth rates be 1.5% for the human population $H$ and 1% for general technology $A$, and let computer prices $P$ halve every two years. Without machine intelligence, world product grows at a familiar rate of 4.3% per year, doubling every 16 years, with about 40% of technological progress coming from ordinary computers. With machine intelligence, the (instantaneous) annual growth rate would be 45%, ten times higher, making world product double every 18 months! If the product shares are raised by 20%, and general technology growth is lowered to preserve the 4.4% figure, the new doubling time falls to less than 6 months.

Such rapid growth may seem so far from historical experience that it should be rejected out of hand. However, an empirical projection of historical trends in world economic growth makes a median prediction that the economy will transition to a doubling time of one to two years around 2025 (Hanson, 1998). While this prediction has large uncertainties, it suggests that we not reject out of hand the predictions of this model of machine intelligence. (Coincidentally, 2025 is also roughly when long-steady computer hardware price trends give cheap computers with hardware as powerful as the human brain, and so is a favorite date for predicting when machine intelligence will arrive for those who believe hardware is the limiting factor (Moravec, 1998).)

## 4. Some Variations

Before considering more complex substitutes for the above model, let us consider some simple variations on it.

We have so far assumed that human labor is proportional to the human population. What if people instead choose the fraction of time they devote to work based on their preferences for consumption, leisure time, and the fulfillment of working?

Imagine that a human owns some constant fraction $\epsilon$ of total output $Y$, due to owning some constant fraction of labor and each form of capital. And imagine that her utility is a simple product of terms for consumption, leisure, and work,

$$\int e^{-\delta t} \left(\epsilon Y(t) + Y_L(t) l(t)\right)^{\xi} l(t)^{\kappa} \left(1 - l(t)\right)^{\zeta} dt,$$

where $l(t)$ is the fraction of time spent working. Work fraction $l(t)$ will stay constant if either the ratio of product $Y$ to wages $Y_L$ stays constant, as it does before machine intelligence, or if wages become negligibly small, as they eventually do after machine intelligence. Naturally, people work fewer hours when wages are negligibly small.

If people consume some constant fraction of total product, then people may demand certain services which can only be performed by humans. There may, for example, be a status value in being served by a real human rather than a machine imitation. In this case, human labor share $\tilde{\alpha}$ doesn't drop all the way to zero, and if $\tilde{\alpha}$ holds constant, wages rise with total world product even after machine intelligence, according to $Y_L = \tilde{\alpha} Y/H$. (Note that the work fraction $l(t)$ can still remain constant.)

Having some exclusively human jobs slows growth somewhat. In our previous numerical example, if $\tilde{\alpha}$ falls from .25 to only .05, instead of to zero, then the economic doubling time falls from 16 years to 27 months, instead of to 18 months.

We have assumed constant exogenous rates of improvement for computer and other technology. Let us now consider the simplest endogenous growth model, learning by doing (Solow, 1997), where technological ability goes as some power of total experience so far, as in

$$A \propto \left[\int_{-\infty}^{t} Y(s) ds\right]^{\phi} \qquad P \propto \left[\int_{-\infty}^{t} M(s) ds\right]^{-\psi}$$

for $\phi, \psi < 1$. For steady growth, i.e., $\ln' Y$ a constant, this implies $\ln' A = \phi \ln' Y$ and $-\ln' P = \psi \ln' Y$.

Substituting these into our growth equation (3), we get

$$\ln' Y = \frac{\tilde{\alpha} \ln' H}{1 - \beta - \phi - \tilde{\gamma}/(1-\psi)}.$$

This equation is valid when this expression is positive. Otherwise, we no longer have diminishing returns and steady growth solutions are no longer possible; our model predicts a continually accelerating growth rate.

Given our previous conservative parameter values, the values $\phi = .23$, and $\psi = .89$ reproduce our familiar economic doubling time of 16 years arises via a 1% annual growth in general technology $A$, and a halving of computer prices $P$ every two years. Using these $\phi, \psi$

values, lowering $\tilde{\alpha}$ just a little, from .25 to .241, reduces the economic doubling time from 16 years to 13 months. (In our exogenous growth model, this just reduced the doubling time to 14 years.) Reducing $\tilde{\alpha}$ further to .24 eliminates diminishing returns and steady growth solutions entirely.

Clearly endogenous growth models are capable of producing much more dramatic implications of substituting machine intelligence for human labor. To remain conservative, however, we stick with exogenous growth models.

## 5. A Continuum of Job Types

In our simplest model, computers stay confined to a small sector of the economy until computers are suddenly equally likely to take on any human job, at which point machine intelligence begins substituting for human labor on a massive scale. In reality, automation seems to increase more incrementally, with computers slowly taking over more types of jobs once done by humans.

Let us now consider a continuum of job types $\theta$, distributed uniformly on the interval $[0, 1]$. In the limit of many small job types each contributing to the total product, we have

$$\ln Y = \ln A + \beta \ln K + \rho \int_0^1 \ln f(\theta) d\theta \tag{4}$$

where $f(\theta)$ is the contribution of job type $\theta$ to the total product.

Each job can be done by either a human or a computer, and the human has a job-type-specific productivity advantage $a(\theta)$. Thus

$$f(\theta) = a(\theta)h(\theta) + m(\theta)$$

where $h(\theta)$ and $m(\theta)$ are the human and computer labor densities across the job types. Total human and computer labor is given by

$$H = \int_0^1 h(\theta) d\theta \qquad M = \int_0^1 m(\theta) d\theta.$$

Note that in this model we do not explicitly distinguish between ordinary computers and machine intelligence. This distinction is implicit instead, in the type of job the computer does. Some job types $\theta$ may require more intelligence than others, and the human vs. computer advantage $a(\theta)$ may then reflect this difference.

Note also that while a computer may substitute for a particular human in a particular job, each job complements all the others, and so each computer is complementing all human workers in all other jobs. Thus this model is not in spirit contrary to our observation of ubiquitous cases where individual computers strongly complement individual human workers.

We again assume a falling computer "price" $P$, so that the expenditure equation becomes

$$Y' = C + K' + P \int_0^1 m'(\theta) d\theta.$$

This again implies $Y_K = I$ and $Y_M = P(I - \ln' P)$.

Our production equation (4) suggests we define a density marginal $Y_{f(\theta)} = \rho Y/f(\theta)$. The chain rule then gives marginal products

$$Y_{l(\theta)} = a(\theta)Y_{f(\theta)} = a(\theta)Y_{m(\theta)}.$$

If arbitrage ensures that marginal products are equal across job types with positive densities, then $h(\theta_1), h(\theta_2), m(\theta_1), m(\theta_2) > 0$ implies $Y_{h(\theta_1)} = Y_{h(\theta_2)}$ and $Y_{m(\theta_1)} = Y_{m(\theta_2)}$. These imply $a(\theta_1) = a(\theta_2)$, however.

Thus if we assume $a'(\theta) > 0$, there can be at most one type $\theta$ where we could have both $h(\theta), m(\theta) > 0$. We call this border type $b$. For all $\theta < b$, $h(\theta) = 0$ and for all $\theta > b$, $m(\theta) = 0$. Humans take the jobs in $[b, 1]$ where their advantage is higher, and computers take the remaining jobs in $[0, b]$. As $b$ increases, automation increases, with computers displacing humans in border jobs.

Arbitrage should ensure that $Y_m$ $(= Y_M)$ is constant in $[0, b]$, and so $bm = M$. Similarly a constant $Y_h$ $(= Y_H)$ in $[b, 1]$ gives $(1 - b)h = H$. At $b$, we get

$$\frac{M}{H} = a(b)\frac{b}{1-b}.$$

Again writing our production equation (4) in terms of marginal products, we get a growth equation

$$(1 - \rho - \beta)\ln' Y = \ln' A - \beta \ln' Y_K - \int_0^1 \ln' Y_f(\theta) d\theta.$$

If we define $a = a(b)$, $\alpha = \rho(1 - b)$ and $\gamma = b\rho$, and again assume a constant interest rate $I$, and constant computer improvement $\ln' P$, this growth equation can be rewritten as

$$(1 - \beta - \gamma)\ln' Y = \ln' A + \alpha \ln' H - \gamma \ln' P - \alpha \ln' \alpha - (\alpha + \gamma)\ln' a. \tag{5}$$

This is essentially the same result as from our simplest model, except that the product shares $\alpha, \gamma$ vary smoothly with time, and there are two extra terms on the right. The far right term slows growth down when $a$ increases rapidly, such as early on when $b$ is near zero. The term before it speeds growth up later when $b$ is near one. Thus, compared to the simpler model, this model can give an even larger difference between growth rates before and after a transition to machine intelligence dominating job types.

We can also examine changes in the human labor wage $Y_h$, by studying the expressions

$$\begin{aligned} \ln' Y_h &= \ln' Y - \ln' H + \ln'(1 - b) \\ &= \ln' Y_m + \ln' a. \end{aligned}$$

The first equation shows that for $b$ remaining far from 1, human wages go up as long as the economy grows faster than the human population, which has been our recent experience. The second equation shows that if $a(b)$ has an upper bound to force $\ln' a(b)$ to approach zero eventually, then human wages eventually fall nearly as fast as computer prices do (assuming constant interest $I$). However, if $a(b)$ is unbounded from above as $b$ goes to 1 and grows quickly enough with $b$, then human wages could continue to rise with time.

The population of machine intelligences rises as

$$\ln' M = -\ln' Y_m + \ln' Y + \ln' b,$$

and all three of these terms are positive, with the first two being large. Thus machine intelligence population rises very fast indeed, and total per-intelligence consumption must fall once machine intelligences are more numerous than humans.

We can bound the contribution of the extra terms in our growth equation (5) by noting that

$$\ln' a - \ln' \alpha = \ln'(Y/H) - \ln' Y_m.$$

Since all four terms are positive, and since the two right hand side terms are large, at least one of the left hand side terms must be large.

This suggests that the transition from human dominated labor to labor dominated by machine intelligence might rather rapid. For example, assume that the relative human advantage $a(b)$ between the 25th percentile job (at $b = .25$) and the 75th percentile job (at $b = .75$) is a factor of three. Then in just four years machines could go from doing 25% to 75% of the job types, if computer prices halved every two years, and per-human product grew an average of 22% per year over this period.

This section has shown that the results of our simplest model are robust to including a continuum of job types, which automation continuously advances along. Machine intelligence can bring dramatically increased world product growth rates, even faster growth in the intelligence population, and rapidly falling human wages.

## 6. Distinguishing Hardware From Software

So far we have lumped together under "capital" $K$ physical capital like plants, human capital like education and training, and computer related capital whose price is not falling rapidly like computer hardware prices are. Are our results robust to distinguishing these capital forms?

To find out, let us retain our previous form for total product,

$$\ln Y = \ln A + \mu \ln K + \rho \int_0^1 \ln f(\theta) d\theta,$$

but let us expand our description of each job type allow it to be performed by either a combination of human labor and human capital, or a combination of computer hardware and computer software. That is, for each $\theta \in [0, 1]$ let

$$f = a h^\sigma e^{1-\sigma} + m^\sigma s^{1-\sigma},$$

where $h(\theta)$ and $m(\theta)$ are human labor and computer hardware densities, and $e(\theta)$ and $s(\theta)$ are human capital and computer software densities.

The parameter $\sigma(\theta)$, assumed to be in $[0, 1]$, says roughly how important education and software is over raw labor and hardware. That is, $\sigma(\theta)$ says roughly how important "intelligence" is for job type $\theta$. We assume that $\sigma$ increases strictly with $\theta$. The parameter

$a(\theta)$ describes the advantage of humans over computers, and we assume that it can be written as

$$a = \nu^\sigma > 1,$$

where $\nu(\theta)$ increases with $\theta$. One way of rationalizing this is to assume that, all else equal between humans and computers, investment in human capital is more productive than investment in computer software, because humans are "smarter," with a better stock of general knowledge and cognitive skills to draw on.

We generalize our spending equation to be

$$Y' = C + K' + \int_0^1 (e'(\theta) + s'(\theta) + Pm'(\theta))d\theta.$$

which says that human capital, software, and other forms of capital all cost the same per unit.

Note that for simplicity we have explicitly assumed only that computer hardware, not software, prices fall especially rapidly compared with generic capital prices. Falling software prices can be modeled via having the human advantage $a(\theta, t)$ fall with time. All the following equations hold for that case as well.

If for some $\theta$ all densities $h, e, m, s$ are positive, then we can show that

$$\sigma^\sigma(1-\sigma)^{1-\sigma} = a f_h^\sigma f_e^{1-\sigma} = f_m^\sigma f_s^{1-\sigma}.$$

This equation cannot hold for two different values of $a$, however, if the relative values of marginals $f_h, f_e, f_m, f_s$ are held fixed across job types. Nor can it hold for two different values of $\sigma$ when $a$ is fixed. Thus we again have a single border job type $b$, with all $\theta < b$ being done by computers and all $\theta > b$ being done by humans. At the border, we have

$$Y_h = Y_m a(b)^{1/\sigma(b)},$$

which implies $Y_h > Y_m$.

We can re-express the contribution of a job type to total product as

$$\begin{aligned} f &= \frac{\rho Y \sigma^\sigma (1-\sigma)^{1-\sigma}}{Y_m^\sigma Y_s^{1-\sigma}} \qquad \text{for } \theta < b \\ &= \frac{\rho Y \sigma^\sigma (1-\sigma)^{1-\sigma}}{Y_h^\sigma Y_e^{1-\sigma}} a^2 \quad \text{for } \theta > b. \end{aligned}$$

Integrating this allows us to write a growth equation just as we did before.

If we define labor shares

$$\pi_s = \int_0^b \sigma(\theta)d\theta, \quad \pi_e = \int_b^1 \sigma(\theta)d\theta, \quad \pi_m = b - \pi_s, \quad \pi_h = 1 - \pi_s - \pi_e - \pi_m,$$

and define product shares $\alpha = \rho\pi_h$, $\beta = \mu + \rho(\pi_s + \pi_e)$, and $\gamma = \rho\pi_m$, then for a constant interest rate $I$ and constant computer price decline $\ln' P$, the growth equation becomes

$$(1-\beta-\gamma)\ln'Y = \ln'A + \alpha\ln'H - \gamma\ln'P - \alpha\ln'\alpha - 2(\alpha+\gamma+\beta-\mu)\ln'a.$$

which is the same as in the last model, except that the last term has a larger coefficient.

The results of the previous model regarding populations and wages carry over directly to this model, except that in those equations the parameter $b$ should be replaced by $\tilde{b} = \pi_m/(\pi_m+\pi_h)$, which is the relevant job ratio for computer hardware relative to human plus computer hardware. Thus all our previous results are robust to explicitly distinguishing human capital, hardware, and software from other forms of capital.

## 7. Discussion

This paper has used standard economic tools to model a non-standard economic assumption, namely that computers will eventually become cheap and capable enough to substitute for most human labor. By distinguishing the complementarity between different job roles from the possibility of substituting machines for humans in particular roles, we can reconcile our historical observation of rising human wages and machines which complement human labor with frequent predictions of eventual machine substitution and consequent falling human wages. Our models show that wages can rise a great deal for a long time before eventually falling dramatically.

These models also suggest that wholesale use of machine intelligence could increase economic growth rates by an order of magnitude or more. These increased growth rates are due to our assumptions that computer technology improves faster than general technology, and that the labor population of machine intelligences could grow as fast as desired to meet labor demand. This second assumption, however, suggests that Malthusian and Ricardian analyzes of population and wages may become appropriate once again. In particular, per-intelligence consumption should fall rapidly.

This analysis has likely underestimated the economic effects of machine intelligence in several ways. For example, we have usually assumed that rates of technological progress do not change when the economy grows faster, except for one demonstration of bigger change from an endogenous growth model. We have also assumed substantial diminishing returns to increasing all production inputs together, even though growth theory models often favor constant or increasing returns for this case. We have not, however, considered the possibility of creating new kinds of jobs.

These results are also robust both to incremental evolution via a continuum of job types, and to explicitly distinguishing human capital, software, and hardware from other forms of capital. If machine intelligences are a real prospect in the foreseeable future, their economic implications would seem to deserve closer scrutiny.

## References


Barro, R. J. (1995). *Economic Growth.* McGraw-Hill, New York.

Hanson, R. (1994). If uploads come first: The crack of a future dawn. *Extropy*, *6*(2), 10–15.

Hanson, R. (1998). Long term growth as a sequence of exponential modes. http://hanson.berkeley.edu/longgrow.html.

Keynes, J. M. (1933). *Essays in persuasion*. Macmillan, London.

Leontief, W. W. (1982). The distribution of work and income. *Scientific American*, *192*, 188–204.

Moravec, H. (1998). When will computer hardware match the human brain?. *Journal of Transhumanism*, *1*. http://www.transhumanist.com.

Nilsson, N., Cook, S., Kay, A., Duchin, F., Boden, M., & Chamot, D. (1983). Artificial intelligence: its impacts on human occupations and distribution of income. In *Proceedings 8th International Joint Conference on Artificial Intelligence* Karlsruhe, West Germany. William Kaufmann.

Nilsson, N. J. (1985). Artificial intelligence, employment, and income. *Human Systems Management*, *5*, 123–125.

Ricardo, D. (1821). *On the Principles of Political Economy and Taxation* (3rd. edition). Murray, London.

Samuelson, P. A. (1988). Mathematical vindication of Ricardo on machinery. *Journal of Political Economy*, *96*(21), 174–282.

Simon, H. A. (1977). *The new science of management decision*. Prentice Hall.

Solow, R. (1997). *Learning from 'Learning By Doing'*. Stanford University Press, Standord, California.