Growth, Automation, and the Long-Run Share of Labor[†]

Debraj Ray ⓡ *Dilip Mookherjee*[*]

September 2021

Abstract. We study the long run implications of workplace automation induced by capital accumulation. We describe a minimal set of sufficient conditions for sustained growth, along with a declining labor share of income in the long run: (i) a basic asymmetry between physical and human capital; (ii) the *technical* possibility of automation in each sector; (ii) a self-replication condition on the production function for robot services; (iv) asymptotic homotheticity (more generally neutrality) of demand, and (v) a minimal degree of patience or intergenerational altruism among a fraction of households. However, the displacement of human labor is gradual, and absolute real wages could rise indefinitely. The results obtain in the absence of any technical progress; they extend to endogenous technical progress even if such progress is not biased *ex ante* in favor of automation.

KEYWORDS: automation, inequality, factor shares, human capital, technical progress

JEL Classification Nos. D33, E24

[†]Mookherjee: Boston University; Ray: NYU and University of Warwick. Neither author has any conflict of interest to disclose. Mookherjee thanks the Department of Economics at NYU for hosting his visit in Fall 2017 when this project was started. Ray acknowledges funding from the National Science Foundation under grant SES-1851758. We are grateful to Joshua Gans, Erik Madsen, Pascual Restrepo and Jeffrey Sachs for useful conversations, and to three anonymous referees for useful comments. Author names are in random order.

[*]Corresponding Author, dilipm@bu.edu; Department of Economics, Boston University, 270 Bay State Road, Boston MA 02215.

1. Introduction

This paper describes a theory of automation and its implications for growth and labor income share in the long run. Our framework is considerably more general than existing models, and generates a number of distinctive predictions. The model features a countable infinity of final goods, and three intermediate goods: machine capital, robots and education (or more accurately, services produced by each of these intermediate goods). In each of these sectors, production takes place by combining machine capital with tasks performed by a combination of human labor and robot services. We impose minimal restrictions on the technology, except that capital and tasks are both essential inputs, and that it is technologically feasible — though not necessarily economically viable — for humans to be dispensed with in every task.[1] The model permits human capital accumulation that allows workers to shift occupations in response to the threat of automation.[2]

We show that under some conditions, sustained declines in labor share can be a consequence of progressive capital-deepening resulting from capital accumulation *alone* (rather than technical progress or rising markups). This result, in line with the evidence in Karabarbounis and Neiman (2014), can happen despite arbitrarily inelastic capital-labor substitution in individual sectors. At the same time, the environment that drives the decline in labor share is also closely connected with sustained economic growth per capita, so absolute wages can grow unboundedly at the same time that the labor *share* converges to zero. We describe the conditions, and explain how the limiting share of labor can be positive when any one of them is violated.

The baseline model deliberately abstracts from technical progress, though we later incorporate endogenously directed technical change. It exhibits the following phenomena:

(a) Endogenous accumulation of both physical and human capital, *but with the relative deepening of physical capital*;

(b) *Ongoing decline in the prices of capital goods relative to human wages*, driven by the deepening in (a);

(c) *Progressive automation* driven by the relative price decline in (b).

Features (a) and (b) are the implications of a fundamental asymmetry between physical and human capital. While individual claims to physical capital in any sector can be replicated and scaled

[1] The distinction between these notions is important, and we return to it in Section 2.2 when we develop the theory.

[2] Apart from constant returns to scale, no substantive restrictions are placed on technology, not even convexity. Production functions can vary across sectors. Human labor could be sector-specific, or migrate across sectors via education or training. Households can accumulate financial wealth can purchase education to move across occupations, and they are permitted to be heterogenous in their tastes, discount factors and initial endowments.

indefinitely, the same is not true for ownership of labor. Humans cannot be bought and sold the same way machines are. Instead, human capital accumulation takes the form of acquiring embodied skills for a specific task or occupation, a capacity always contained in *one* physical self. So workers can invest in human capital to an unbounded degree, but are subject to diminishing returns in the acquisition of skill within any one task, and to the possibility that new skills may be needed to switch occupations or tasks. Consequently, the returns to human capital acquisition are determined endogenously, and the pattern of household demand across goods produced by different sectors plays a central role in this determination.

In this setting, we first provide conditions for positive long run growth in per capita income: a *self-replication property* in the technology for producing “robots" or digital services, and a minimal threshold for *patience* or intergenerational altruism in preferences. The self-replication property, which we discuss in detail below, implies that the production of robots will be fully automated if the capital rental price falls sufficiently relative to the unit cost of tasks. It isn’t a universal condition, but does holds automatically in familiar settings such as Cobb-Douglas production. Self-replication bounds the price of robots relative to capital rentals, and we show that this condition, along with patience, generates long run growth as the economy is progressively freed from the constraint of a given endowment of human labor.

But endogenous growth is not the only implication. Because robot prices are bounded by self-replication, they must decline relative to human wages in a growing economy, for that growth must arise precisely from the accumulation of physical capital relative to human labor, inclusive of human capital accumulation. Therefore, if automation is *technically feasible* in any sector, it will also be a long run *economic outcome*, as long as that sector expands with the economy. By itself, this does not erode labor share, as humans could move from sector to sector. If, however, preferences are *asymptotically homothetic*, then we prove that this intersectoral movement cannot prevent a sustained decline in the labor share. Making transparent these conditions for the declining labor share is a central goal of our paper.

The share decline must perforce be gradual if there exist occupations where humans are sufficiently productive relative to robots: for then there are always sectors that are yet to be automated. But labor movement across sectors only attenuates the decline, without negating it. If preferences are asymptotically homothetic, there is just not enough demand to sustain the persistent scaling of human capital needed to ward off the decline. Section 3.8 argues that other conditions on preferences also deliver a similar result. We emphasize “preference neutrality" in particular: a condition stating that preferences do not particularly favor human-friendly sectors — nor do they necessarily disfavor them.

Thus automation is a double-edged sword. It is an engine of income growth. And yet that same engine causes the labor income share to asymptotically vanish. This dichotomy explains why a vanishing labor *share* can co-exist with sustained growth in *absolute* human wages. Proposition 2 formalizes this intuition by providing conditions for some human wages to grow without bound: the existence of essential sectors and occupations in which humans have sufficiently high marginal productivity relative to robots even as they are close to full displacement. If additionally the costs of occupation-switching are bounded, all human wages grow without bound. Automation can raise all boats — only not all at the same rate, with wage incomes growing slower than capital incomes.

Sections 3.5–3.8 explain why each of these conditions is required for the results, by dropping them one by one. This helps identify various pathways for the share of labor to remain positive in the long run: (a) a failure of the self-replication condition in the robot sector, (b) the impossibility of fully automating some sectors, (c) the possibility of unbounded human capital acquisition within occupations, and (d) the failure of asymptotic homotheticity in preferences.

The analysis so far abstracts entirely from technical progress, relying entirely on changes in prices of capital goods relative to human wages. Certainly, the labor share could remain positive if technical progress is biased in favor of humans. For instance, Acemoglu and Restrepo (2018) restrict technical progress to enlarge only the productive capacities of human labor. The asymmetry is then built in by assumption, rather than endogenously explained — giving rise to the question of the kind of bias that might emerge when the opportunities for technical progress are more evenly distributed between human labor, robots and capital services. To explore this question, Section 4 extends our model to permit directed technical progress in machine, human and robot productivities, but we explicitly assume no technological bias, either in favor of humans or against them. Such a model could also be reinterpreted (with a hedonic reinterpretation of the commodity space) as one in which new goods are created. In such a symmetric setting, with technical progress equally sensitive to the prices of machines, robots and humans, we show that our long run distributional implications continue to be robust. Progressive capital deepening ensures that in equilibrium the derived demand for innovations in capital productivity cannot be surpassed by those for innovations in human productivity. Hence technical progress cannot be biased in favor of humans.

While our motivation is primarily conceptual and intended to guide largely speculative thoughts about the future, our theory provides a potential explanation for the recent decline in labor shares documented by Karabarbounis and Neiman (2014). As Section 5 clarifies, such a theory can be distinguished from explanations based on capital-augmenting technical progress, human capital accumulation, rising markups and market concentration or declining bargaining power of labor unions. Its relevance lies in the evidence provided by Karabarbounis and Neiman (2014), that a substantial fraction of the decline in labor share worldwide is explained by falling capital prices, *even after controlling for capital-augmenting technical progress, markup rates and the skill composition of labor*. Their theoretical explanation for this result is based on elastic capital-labor substitution, an assumption which runs contrary to evidence provided in industry panel studies; see, e.g., Chirinko and Mallick (2014). Our model shows that a declining labor share can result from capital deepening even in the presence of inelastic capital-labor substitution in most sectors.

It is important to clarify that we do not address the question of inequality in the *personal* distribution of incomes. Nor do we argue that a growing functional divergence between capital and labor incomes *must* imply growing inequality in personal incomes. These issues require analyses of the composition of household investment between financial wealth and human capital, and in inequality of labor incomes. Suitable applications and extensions of our model are needed to study these questions, as elaborated in the concluding section.

Section 2 presents the baseline model. The main results are in Section 3, with related lines of discussion. Section 4 studies technical progress. Section 5 discusses the connections to existing literature in detail, while Section 6 concludes. Proofs are collected in an Appendix.

2. Baseline Model with No Technical Progress

2.1. Production. There is a countable collection I of consumption goods, indexed by i. In addition, there are three intermediate good sectors producing education, robot services, and machine capital. The index j serves as generic notation for any of these (consumption or intermediate good) sectors. Everything is producible, with the exception of raw human labor. That endowment is fixed (or normalized if population is growing), but human *capital* evolves as individuals make educational investment decisions, thereby moving across occupations.

A good is produced by combining machine capital with a set of tasks performed by a combination of robot and human services. For a human, Any specific task in a given sector constitutes an *occupation*. We shall use the terms "task" and "occupation" interchangeably. The set O_j comprises the set of tasks or occupations in sector j, and $O \equiv \cup_j O_j$ denotes the set of all occupations in the economy. In sector j, let k_j denote machine capital and $\boldsymbol{\lambda}_j = \{\lambda^o\}$ a finite vector of task quantities (indexed by $o \in O_j$). These combine to produce output y_j according to a production function:

$$y_j = f_j(k_j, \boldsymbol{\lambda}_j), \tag{1}$$

where f_j is increasing, smooth, and linearly homogeneous, with unbounded steepness at zero in each input, and $f_j(k, \boldsymbol{\lambda}) = 0$ when any input is 0.[3] No curvature restrictions are imposed.

[3]This condition bounds substitution across inputs near the "axes," not necessarily elsewhere.

The quantity λ^o of task $o \in O_j$ performed in turn depends on robot and human services employed in that task, according to:

$$\lambda^o = \lambda^o(h^o, r^o), \tag{2}$$

where h^o is human input, r^o is robot services, and λ^o is increasing, smooth and linearly homogeneous with $\lambda^o(0,0) = 0$ (again, with no assumption on curvature).

Nothing of substance is lost by presuming that machine capital and robot services are homogeneous and can move freely across tasks and sectors, so we assume this.[4] In contrast, such considerations are important for humans. To enter a given occupation, an individual may require a suitable skill which can be acquired via education. Individuals will be born with some innate distribution of human capital, represented by occupations that they can work in without any formal education, and they can decide to augment the set of occupations they are eligible to work in by acquiring necessary education. Specific assumptions regarding educational requirements are given below.

2.2. The Feasibiity of Automation. We presume that it is feasible to automate every task; that is, for each o, $\lambda^o(0, r) > 0$ for some $r > 0$. It should be noted that this presumed *technological* feasibility of automation does not imply its *economic* viability. For instance, if $\lambda^o(h, r) = \nu r + \mu h + r^\alpha h^{1-\alpha}$ for for $\nu > 0$, $\mu > 0$, and $\alpha \in (0, 1)$, then humans would be perennially employed in every task, no matter what factor prices are.

More generally, economic viability depends on prices. If a burger chain is not automated, does it mean that it is fundamentally impossible at the current state of knowledge to automate hamburger production, or is it because wage conditions dictate that automation is not currently *economically* viable? We would submit that it is the latter. But the technical feasibility of automation goes far beyond routine tasks. Scott Santens (2016) has argued that

> "All work can be divided into four types: routine and nonroutine, cognitive and manual ... Machines that can learn mean *nothing* humans do as a job is uniquely safe anymore. From hamburgers to healthcare, machines can be created to successfully perform such tasks with no need or less need for humans, and at lower costs than humans ... A world with Amelia and Viv — and the countless other AI counterparts coming online soon — in combination with robots like Boston Dynamics' next generation Atlas portends, is a world where machines can do all four types of jobs and that means serious societal reconsiderations ... These exponential advances, most notably in forms of artificial intelligence limited to

[4] This assumption is inessential, in the sense that similar results will apply if we extend the technological assumptions below to the production of different sector-specific capital goods and robots. The growth result requires both the feasibility of automation and self-replication to hold for at least one capital good in an essential sector and the corresponding robots needed both directly and indirectly to produce this capital good. The distributional results will additionally require self-replication to apply to the production technology for every type of robot.

specific tasks, we are entirely unprepared for as long as we continue to insist upon employment as our primary source of income."

Thus the view we adopt here is this: the technology is already upon us. The question is one of the economic implementation of that technology. That said, in Section 3.7 we discuss how our results are modified if full automation is not technologically feasible in some tasks or sectors.

2.3. Prices. Within any date, machine capital services serve as numeraire: the rental price of k is set to 1. The collection $\boldsymbol{w} = \{w^o\}$ for $o \in \cup_j O_j$ is the wage system. Output prices are $(\mathbf{p}, p_r, p_e, p_k)$ for final goods, robot services, education, and capital. By constant returns to scale and the assumption of a competitive economy, all prices will equal unit costs of production for any sector with strictly positive output:

$$p_j \leq c_j(1, \boldsymbol{q}_j), \text{ with equality if } y_j > 0, \tag{3}$$

where 1 is the return to machine capital, $\boldsymbol{q}_j$ is the price vector of occupations in sector j, and c_j is the *unit cost function*, dual to the function f_j.[5] The prices of occupations, in turn, come from a second collection of unit cost functions $\{c^o\}$ for each occupation in that sector:

$$q^o = c^o(w^o, p_r). \tag{4}$$

2.4. Factor Demands and Automation. In each sector, machine capital and task levels are chosen to maximize profits, satisfying familiar first-order necessary conditions when an input is positive. The mapping from prices to human and robot demand then flows through the aggregators $\boldsymbol{\lambda}_j$. Consider the sub-problem where for any task o, the human-robot input mix is chosen to minimize the unit cost of producing the aggregator λ^o. By the linear homogeneity of λ^o, these depend only on the ratio of wages to price of robot services $\zeta^o \equiv w^o/p_r$. (Machine capital is not used in task creation and does not enter the picture here.) The *automation index* a^o tracks the vulnerability of occupation o to the robot threat, and is given by

$$a^o(\zeta) \equiv \min_{(r^o,h^o)} \left\{ \frac{r^o}{h^o\zeta + r^o} \middle| (r^o, h^o) \text{ minimizes unit cost under factor price ratio } \quad \zeta^o = \zeta \right\},$$

taking values between 0 and 1. We can extend this definition to the sector as a whole. For any wage vector $\boldsymbol{w}$ and robot price p_r, the above unit cost problems generate an input vector $\boldsymbol{q}_j$ for the aggregators in that sector. With these, solve the unit cost problem for the output of sector j. We

[5]Our results easily extend to monopolistic competition with CES preferences, which generates a constant profit markup in all sectors. Profits would appear in that setting, so national income would be the sum of returns to capital, to workers and profits. Our distributional results would continue to apply.

can then define the automation index of sector j by

$$a_j(\boldsymbol{w}, p_r) \equiv \min \sum_{o \in O_j} \frac{q^o \lambda^o}{\sum_{o' \in O_j} q^{o'} \lambda^{o'}} a^o \left(w^o/p_r\right),$$

where the minimum is taken over all aggregator vectors $\boldsymbol{\lambda}_j$ that solve the unit cost problem. These aggregator vectors depend on all factor prices, including the rental rate of capital, but this last term has been normalized to 1 and so does not appear explicitly here.

2.5. Accumulation. The aggregate stock of capital $K(t)$ evolves according to

$$K(t+1) = (1-\delta)K(t) + y_k(t), \tag{5}$$

where $\delta \in [0,1]$ is a constant, sector-independent depreciation rate for physical capital.[6] Only machine capital is formally durable, but durable robots are implicitly included by embedding them in physical capital in the robot sector, where they produce services under the robot production function f_r (along with other occupational inputs such as maintenance).

The stock of raw human labor is given (or normalized if population grows exogenously). But human capital can change endogenously with education. There is some initial allocation of humans across occupations. There could be a “null occupation" where individuals without initial skill can be placed, or can freely “drop out" to. An individual can move from occupation o to occupation o' (both in $\cup_j O_j$ but perhaps in the same or different sectors) at an educational cost of $e(o, o')$ times p_e, the endogenous unit cost of education. Human capital might depreciate; i.e., e_{oo} might be positive for some or all o. We place no restriction on the education needed to switch occupations, so the model captures both inflexible occupational specificity, or complete flexibility (with zero switching costs), or anything in between. All skill premia will be endogenously determined. In the baseline model we do not allow humans to augment their skill *within* any given occupation. Section 3.6 below will discuss how the analysis can be extended when we do allow within-occupation training, provided such training is subject to increasing marginal costs.

2.6. Preferences. A continuum of infinitely lived individuals, indexed by ι, is divided into a finite set of types, indexed by m. Each m has a one-period increasing, continuous,[7] strictly concave utility u_m on vectors of final goods, and a discount factor $\beta_m \in (0,1)$. Infinite lives can also be seen as a sequence of generations bound together by altruism. For any pre-determined current expenditure z on final goods and price vector $\mathbf{p}$, her chosen bundle maximizes $u_m(\mathbf{x})$, subject to $\mathbf{px} \leq z$. That generates a demand function $\mathbf{x}_m(\mathbf{p}, z)$. Denote by $v_m(z, \mathbf{p})$ the corresponding indirect utility

[6] The model can be extended to incorporate sector specificity of capital services and depreciation rates.

[7] The continuity of preferences or demand, here and everywhere else, will be taken relative to the pointwise or product topology on sequences of goods or price vectors.

function. We assume u_m is such that for every $\mathbf{p}$, the indirect function v_m is increasing, concave and differentiable, with unbounded steepness in z at zero.

Say that preferences satisfy *asymptotic homotheticity* if for every m:

$$\lim_{z\to\infty} \frac{\mathbf{x}_m(\mathbf{p}, z)}{z} = \mathbf{d}_m(\mathbf{p}) \text{ for some function } \mathbf{d}_m \tag{6}$$

for every $\mathbf{p} \gg 0$, where: (i) $\mathbf{d}_m$ is continuous on any bounded sequence of price vectors with strictly positive pointwise limit, and (ii) if there is a sequence $\{\mathbf{p}^n\}$ with some p_i^n converging to zero, then $\liminf_n d_{mi}(\mathbf{p}^n) > 0$ for at least one such i.

2.7. Household Optimization. At the start of any date, an individual has some financial wealth (representing her existing claims on capital or debt), and one unit of human labor along with a starting occupation. At date 0, her financial assets are nonnegative, and she can also work in a subsistence activity at any date to earn some small, exogenous, strictly positive income $\underline{w}$. We ignore the subsistence activity as it will get swamped in a growing economy: it is an expedient device to ensure a positive lower bound to human wages in all occupations.

At each date, individuals inelastically supply labor, make occupational choices (possibly with educational requirements), and implement consumption and savings decisions at endogenous prices, all within an infinite-horizon setting with perfect foresight. Given a dated price-wage system for goods, capital, and occupations, an individual of type m with initial (date-0) endowments of financial wealth $F(0) \geq 0$ and human capital (in occupation $o(-1)$) maximizes[8]

$$\sum_{t=0}^{\infty} \beta_m^t v_m(z(t), \mathbf{p}(t)), \tag{7}$$

by choosing a path of financial wealths $F(t)$ and occupations $o(t)$ at educational cost

$$E(t) \equiv e(o(t-1), o(t)), \tag{8}$$

along with current expenditure $z(t)$, subject to the date t budget constraint:

$$F(t) + w^{o(t)}(t) = z(t) + p_e(t)E(t) + \frac{F(t+1)}{\gamma(t)}, \tag{9}$$

and the no-Ponzi condition $\liminf_t F(t) \geq 0$. To accommodate imperfect capital markets, we impose $F(t) \geq B_m$ for all t, a borrowing limit that can be set arbitrarily high. Note that $\gamma(t)$ is the "return factor" on financial wealth at date t, and that:

$$\gamma(t) = \frac{1 + (1-\delta)p_k(t+1)}{p_k(t)}. \tag{10}$$

[8]We allow for heterogenous endowments and behavior within m, but drop the index ι here for ease in writing.

To understand (10), note that one unit of wealth can purchase claims to $\frac{1}{p_k(t)}$ units of physical capital at t. Each such unit generates a rental income of 1, then depreciates to yield $(1-\delta)$ units of physical capital worth $(1-\delta)p_k(t+1)$ at the next date.

A sufficient condition for the household optimization problem to be well-defined is that all utility functions are bounded. But well-known weaker conditions can be imposed; for instance, when utility functions have a well-defined tail elasticity.

2.8. Equilibrium. Given initial $K(0)$, an allocation of financial claims $\{F_\iota(0)\}$, and initial human capital $\{o_\iota(-1)\}$ (varying across or within types), an *equilibrium* is a sequence of wages $\{\boldsymbol{w}(t)\}$, prices $\{\mathbf{p}(t), p_r(t), p_e(t), p_k(t)\}$ and quantities $\{F_\iota(t), z_\iota(t), E_\iota(t), j_\iota(t), k_j(t), r_j(t), h_j(t), y_j(t)\}$, all non-negative and finite, such that:

A. Individuals maximize utility as described in (7)–(10), with $F_\iota(0) = p_k(0)k_\iota(0)$ for all ι, and firms maximize per-period profits at every date, with (3) holding.

B. The final goods markets clear: at every date, and for every final good i:

$$\sum_m \int_{\iota \in m} x_i(z_\iota(t), \mathbf{p}(t)) = y_i(t), \tag{11}$$

C. The robot market clears; for each t:

$$y_r(t) = \sum_i r_i(t) + r_r(t) + r_e(t) + r_k(t). \tag{12}$$

D. The labor market clears; for each t and each occupation o in sector j:

$$h^o(t) = \text{ Measure of } \iota \text{ such that } o_\iota(t) = o, \text{ with } w^o(t) \geq \underline{w} \text{ whenever } h^o(t) > 0. \tag{13}$$

E. The capital market clears; for each t, $K(t)$ evolves as in (5), with:

$$K(t) = \sum_i k_i(t) + k_r(t) + k_e(t) + k_k(t), \tag{14}$$

and the undepreciated capital stock plus rental income on it is willingly absorbed:

$$[1 + (1-\delta)p_k(t)]K(t) = \sum_m \int_{\iota \in m} F_\iota(t), \tag{15}$$

F. Finally, the education market clears; that is, for every t:

$$y_e(t) = \sum_m \int_{\iota \in m} E_\iota(t), \text{ where } \{E_\iota(t)\} \text{ satisfies (8).} \tag{16}$$

Per-capita national income (gross) is given by the expenditure on all final goods, plus investment in new capital goods and education:

$$Y(t) = \sum_i p_i(t)y_i(t) + p_e(t)y_e(t) + p_k(t)y_k(t). \tag{17}$$

In this paper, we do not go into the technicalities of equilibrium existence.

3. Long Run Growth, Automation and the Declining Labor Share

3.1. An Illustrative Example. There is a single occupation in each sector, so we use j to index these. There is one final good with production function $y_1 = k_1^{1/2}\lambda_1^{1/2}$, a capital goods sector with $y_k = k_k^{1/2}\lambda_k^{1/2}$, and a robot sector that has a CES production function with elasticity $1/2$:

$$y_r = \left[\frac{1}{2}k_r^{-1} + \frac{1}{2}\lambda_r^{-1}\right]^{-1}.$$

Humans and robots are substitutable at a constant rate ν everywhere: $\lambda_j = h_j + \nu r_j$ for all j. Humans move freely across sectors, so there is no education and just a single wage w. Then the occupational price q is w if there is no automation, and $\nu^{-1}p_r$ if there is (partial or full) automation. In the final good and machine sectors, the unit cost function is $c_1(1, q) = c_k(1, q) = \sqrt{q}$, while in the robot sector it is $c_r(1, q) = \frac{1}{2}\left[1 + \sqrt{q}\right]^2$. Everyone has the same one-period utility $u(x) = \ln(x)$, with discount factor $\beta \in (0, 1)$.

To track equilibrium paths, notice that at any date, robot prices must satisfy

$$p_r(t) \leq c_r(1, q_r(t)) = \frac{1}{2}\left[1 + \sqrt{q(t)}\right]^2. \tag{18}$$

with equality if the robot sector is active.

Case 1: $\nu \leq 1/2$. Then automation cannot ever occur. For if it did at any date t, then $q(t) = \nu^{-1}p_r(t)$. Substituting this into (18) which now holds with equality, we see that

$$p_r(t) = \frac{1}{2}\left[1 + \sqrt{\nu^{-1}p_r(t)}\right]^2 > \frac{1}{2}\nu^{-1}p_r(t),$$

which contradicts $\nu \leq 1/2$. So at every date the robot sector shuts down. The economy effectively reduces to a standard neoclassical growth model with a single consumption and capital good with aggregate Cobb-Douglas production, so there is no long run growth, while the share of labor in national income is 50% at every date.

Case 2: $\nu > 1/2$. Then, if the economy exhibits sustained growth of per-capita income — as it will if some household types are patient enough — all sectors j that grow must be "asymptotically automated": $a_j(t) = a_j(w(t), p_r(t)) \to 1$ as $t \to \infty$. For suppose not; then $a_j(\tau)$ must be bounded away from 1 in at least one growing sector j along a subsequence $\{\tau\}$ of dates. Since the total amount of human labor in the economy is bounded, so must be the *overall* occupational input in that sector. Then sustained growth implies that machine capital used in j — and hence the ratio of machine capital to occupational inputs — grows without bound, implying $w(\tau) \to \infty$. In the absence of full automation, unit occupational labor cost $q_j(\tau)$ will equal $w(\tau)$, and also converge to ∞. By (18),

$$p_r(\tau) \leq \frac{1}{2}\left[1 + \sqrt{q(\tau)}\right]^2 = \frac{1}{2}\left[1 + \sqrt{w(\tau)}\right]^2,$$

so that along the same subsequence,

$$\frac{\nu^{-1} p_r(\tau)}{w(\tau)} \leq \frac{1}{2\nu}\left[\frac{1}{\sqrt{w(\tau)}} + 1\right]^2 \to \frac{1}{2\nu} < 1 \text{ as } t \to \infty,$$

but that would imply $q_j(\tau) \leq \nu^{-1} p_r(\tau) < w(\tau)$ for large τ, a contradiction.

To provide some intuition, note that if $\nu > 1/2$, it is possible to dispense with humans altogether, and still produce robots at a finite unit cost (using machines and robots). More precisely, there exists $p_r^* < \infty$ satisfying $p_r^* = \frac{1}{2}\left[1 + \sqrt{\nu^{-1} p_r^*}\right]^2$, if and only if $\nu > 1/2$. Then p_r^* is an upper bound to the price of robots, which places a limit on human wages in this example. it follows that in any growing economy, the share of wages in national income must converge to 0 in the long run. Of course, in a more general setting, human wages will not necessarily be bounded even if robot prices are, and one will need to explore conditions under which growth will occur. Most crucially, we need to understand the underlying production conditions that generate the two cases described above. Our first task for the general model is to investigate this question.

3.2. Self-Replication. By the assumed feasibility of automation for every task, we know in particular that for every $o \in O_r$, $\lambda^o(0, r) > 0$ for some $r > 0$. By linear homogeneity, $\lambda^o(0, r)/r$ is independent of r for $r > 0$. Call this ratio ν^o. It represents the unit productivity of robots in the hypothetical event that the task in question is fully automated. Notice too that in that event, the unit cost of task o would equal $(\nu^o)^{-1} p_r$.

In order to interpret the condition below, temporarily switch the numeraire to robot services ($p^r \equiv 1$) rather than capital services. Consider unit cost minimization in the robot sector, with each task o in O_r priced at $(\nu^o)^{-1}$ per unit, and the capital rental rate is η. Consider the limit unit cost when the capital rental rate converges to zero:

$$\lim_{\eta \to 0} c_r\left(\eta, \{(\nu^o)^{-1}\}\right).$$

Observe that this limit depends only on the technology of the robot sector.

PROPOSITION **1.** *Suppose the robot sector satisfies the following self-replication condition:*

$$\lim_{\eta \to 0} c_r\left(\eta, \{(\nu^o)^{-1}\}\right) < 1. \tag{19}$$

Then there is a nonempty compact set P^* of strictly positive solutions to the equation

$$p_r = c_r\left(1, \{(\nu^o)^{-1} p_r\}\right), \tag{20}$$

and in equilibrium, $p_r(t) \leq \sup P^* < \infty$ for all t: the robot price is bounded relative to the rental on capital. If at any t, the robot sector is automated, then $p_r(t) \in P^*$.

We illustrate the argument graphically. Revert to using capital rental services as the numeraire. Because ν^o units of occupational input in o can be produced by a single robot unit, it must be that $q^o \leq (\nu^o)^{-1} p_r$. This option imposes an upper bound to the price of robot services:

$$p_r = c_r\left(1, \{q^o\}\right) \leq c_r\left(1, \{(\nu^o)^{-1} p_r\}\right). \tag{21}$$

Figure 1 depicts $c_r(1, \{(\nu^o)^{-1} p_r\})$. Because f_r has unbounded steepness in machine capital at zero, c_r lies above the 45^0 line for all strictly positive p_r sufficiently close to zero. At the same time, the self-replication condition (19) plus linear homogeneity guarantees that c_r ultimately dips and stays below the 45^0 line; see Panel A. Then P^* is the set of intersections with the 45^0 line, as described by (20). It is nonempty and compact,[9] and (21) is equivalent to the assertion that $p_r(t) \leq \sup P^*$ for all t in any equilibrium. So the price of robot services (relative to machines) is bounded above if self-replication holds. If the robot sector *is* automated, then that price must be one of the solutions in P^*, a result which can be viewed as a variant of the Nonsubstitution Theorem (Arrow 1951, Samuelson 1951). Of course, automation may never be full but only asymptotic, in which case the robot price converges to some element of P^*.

Conversely if self-replication fails, a non-zero solution to (20) could fail to exist, as shown in Panel B, and this will *necessarily* happen if f_r is quasi-concave. Then the robot producing sector can never be automated, and the price of robot prices is unbounded; see Proposition 3.

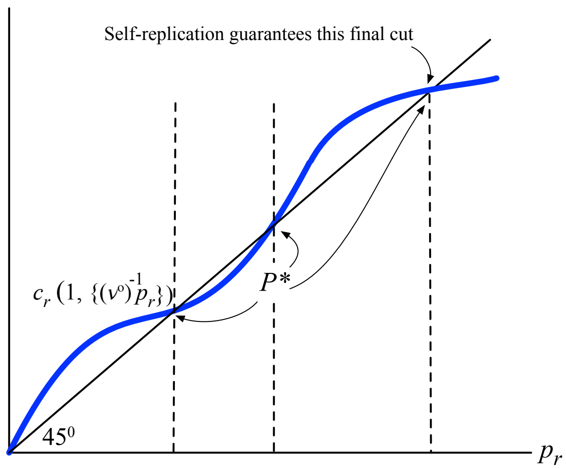
The self-replication condition (19) can be illustrated in the special CES class, assuming just one task in the robot sector. We have:

$$f_r(k, \lambda) = \left[\alpha k^{\frac{\sigma-1}{\sigma}} + (1-\alpha)\lambda^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}},$$

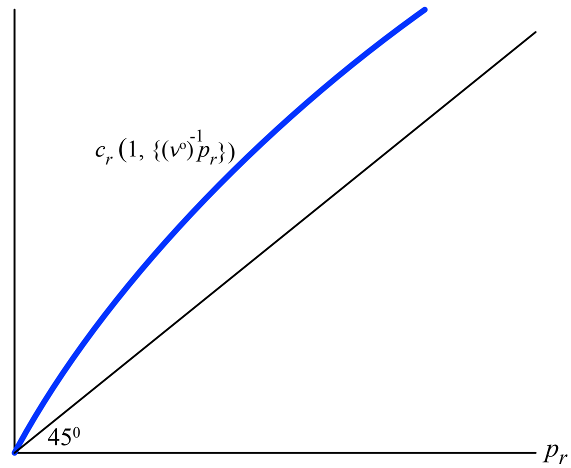
with $\alpha \in (0,1)$ and the elasticity of substitution $\sigma \geq 0$. The unit cost function is

$$c_r(\eta, \{\nu^r\}^{-1}) = \left[\alpha^\sigma \eta^{1-\sigma} + (1-\alpha)^\sigma \{\nu^r\}^{\sigma-1}\right]^{1/(1-\sigma)}.$$

[9] If f_r is quasi-concave, then P^* is a singleton — there is a unique positive solution to (20).



(a) Self-Replication Holds



(b) Self-Replication Fails

Figure 1. Replication and the Bound on Robot Price

where η denotes the capital rental rate and ν^r denotes the productivity of robots in the production of robots when it is fully automated. So our limit equals zero when $\sigma \geq 1$, which includes the Cobb-Douglas case (i.e., "enough" substitution is available). But it is positive when $\sigma < 1$. For instance, if the production function is "almost" Leontief, labor costs will matter for unit cost no matter how cheap machines are. In this latter case, (19) does restrict the value of ν_r. Specifically, self-replication reduces to the capital-labor substitution elasticity exceeding some lower bound smaller than one:

(22) $$\text{Either } \sigma \geq 1, \text{ or } \sigma \in (0,1) \text{ and } \nu_r > (1-\alpha)^{\sigma/1-\sigma}.$$

3.3. Automation and the Declining Labor Share Under Long Run Growth. We now present our main result: under sufficient household patience, the self-replication condition in the robot sector has strong implications for long run growth, automation and income distribution.

Assume that self-replication holds. Define the *patience condition* as follows:

(23) $$\beta_m \left[(1-\delta) + \frac{1}{c_k\,(1, \{(\nu^o)^{-1} \sup P^*\})} \right] > 1.$$

for some type m, where o ranges over O_k and P^* is defined as in Proposition 1.

THEOREM **1.** *Assume the robot self-replication condition* (19) *holds, as well as the patience condition* (23)*. Then:*

(i) *Per-capita national income grows without bound:* $Y(t) \to \infty$.

(ii) *Every sector j that grows exhibits asymptotic full automation:*

(24) $$a_j(\boldsymbol{w}(t), p_r(t)) \to 1 \text{ as } t \to \infty.$$

(iii) *If in addition household preferences are asymptotically homothetic, the share of human labor in national income converges to zero as $t \to \infty$, and that of physical capital converges to 1.*

We sketch the underlying argument; a formal proof is in the Appendix. Part (i) states that per capita income grows (without bound) if self-replication and patience hold. The first difficulty is to account for moving capital prices. While bounds can be placed on these prices, there will in general be capital gains (or losses). To sidestep the spikes of accumulation and decumulation that could arise from these anticipated gains and losses, we cumulate the relevant Euler equations for financial wealth. Recalling the indirect utilities v_m and $\gamma(t)$, the equilibrium rate of return on financial assets, we have:

(25) $$v'_m(z_m(t), \mathbf{p}(t)) \geq \beta_m \gamma(t) v'_m(z_m(t+1), \mathbf{p}(t+1)),$$

with equality holding if financial wealth is actively accumulated. From (10),

$$\gamma(t) = \frac{1 + (1-\delta)p_k(t+1)}{p_k(t)} = \left[\frac{p_k(t+1)}{p_k(t)}\right]\left[(1-\delta) + \frac{1}{p_k(t+1)}\right],$$

where the second equality decomposes the return into the product of capital gains and the rental income (augmented by any undepreciated capital) on a unit of wealth. If we compound the Euler inequality in (25) over dates $0, \ldots, t$, where $t \geq 2$, then we have

$$v'_m(z_m(0), \mathbf{p}(0)) \geq \beta_m^{t-1} \frac{(1-\delta)p_k(t) + 1}{p_k(0)} \left\{\prod_{\tau=1}^{t-1}\left[(1-\delta) + \frac{1}{p_k(\tau)}\right]\right\} v'_m(z_m(t), \mathbf{p}(t)),$$

which eliminates temporary spikes and dips in capital gains. The key observation is that the self-replication condition implies a finite upper bound to the price of machines, given the option to automate their production: $p_k(\tau) \leq c_k(1, \{(\nu^o)^{-1} p_r(\tau)\}) \leq c_k(1, \{(\nu^o)^{-1} \sup P^*\})$. In turn this limits the extent to which the value of capital goods can depreciate, implying a positive lower bound to the return to capital: $\left[(1-\delta) + \frac{1}{p_k(\tau)}\right] \geq \left[(1-\delta) + \frac{1}{c_k(1,\{(\nu^o)^{-1}\sup P^*\})}\right] \equiv (1 + \underline{r})$, say. The patience condition (23) then implies $\beta_m(1+\underline{r}) > 1$. Hence:

(26) $$v'_m(z_m(0), \mathbf{p}(0)) \geq \frac{\beta_m \left[\beta_m(1+\underline{r})\right]^{t-2}}{c_k\left(1, \{(\nu^o)^{-1}\sup P^*\}\right)} v'_m(z_m(t), \mathbf{p}(t)).$$

which implies that $v'_m(z_m(t), \mathbf{p}(t)) \to 0$ as $t \to \infty$. Further bounds on equilibrium prices $\{\mathbf{p}(t)\}$ (see Appendix) then imply that the consumption of type-m households must grow. With bounded debt, the same is true *a fortiori* for overall per-capita consumption and income.

Part (ii) asserts that any sector that grows exhibits asymptotic full automation in the sense of (24). If a sector grows, at least one of its task levels must grow with it — a consequence of unbounded steepness with respect to at least one task, in conjunction with self-replication (so the price of occupations is bounded relative to capital). But the total available supply of raw *human* labor is bounded. Therefore each growing task or occupation o within the sector must either have human labor equal to zero, or $w^o/p_r \to \infty$. In either case Lemma 3 in the Appendix implies that its automation index must converge to 1. Because all bounded occupations become insignificant relative to the growing ones, the result follows.

Part (iii) uses a more subtle argument. It is possible that there is no *uniform* threshold for automation — at any human wage, there could always be productive sectors where humans continue to be a desirable presence. In fact, humans may well be persistently present in every occupation, asymptotically automated or not,[10] but with asymptotic automation their income share cannot be preserved. However, non-uniform automation thresholds open the possibility of “human shelters" that provide opportunities for humans to stay ahead of automation waves. To do so, they must perennially accumulate human capital and move into occupations where human employment and wages are less threatened by automation. Indeed, in these relatively protected sectors, the human wage could be very high. In Proposition 2 below, we provide conditions under which in any equilibrium with growth, the highest human wage across all sectors grows *unboundedly* over time. If humans acquire the skills to enter these yet-to-be-automated sectors, their wages might conceivably grow in step with per capita income.

At this point, the endogeneity of prices and wages takes center stage. The willing absorption of humans into sectors requires that there be adequate demand for their outputs. Under the usual efficiency-units approach, this demand question is eliminated by construction: relative wages cannot change over sectors that are thus aggregated with brute force. With an endogenous wage structure, this is no longer the case. Part (iii) shows that if demand is asymptotically homothetic, then the economy runs out of steam in its ability to shelter labor. For the human wage share to stay positive in the long run, household expenditures shares on yet-to-be automated sectors must remain sizable. Under asymptotic homotheticity, this cannot happen: wage incentives do not climb at the required pace.

3.4. Long-Run Wages. The discussion in Section 3.3 suggests that a vanishing share of labor income could co-exist with unbounded growth in human wages. When Theorem 1 applies, universal automation ensures that prices of all consumer goods are bounded. Hence real wages relative to

[10]To see why, consider the example of an λ_j function provided just after equation (2).

any consumer price index are unbounded if and only if wages (as defined here) are unbounded. In this section we study conditions for this outcome.

Two forces could make for growing human wages even as their relative share declines. The first has to do with the rate of robot-substitution for humans in some *given* occupation. The second has to do with human movement *across* occupations as automation becomes more pervasive. Both are summarized in a single sequence of numbers. Consider a more general version of our example in Section 2.1 for the production of the input in occupation o:

$$\lambda^o = \lambda^o(h, r) = \nu^o r + \mu^o h + g^o(h, r),$$

where $\nu^o > 0$, $\mu^o \geq 0$, and g^o is a standard production function with $g^o(0, r) = g^o(h, 0) = 0$. Such an occupation could become automated, but if g^o has unbounded steepness in h at 0, *only* asymptotically so: human labor can never be fully dispensed with at any (finite) wage. More generally, let θ^o denote the limiting marginal rate of substitution of humans for robots as the ratio of human labor to robot services in occupation o converges to zero:

$$\theta^o \equiv \lim_{h/r \to 0} \frac{\partial \lambda^o / \partial r}{\partial \lambda^o / \partial h}.$$

This measures the “local relative efficiency" of robots relative to humans in occupation o as human labor vanishes. In the example, if g^o has unbounded steepness in h, $\theta^o = 0$. Otherwise, θ^o is determined by the slopes of the two functions $\lambda^o(0, r) = \nu^o r$ and $\lambda^o(h, 0) = \mu^o h$, and the limiting marginal rate of substitution from g^o. Our sequence of interest is $\{\theta^o\}$, where o ranges over all occupations in a subset to be described precisely below.

To highlight these forces, we place additional restrictions on education and preferences. Specifically, we assume that the education function is uniformly bounded: $\sup_{o,o'} e(o, o') < \infty$. In addition to asymptotic homotheticity of preferences, we presume that limiting demand has *full support*: for each type m and each price $\mathbf{p} \gg 0$, $d_{mi}(\mathbf{p}) > 0$ for $i \in I$. Finally, define $O_{-e} \equiv O - O_e$ to be the set of all occupations except those that pertain to the education sector.

PROPOSITION **2.** *Suppose that the conditions of Theorem 1 hold, including asymptotic homotheticity of preferences. In addition suppose limiting demand has full support and that the education function is uniformly bounded on $O \times O$. For every individual ι, let $w_\iota(t) = w^{o_\iota(t)}(t)$ be the human wage she receives at date t.*

(a) *If* $\inf_{o \in O_{-e}} \theta^o > 0$, *every human wage is bounded.*

(b) *If* $\inf_{o \in O_{-e}} \theta^o = 0$, *every human wage must grow without bound.*

Under the additional restriction of full support on preferences and uniformly bounded educational requirements, Proposition 2 provides a complete characterization of when all human wages can grow in an unbounded fashion. The limit condition on $\inf_{o\in O_{-e}} \theta^o$ captures both the possibility that the marginal product of human labor can climb in a particular occupation (even as that occupation progressively succumbs to automation), as well as the possibility that there is human “protection" available *across* occupations. When this term is strictly positive, neither the occupation-specific nor the cross-occupation protection is available, and limiting human wages are bounded. (While the education sector is exempt from this condition, we show that it can do nothing to overturn this result.)

On the other hand, when $\inf_{o\in O_{-e}} \theta^o = 0$, there is protection either from occupation-specific steepness in the marginal product of human labor, or relative cross-occupational proclivities to automation. In this second case, wages can grow in some occupations — or over some sequence of occupations — and then the boundedness of educational costs allows all individuals to participate in that growth.[11]

A bounded education function is used in these arguments. Significant progress can be made without it. It is possible to characterize sustained growth in *occupational* human wages, rather than personal wages, by dropping the uniform boundedness assumption. We stick to the present formulation as it is stronger and cleaner. In defense of the boundedness assumption, it should be noted that while we work with an infinity of final goods sectors, there is no need to suppose that the corresponding occupations will be dramatically different. For instance, two managerial roles in very different sectors could be very similar. The following mental picture may be useful. Think of occupations as belonging to some compact set in an abstract space C of characteristics, with the education function continuous on $C \times C$. Each sector draws a finite set of points from this space for its associated set of occupations. This formulation allows for fine distinctions across sector-specific occupations, as it should, while still retaining the uniform boundedness of educational in moving from one occupation to another. “Distant" goods need not be produced by “distant" sets of skills.

As a final remark: part (b) of Proposition 2 illustrates the fact that ubiquitous automation need not result in stagnation of real wages. On the contrary, automation could be the engine of long-run growth in wages. Observe that in the absence of robots, the patience condition would not hold; it is easy to write down examples where per capita income and wages would be bounded.

[11]Certainly, including education in part (a) of the proposition would still give us a sufficient condition for bounded wages. But that condition would not have been *necessary*. That is, if education is included as one of the sectors in the condition of part (b), that condition would not be sufficient for unbounded wage growth.

Hence automation can boost the living standards of workers, though the growth in wages would be outstripped by growth in capital incomes when Theorem 1 holds.

3.5. Failure of Self-Replication. A failure of self-replication means that robot prices cannot be severed from human wages. Human workers are indispensable in the production of robot services, so the price of the latter climbs with wages as labor scarcity grows. The scope for automation is then limited. The robot self-replication property is formally *necessary* for Theorem 1. The example in Section 3.1 already illustrates this point, but we elaborate on the distributional consequences when self-replication is violated.

To develop this argument, we place some restrictions on our general environment. Once again, without any real loss of generality, we assume just one occupation per sector, indexing occupations by their sector index. The first restriction is a general version of the condition that the production function f_j defined on capital and tasks has an elasticity of substitution smaller than 1 in every sector j. For any sector j, and any effective price of task q, consider the set $\Xi_j(q)$ of ratios of task service to machine capital $\xi = \lambda/k$ that minimize unit cost of production, and let

$$\Lambda_j(q) \equiv \min_{\xi \in \Xi_j(q)} \frac{q\xi(q)}{1 + q\xi(q)}$$

be the lowest ratio of the payment to sector j's task to the total sectoral cost outlay. Temporarily think of the production function f_j in this sector as CES with elasticity of substitution lower than 1. Then we know that $\Lambda_j(q)$ is increasing in q, with $\Lambda_j(0) = 0$ and $\Lambda_j(\infty) = 1$. In particular, given any lower bound $q_- > 0$, we have

$$\inf_{q \geq q_-} \Lambda_j(q) > 0.$$

In our more general setting without constant elasticity (or indeed concavity), we impose the above condition, and uniformly so across sectors:

$$\inf_j \inf_{q \geq q_-} \Lambda_j(q) > 0. \tag{27}$$

Next, we make additional assumptions on the production function for robots. We assume that it is strictly quasiconcave, in addition to being linearly homogeneous. We assume further that $\lambda_r(h, 0) > 0$ for some $h > 0$, so that the occupational aggregate in the robot sector can be produced by humans alone. This restriction is analogous to the feasibility of automation, though assuming it or not makes no difference to Theorem 1. Call such a technology *regular*.

PROPOSITION **3.** *Suppose that (27) holds and the robot production function is regular. Then, if the self-replication condition fails, in any equilibrium the share of human labor in national income is bounded away from zero.*

Proposition 3 shows that *both* the asymmetry of human and physical capital accumulation *and* the self-replication condition are needed for our results. Indeed, the latter condition is logically necessary in a broad class of environments. Without it, robot prices cannot be divorced from the wages of human labor. As labor becomes more expensive, so do robots, and the forces of automation are attenuated — sufficiently attenuated, as it turns out, under the conditions of Proposition 3 so that the share of human labor does not decline in a sustained way over time.

3.6. Within-Occupation Human Capital. We now discuss the asymmetry between the accumulation of physical and human capital in the preceding analysis, and extend the theory to incorporate the acquisition of intra-occupational skill. First note that the device of several occupations within a sector can be interpreted to mean that these are different skill levels within the same job. As long as there is a finite (or even compact) set of such skill levels, the theory already accommodates such cases, by redefining different levels of skill as different occupations. However, that is still in contrast to the unbounded scope for accumulation of physical capital within any sector. Could our model be extended to similarly accommodate the unbounded accumulation of skill within a sector?

We already know that the answer cannot be an unqualified yes: there are macroeconomic models which generate balanced labor income shares once human capital can be accumulated to an unbounded degree in efficiency units, with no changes in relative prices. So studying this extension will help identify the precise nature of the asymmetry needed between physical and human capital accumulation in our model.

For expositional clarity, we revert to the common-sense notion of an occupation, and do not interpret varying levels of skill as constituting distinct occupations. We extend our model to allow workers to acquire varying levels of skill within any given occupation, and place no upper bound on the amount of such skill that can be accumulated. We model skill in the conventional manner, as a certain number of efficiency units. Let the production function for task o in some sector be $\lambda^o(\mu^o h^o, r^o)$, where μ^o is the productivity of a human in that occupation. Wages are paid per unit of productivity, just as in the standard model based on efficiency units, so the income of a person with productivity μ^o is $w^o\mu^o$, where w^o is the occupation-specific “efficiency unit human wage."

Everything else in the model is kept unchanged, but we now need to specify the technology of productivity acquisition. To this end, we extend the education function as follows: let $e(\mu, \mu', o, o')$ denote the units of education needed to move from “starting productivity" μ in occupation o to “destination productivity" μ' in sector o', *where o could be equal to o'*. In particular, one can both invest within an occupation and across occupations, generally with heterogeneous cost implications. Moreover, continued on-the-job education can depend on baseline levels of productivity already acquired in that sector.

Assume that e is smooth in its first two arguments with partial derivatives e_1 (typically negative) and e_2 (typically positive). We place the following substantive restrictions on e:

(H.1) For any o and $S > 0$, there is $M < \infty$ such that $e_2(\mu, \mu, o, o) \geq S$ for all $\mu \geq M$.

(H.2) For any o, there is $L^o \geq 0$ with $e_1(\mu, \mu', o, o') \in [-L^o, 0]$ for all (μ, μ') and o'.

(H.3) For each occupation o', there is a bound $\hat{\mu}^{o'}$ such that for every starting $o \neq o'$ and productivity μ, $e(\mu, \mu', o, o') = \infty$ for $\mu' \geq \hat{\mu}^{o'}$.

(H.1) states that *within* any occupation, the marginal cost of skill acquisition becomes very high as baseline productivity increases. (H.2) states that while a higher starting productivity may bring down the cost of achieving any destination productivity in the same or different occupation, the marginal savings are bounded. (H.3) states if an individual is switching occupations, there is some upper bound to the productivity with which she can immediately start in the new occupation. Of these three, the one that matters the most is (H.1). This condition does not automatically seal off unbounded skill accumulation, because the price and wage structure also matters: the returns to skill may grow fast enough to outpace the rising marginal cost. But as we shall now see, the self-replication condition prevents such an outcome.

PROPOSITION **4.** *Suppose that within- and cross-sector human capital are accumulated via an education function satisfying H.1–H.3. Suppose, moreover, that the self-replication condition (19) is satisfied, and preferences are asymptotically homothetic. Then, if (23) holds, there is sustained per-capita income growth, and the income share of labor goes to zero.*

The Appendix contains a detailed proof; we describe the main step here. Under self-replication, each sectoral price is bounded below and above over time by strictly positive, finite numbers, just as before; see Lemma 2. But *wages* will not generally be bounded. We separate two cases.

In the first, the unit cost of some task grows; see the formal proof for precise statements regarding subsequences, etc. But then, the feasibility of automation allows us to prove that the share of human labor income in total factor bill for that task must converge to zero; see Lemma 3. The second possibility is that the unit cost of some task is bounded. Then (H.1) chokes off the incentive to acquire within-occupation productivity, given that the price of education is bounded below. The gains from such acquisition include direct wage benefits from the associated occupation, as well as cost savings on *future* investments, but these are all bounded, by our conditions on the education function. At the same time, the cost of incremental productivity climbs without bound. These observations ensure that when the task unit cost is bounded, so is productivity per person. With this boundedness result in hand, we can essentially follow the existing line of proof in Theorem 1 to obtain our previous result.

3.7. Sectors Where Full Automation is Infeasible . Our distributional analysis so far assumed that full automation is technically feasible in every sector. Certainly, our results could have been stated in weaker fashion, in the sense that automation occurs in every growing sector for which automation is technically feasible, provided the self-replication condition. But we are also interested in the *macroeconomic* distributional consequences when some sectors or occupations are "protected;" that is, $\lambda(0, r) = 0$. Examples might include "live music" or "hand-made pottery," with a human element in production by the very nature of the good. Of course, it is still possible that the ratio of human labor to robot services could become vanishingly small over time. In the live-music example, it might be possible to increase the size of the audience without bound for any live concert, and "hand-made pottery" could be judiciously redefined to include minimal human intervention. The debate is philosophical and possibly endless, as anyone who's seen *Blade Runner* or familiar with the Turing test will know.

For expositional simplicity, assume there is just a single task/occupation in each sector, and so use j to index occupations as well. Say that sector (or occupation) j is *unprotected* if $\lambda_j(0, r) > 0$ for $r > 0$, as assumed so far, and *protected* if $\lambda_j(0, r) = 0$ for all $r \geq 0$. When preferences are asymptotically homothetic, say that the asymptotic demand system $\mathbf{d}_m(\mathbf{p})$ is *elastic* if for any subset Q of sectors, $\sum_{i \in Q} p_i^n d_i(\mathbf{p}^n) \to 0$ along any sequence of prices $\{\mathbf{p}^n\}$ for which $p_i^n \to \infty$ for every $i \in Q$, while prices of all goods not in Q are bounded above.

PROPOSITION **5.** *Suppose that all intermediate goods sectors and some final goods sectors are unprotected, and that the self-replication and patience conditions hold. Then:*

(i) *Per-capita national income grows without bound.*

(ii) *For every unprotected sector on which expenditure grows, there is asymptotic automation and the output price is bounded.*

(iii) *For every protected sector on which expenditure grows, there is asymptotic automation and the output price is unbounded.*

(iv) *Suppose that preferences are asymptotically homothetic, and that the expenditure shares of all sectors converge to a limit expenditure share vector. Then the limit share of human labor in national income is bounded above by the asymptotic share of expenditure on protected sectors. Moreover, if the demand system for every type is elastic, the share of human labor in national income converges to zero.*

We omit a formal proof; much of it follows ground already covered. Part (i) follows from exactly the same argument as Theorem 1(i), which relies only on capital and robot sectors being unprotected in conjunction with the patience condition. Part (ii) is a special case of Theorem 1(ii), noting that the growth of output *value* is the same as the growth of *physical* output — prices must be bounded. Part (iii) is new. There are two cases: either the price of the protected good grows (without bound), or its physical output does. Under the former, the sectoral unit cost of the corresponding task must grow — and so too must human wages, given that robot prices are bounded (by self-replication). Asymptotic automation then follows from Lemma 3 in the Appendix.

In the latter case, with output growing, there are two possibilities: (a) The level of tasks *also* grows in that sector, but then we have asymptotic automation, given that the stock of raw labor is bounded. Moreover, since the sector is protected, the unit cost of the task must grow, and so must the price of final output. (b) The volume of tasks in the sector is bounded, but then capital must grow, but this can only be a consequence of an over-increasing task price. The latter in turn can only happen if the human wage grows, and once again we obtain asymptotic automation. Moreover, the price of the final output must grow without bound.

Exactly the same argument as in Theorem 1(iii) shows that the labor share of income generated from all unprotected sectors converges to zero. Therefore the overall labor share in national income must be asymptotically bounded above by the asymptotic share of expenditure on protected sectors. Finally, observe that prices in all unprotected sectors are bounded (Lemma 2(ii)) and all prices in growing protected sectors are unbounded. Moreover, by assumption, all protected goods are final goods. Then, with an elastic demand system, the expenditure share on all protected goods must fall to zero, and by the upper bound just established, so must the share of human labor in national income, establishing part (iv). Intuitively, protected goods are subject to “Baumol’s cost disease" and become infinitely expensive relative to unprotected goods. Because the share of household expenditures on protected goods disappears over time, the wage shelter afforded by these sectors evaporates in the long run.

3.8. Non-Homothetic Preferences. Return now to the benchmark case without protected sectors. Recall that in this case, the asymptotic homotheticity of preferences precludes a shift in demand composition that’s strong enough to adequately absorb humans into “relatively protected" occupations; that is, strong enough so that the limiting income share of labor is positive. To what degree can these implications of homotheticity be extended to more general preference profiles?

Suppose that preferences are non-homothetic, and demand persistently shifts over the space of goods with rising income. If those shifts occur precisely in favor of goods where humans are harder to displace (e.g., where θ_i defined in Section 3.4 is large), then it is possible for the long run labor share to be bounded away from zero.[12] Here is a heuristic description of the forces at play. As in the proof of Theorem 1 (see Appendix), define the share of human labor income generated in any active sector j at any date t by

$$\Psi_j(t) = \frac{\sum_{o \in O_j} w^o(t) h^o(t)}{p_j(t) y_j(t)}.$$

If $\ell(t)$ denotes the overall share of human labor in national income, it follows that

$$\ell(t) = \sum_{i=1}^{\infty} \Psi_i(t) s_i(t) + \frac{\sum_{j=e,r,k} \Psi_j(t) p_j(t) y_j(t)}{Y(t)} \tag{28}$$

at every date t, where $s_i(t)$ is the aggregate share of final goods expenditures at date t on good i, and it is understood that the sum is taken only over active sectors. Part (ii) of Theorem 1 speaks to growing automation in any active sector, and indeed we show in the Appendix that if $y_j(t) \to \infty$ in any sector, the corresponding human share $\Psi_j(t)$ converges to zero. So we can ignore the last three terms in (28): either $\Psi_j(t) \to 0$ or the sectors become insignificant as a share of (growing) national income. Everything therefore hangs on the question of whether

$$\ell(t) \simeq \sum_{i=1}^{\infty} \Psi_i(t) s_i(t) \tag{29}$$

converges to zero or not. Because $\Psi_i(t) \to 0$ for every growing sector i, and $s_i(t) \to 0$ for every non-growing sector (national income is growing), we have pointwise convergence to zero for each term in the series above, but not necessarily *uniform* convergence. The overall tension is summarized in the possibility that over time, the shares $\{s_i(t)\}$ will assign progressively greater weight to the protected sectors, leading to an asymptotically positive infinite sum even though each term in it converges to zero. Homotheticity eliminates this possibility: under it, the sequence of share vectors $\{s_i(t)\}$ has a limit which is also a share vector. Then the infinite sum must converge to zero; see Lemma 4. This is why homothetic preferences cannot allow a positive asymptotic labor share. But it also raises the question of whether some property different from or weaker than homotheticity will also suffice for the same result.

Certainly, some conditions on demand will be needed, otherwise, the following dynamic process could form an equilibrium (accompanying restrictions on primitives can be provided to generate

[12] Hubmer (2018) and Jaimovich, Rebelo and Wong (2019) provide evidence that higher incomes are associated with a shift of spending in favor of more labor intensive goods. Calibrating a neoclassical model to data from the US since the 1950s, Hubmer (2020) argues that secular trends in the labor share can be explained by the tradeoff between a rising share owing to non-homothetic preferences and a falling share owing to capital-labor substitution. Comin, Danieli and Mestieri (2019) describe non-homotheticities in demand which raise the share of services and lower that of agriculture, and are associated with rising wage polarization. They do not investigate the implications for the decline in overall labor income share. Karabarbounis and Neiman (2014) on the other hand argue that this decline is mainly intrasectoral, and not driven by changing intersectoral composition.

such an outcome). Suppose that all individuals are identical in type and initial conditions, with "modified Cobb-Douglas" preferences, so that for income y, expenditures at income y are equally divided over sectors $1, \ldots, n(y)$, where $n(y)$ is some nondecreasing step function that expands the relevant set of consumption goods as y increases. Suppose, moreover, that the common discount factor satisfies the patience condition (23). Then any equilibrium induces an aggregate expenditure share vector uniform over sectors $1, \ldots, n^*(t)$, for some nondecreasing, unbounded step function $n^*(t) = n(y(t))$. Suppose that the technology is such that $\Psi_i(t) = 0$ for $i < \left\lfloor \sqrt{n^*(t)} \right\rfloor$, but $\Psi_i(t) = a \in (0, 1)$ for $i \geq \left\lfloor \sqrt{n^*(t)} \right\rfloor$, and for some constant a (think of this as 1 minus the share of machine capital in a Cobb-Douglas production technology). That is, higher-index goods are automated later, while at the same time the consumption basket leans towards such goods. Then

$$\ell(t) \simeq \sum_{i=1}^{\infty} \Psi_i(t) s_i(t) \geq \frac{a[n^*(t) - \sqrt{n^*(t)}]}{n^*(t)} = a - \frac{a}{\sqrt{n^*(t)}} \to a \text{ as } t \to \infty.$$

Notice how expenditures spread out over goods linearly in $n^*(t)$, while automation proceeds "at the rate of $\sqrt{n^*(t)}$." So the expenditure share effect neutralizes the automation effect.

That said, the example clearly indicates that it takes quite a bit for this particular escape hatch to be pried open. If the demand share of yet-to-be-automated goods is persistent and per-capita income is growing, such sectors must also experience growing revenue. As these sectors are not yet automated, the prices of their outputs will generally rise without bound.[13] If as in Proposition 5(iv), demand is price-elastic, consumer expenditure will shares will progressively shift away from those sectors. But there is no need to go that far: even if expenditure shares are generally *uncorrelated* with automation patterns — rather than negatively correlated as just discussed — it will become impossible to prevent a vanishing labor share in the economy as a whole. It is in this sense that "preference neutrality" towards protected and unprotected sectors implies an overall inability to ward off the decline in human labor share.

Phrased in the light of (29), our results might appear to be a mathematically arcane implication of the relative speeds of convergence across double infinities (in goods and time), but actually involves an important economic issue. The potential space of goods is infinite, in the sense that the future can always bring new commodities into being where humans are (at least temporarily) not displaced by robots. And time is also infinite, resulting in an open-ended horizon where every sector is exposed to possible automation in the future. The relative speeds of the two processes determine the asymptotic labor share in the economy.

[13]The only exception to this will occur if, fortuitously along the very same sequence, the productivity of the occupational aggregates in final goods production also rises without bound.

4. Technical Progress

We extend the theory to incorporate directed technical progress. "Directedness" means that technical progress is geared to input scarcity. The key assumption we make is that the opportunities for such progress are symmetric across all inputs and sectors. This is not to deny the possibility that the very nature of science and technology might generate exogenous biases in certain directions. But studying the effect of such predetermined biases would not need a theory. If they were to favor unbridled automation, our earlier results would be a foregone conclusion. If they favored the augmentation of human quality over robots, that would raise the share of humans in national income instead.

Directed change generally points to a "balanced-growth" view of technical progress; see Acemoglu and Restrepo (2018, 2019), with antecedents that include Hicks (1932), Salter (1966), Galor and Maov (2000), and Acemoglu (1998, 2002), among many others. Acemoglu and Restrepo (2018) generate balanced growth by assuming that new tasks lie entirely in the human domain, providing temporary protection from the robot invasion. But the robots are also hard at work, automating existing tasks and perennially chasing the moving human frontier. In equilibrium, balance is achieved between these two forces. This approach, while genuinely insightful, raises many questions. Why can't new tasks that favor *robots* also appear on the frontier? Or (the flip side): why cannot technical progress allow humans to recover their edge in old tasks? And what if there is technical progress in machine capital?

In this section, we enlarge the range of possible directions of technical progress to incorporate changes in all inputs, and presume that R&D has symmetric potential in each direction. *Actual* progress will be determined by endogenous factor price dynamics. To simplify the exposition, we assume (a) a finite number of final consumer goods; (b) one task per sector, and (d) linear substitution between humans and robots within each category. Moreover, we restrict attention to equilibria with long run per-capita capital accumulation in natural units, without deriving this from underlying rates of time preference of households, and we suppose that the self-replication condition holds.[14]

Under these conditions, Theorem 2 reasserts the finding of a vanishing labor income share.

4.1. Framework. Let $\pi^F(t)$ denote the economy-wide productivity (or efficiency units per natural unit) of factor $F = k, r, h$ at date t.[15] With one occupation per sector, j indexes both occupation

[14] More precisely, we assume the self-replication condition holds at the initial date.

[15] This common productivity can be relaxed to allow sector-specific productivity improvements in each input, with positive cross-sector spillovers.

and sector, and the sector j production function at date t can be written as

$$y_j(t) = f_j(\pi^k(t)k_j, \pi^h(t)\nu_j h_j + \pi^r(t)r_j),$$

incorporating our assumption of linear substitution across humans and robots in each sector: ν_j captures the comparative productivity of humans relative to robots in sector j (this is not subsumed in the common π-terms). The same assumptions are made on f_j as before. Assume that the self-replication property for the robot sector holds at date 0; that is,

$$\pi^r(0) > \lim_{\eta \to 0} c_r(\eta, 1). \tag{30}$$

All our results extend to any competitive equilibrium in which (30) holds at some t along the equilibrium path, but we avoid an assumption on the endogenous variable $\pi^r(t)$.

4.2. R&D. At each date, R&D in each principal factor F is conducted by a short-lived F-specific inventor whose activities and returns are external to the economy in question.[16] This inventor may be the winner of a prior technological competition or race among potential inventors for factor F improvements at that date. As will become evident, our results extend to a setting where there is a single inventor who simultaneously carries out R&D across multiple factors — i.e., the ability to coordinate R&D across different directions makes no difference.

The F-specific inventor can raise the productivity of F by a factor $(1 + \rho)$ across dates t and $t + 1$, at cost $\kappa(\rho)$. Therefore

$$\pi^F(t + 1) = (1 + \rho^F(t + 1))\pi^F(t) \tag{31}$$

where $\rho^F(t + 1)$ is the rate of productivity improvement of factor F at $t + 1$. It is endogenous and lies in some compact interval $[0, \bar{\rho}]$ where $\bar{\rho} < \infty$. The cost function is strictly increasing, differentiable and convex, with $\kappa(0) = 0$ and $\kappa'(\rho)$ bounded on $[0, \bar{\rho}]$. Under our already-discussed symmetry postulate, the same cost function applies to all three inputs.

Each short-lived inventor owns property rights over the improvement, and so earns a license fee levied on all firms that make use of the improved process at $t + 1$. The fee is levied per (natural) unit of the factor employed by the firm at $t + 1$. Rights expire at the end of $t + 1$ and is freely available to all producers from $t + 2$ onwards.

Each inventor takes factor prices as given, as in the competitive innovation models of Grossman and Hart (1979) and Makowski (1980). Denote the price of factor F in sector j at $t+1$ by $\omega_j^F(t+1)$

[16] We can integrate these inventors into our economy by providing them with a technology that depends on machine capital and human/robot labor. We avoid that recursive extension here. (One difference: this sector will not be perfectly competitive, with profits constituting a positive fraction of national income.) The extent to which humans can be replaced by robots in R&D is then a determinant of the labor income share, as in other sectors.

and its corresponding employment in j by $x_j^F(t+1)$. The maximum unit license fee $L_{j,t+1}^F$ at date $t+1$ that the inventor can charge to producers in sector j is then:

$$L_j^F(t+1) = \omega_{j,t+1}^F \rho^F(t+1) \tag{32}$$

Intuitively, the "effective factor price" for licensees must rise by exactly the same rate as the proprietary productivity advance,[17] so the total fee from sector j equals $L_j^F(t+1)x_j^F(t+1) = \rho^F(t+1)E_j^F(t+1)$, where $E_{j,t+1}^F \equiv \omega_j^F(t+1)x_j^F(t+1)$ denotes the factor bill for F in sector j. Consequently, the *net* return to our inventor equals

$$\rho^F(t+1)\sum_j E_j^F(t+1) - \kappa(\rho^F(t+1)),$$

implying that optimal R&D generates an improvement rate satisfying the first order condition

$$\sum_j E_j^F(t+1) = \kappa^{'}(\rho^F(t+1)) \tag{33}$$

The same first-order condition holds even when the same inventor controls R&D in more than one factor, since the overall payoff is just the aggregate of payoffs from each factor.

4.3. Equilibrium. An equilibrium extends the definition of competitive equilibrium in Section 2.8. Because licensees transfer all surplus to the inventor, current production decisions are the same as they would have been in the absence of license purchases, but based on the technology in the public domain at the previous date. We eschew the straightforward details of this definition. Informally, an equilibrium is a sequence of wages $\{\boldsymbol{w}(t), w_r(t), w_e(t), w_k(t)\}$, prices $\{\mathbf{p}(t), p_r(t), p_e(t), p_k(t)\}$, quantities $\{F_m(t), z_m(t), e_m(t), j_m(t), k_j(t), r_j(t), h_j(t), y_j(t)\}$ for every person and every sector, and productivities $\{\pi^F(t)\}$ for factor $F = k, r, h$, such that:

(a) Given the sequence of productivities, the remaining sequence of outcomes constitutes a competitive equilibrium (i.e., all factor and product markets clear); and

(b) At every date, given equilibrium prices, all productivity changes and fees are the outcome of optimal R&D activities, as described above.[18]

4.4. Automation and the Vanishing Labor Share with Technical Progress. We now arrive at the main result of this section.

[17] One efficiency unit of the factor costs $\omega_j^F(t+1)/\pi^F(t)$ for someone without access to the improved process, and $\omega_j^F(t+1)/[(1+\rho^F(t+1))\pi^F(t)]$for someone with access. The difference in unit cost is $[\omega_j^F(t+1)\rho^F(t+1)]/[\pi^F(t)(1+\rho^F(t+1)))]$, so this can be sucked out as a license fee per efficiency unit. Multiplying by the number of efficiency units $\pi^F(t)(1+\rho^F(t+1))$ made possible by the advance, we obtain expression (32).

[18] In particular, they satisfy the first-order condition (33) with equilibrium factor bills corresponding to market clearing relative to the production functions based on public domain technology at the previous date.

THEOREM **2.** *Assume the self-replication condition* (30)*, and all other conditions stated above in this section. Then in any equilibrium which exhibits unbounded accumulation of machine capital, the income share of human labor in the economy must converge to zero as* $t \to \infty$*.*

Theorem 2 resurrects our earlier prediction, and continues to highlight the effects of asymmetry across human and physical capital accumulation. The theorem now makes a stronger assumption on growth, asking that capital be accumulated in equilibrium. It is stronger, because technical progress induces a downward drift on prices (relative to incomes), which is an “automatic" — albeit endogenous — source of real income growth. For machine capital to be willingly accumulated despite this drift, the degree of patience must clear a higher threshold (a sufficient condition will depend on the maximal rate $\bar{\rho}$ of technical progress).

The proof of Theorem 2 is intuitive enough to be provided in the main text. First observe that under market-clearing, aggregate expenditure on capital services $E^k(t)$ equals aggregate supply of machine capital $K(t)$ in natural units (since machine capital is the numeraire). Hence $K(t) \to \infty$ implies the factor bill for capital services grows without bound, and therefore the rate of productivity improvement of capital services attains the upper bound $\bar{\rho}$ after some date. Therefore $\frac{\pi^h(t)}{\pi^k(t)}$, the productivity of human relative to capital services, is bounded. The asymmetric growth in endowments in natural units between machine capital and human labor generates a bias (at least weakly) in technical progress in favor of capital.

Next, the price of robot services relative to capital services in efficiency units is bounded:

LEMMA **1.** *In any equilibrium, there exists* $B < \infty$ *such that for all* t*:*

$$\frac{\pi^k(t)}{\pi^r(t)} p_r(t) < B \tag{34}$$

The lemma is a consequence of the self-replication condition, which implies an upper bound to the equilibrium price of robots at each date, relative to the capital; this is the analogue of Proposition 1. The bound $p_r^*(t)$ satisfies

$$p_r^*(t) = c_r\left(\frac{1}{\pi^k(t)}, \frac{p_r^*(t)}{\pi^r(t)}\right) \tag{35}$$

It is easily checked that $\frac{\pi^k(t)}{\pi^r(t)} p_r^*(t)$ is decreasing in $\pi^r(t)$.[19] Since the productivity of robots can only increase over time, the upper bound is non-increasing across dates. Hence $\frac{\pi^k(0)}{\pi^r(0)} p_r^*(0)$ is an upper bound on the price of robots in efficiency units at *every* date.

[19](35) is equivalent to $p_r^*(t)\pi^k(t) = c_r(1, \frac{p_r^*(t)\pi^k(t)}{\pi^r(t)})$, so $p_r^*(t)\pi^k(t) = \psi(\pi^r(t))$ where $\psi(y)$ solves for p in the equation $1 = c_r(\frac{1}{p}, \frac{1}{y})$. Clearly ψ is non-increasing. Therefore $\frac{\pi^k(t)}{\pi^r(t)} p_r^*(t) = \frac{\psi(\pi^r(t))}{\pi^r(t)}$ is decreasing in $\pi^r(t)$.

Combining these two observations, we infer that the wage rate earned by humans must be bounded in every sector, owing to the threat of automation. In any sector j that employs human labor at any date t, humans must be cost-effective relative to robots:

$$\frac{w_j(t)}{\nu_j \pi^h(t)} \leq \frac{p_r(t)}{\pi_j^r(t)} < \frac{B}{\pi^k(t)},$$

where the second inequality follows from Lemma 1. It follows that

$$w_j(t) \leq \frac{\pi^h(t)}{\pi^k(t)} \nu_j B < \infty. \tag{36}$$

With finitely many sectors, ν_j is bounded,[20] and so, by (36), are human wages. So the national share of human labor income must converge to 0 in the long run, as $K(t) \to \infty$.

Now, while it is true that human wages are bounded, this is only relative to our chosen numeraire, which is the rental rate on machine capital in natural units. Because technical progress occurs in all sectors, machine capital becomes highly productive over time, which leads to a progressive decline in the prices of final goods, relative to the same numeraire. While human wages are bounded above in that numeraire, as just shown, they are also bounded *below*, and so by any measure of the cost-of-living — that is, relative to any index number defined on the basket of final goods — real incomes must diverge to infinity. The fact that the share of human labor share nevertheless converges to zero reveals again the contrast between absolute and relative behavior in human incomes, as discussed in earlier sections.

5. Relation to Existing Literature

Our model is distinct from most existing literature on long-run income distribution, in that it generates a novel set of long-run predictions, and its generality reveals the fundamental assumptions that respectively drive different predictions. As already noted, our model allows for multiple goods produced under diverse technologies with no substantive restrictions on them, not even convexity. Human labor could be sector-specific, or migrate across sectors via education or training, and in particular, workers can react to the threat of automation by switching to sectors where humans are harder to displace. In terms of outcomes, automation and the progressive displacement of humans occur as a consequence of capital deepening alone, even without any technical progress. Under a set of minimal and transparent sufficient conditions, the share of labor converges to zero in the long run. Moreover, balanced growth can occur when any of these conditions fail to apply.

[20] With infinitely many sectors and unbounded ν_j, demand composition will matter as in previous sections.

A possible reaction to our exercise is that it is "just" an Ak model. Certainly, our economy behaves "as if" it has an asymptotically Ak aggregate production function, which permits long run growth as in Rebelo (1990) or Jones and Manuelli (1991). However, our main interest is in the long run functional distribution between capital and labor, and in this respect it shows how the deeper disaggregated structure matters. To elaborate, note that an asymptotic Ak model can co-exist with both a positive or an ever-declining labor share.[21] Indeed, as Jones and Manuelli (1991, fn. 2) argue, *any* long run labor share between 0 and 1 can be generated with suitable parametric assumptions on the class of Ak aggregate production functions they study. Without an underlying theory of the underlying *disaggregated* economy, how it evolves with progressive automation, and a consideration of more primitive forces, it is not possible to make strong predictions regarding the asymptotic human labor share. Do all sectors eventually get automated, or just a subset? And even if the former is asymptotically true, are there not sectors that are yet to be automated at any finite date — and might wages in such sectors conceivably keep pace with capital income? Additionally, what if workers can invest in human capital to "compete" with robots? Answering these questions requires a more careful examination of micro-foundations, which constitutes the core of this paper. And as we show, the answer does not depend on technological assumptions alone: demand composition also matters.

We now discuss the literature on automation and its consequences for income distribution. In some models (such as Aghion, Jones and Jones (AJJ, 2019)), automation results from technical progress, rather than a fall in the relative price of capital goods. AJJ extend the task-based setting of Zeira (1998) where automation (occurring at an exogenous rate) is akin to an increasing capital share in an aggregate production function, resulting in a declining labor share over time. On the other hand, capital accumulation increases labor share owing to "Baumol's cost disease," i.e., inelastic capital-labor substitution. These two effects run counter to one another. Hence the long run share of labor can be positive in AJJ if the cost-disease effect outweighs the automation effect. In our setting, automation is endogenous and can occur even in the absence of technical progress, as a consequence of progressive capital deepening.[22] If the self-replication condition holds, this

[21] Endogenous growth models such as those in Romer (1986) and Alesina and Rodrik (1993) generate a positive labor share via private diminishing returns, coupled with nondecreasing society-wide returns via externalities or infrastructure. A positive human share occurs in Ak models in which human and physical capital keep pace with each other, as in Lucas (1988) or Mankiw, Romer and Weil (1992). A stable share can also arise from parametric restrictions in aggregative models with automation, as in Aghion, Jones and Jones (2019), if automation and inelastic capital-labor substitution happen to mutually neutralize their opposing effects on the labor share. On the other hand, an aggregate Ak production function can equally well generate a zero share for labor income, as in the Harrod-Domar model or its asymptotic variants.

[22] Indeed, Zeira's original model featured replacement of labor by capital owing to adoption of labor-saving technologies when capital prices are low relative to wages, just as in our model. However, the focus of Zeira (1998) was on the role of economy-wide factor endowments on per capita income (rather than income distribution) in a cross-country setting, and it did not endogenize the supply of labor-saving technologies.

induced automation effect is powerful enough to drive the long run labor share to zero, *even despite* inelastic capital labor substitution in all sectors.

Acemoglu and Restrepo (AR, 2018) also extend Zeira's approach to study the distributional implications of automation.[23] Their model has one final good, produced by a continuum of tasks, each of which is produced either by robots or humans. There is a task threshold above which tasks can only be performed by humans. Technical progress enlarges the set of tasks that lie above this threshold, and so is effectively restricted to be in favor of humans (in contrast to AJJ). Below the threshold, robots can substitute for humans depending on relative factor prices; hence capital accumulation tends to lower labor share in AR (again in contrast to AJJ). As in AJJ, a long-run positive share emerges, but for opposite reasons.

As in AR, our model generates automation and a declining labor share as a consequence of capital deepening. However the greater generality of our model reveals the underlying microfoundations for this result. First, our use of multiple sectors shows how self-replication in the robot sector spills over to all sectors, an issue that does not arise in an aggregative setting. Second, we make explicit the role of occupational diversity in sustaining human capital accumulation. In so doing (and again by invoking a multiplicity of sectors) we show the fundamental role played by the composition of demand; specifically, the asymptotic homotheticity of preferences as individual income climbs. Finally, our formulation of technical progress allows for capital as well as human productivity improvements, and does so in an ex ante unbiased fashion. The direction of technical progress is then driven by endogenous innovation incentives. The extension of our baseline model in Section 4 provides an illustration of plausible circumstances where technical progress ends up not being directed in favor of humans.

Benzell et al (2018) present a model with two final consumption goods. “Robots" or “code" represent a durable capital asset used to produce a material good, while labor is used to produce a useful service. An increase in the stock of robots causes the relative price of the service and hence wages to fall, owing to induced changes in demand composition. This paper therefore shares some common features with ours: automation is induced by changes in factor prices and demand composition matters. However in their model, in the long run there is no growth and the share of labor is positive. These differences owe to their assumption that there is no scope for robots to displace humans in the production of services, and a different model of savings (an OLG specification without parental altruism) that prevents any long run growth.

[23] Hemous and Olsen (2020) extend Acemoglu and Restrepo (2018) to incorporate skilled and unskilled labor, and focus on the implications of automation for wage inequality, an issue we ignore.

Caselli and Manning (2019) study the consequences of automation on levels of real wages, rather than inequality of factor share between labor and capital. They compare steady states of a model with multiple final/intermediate good and types of labor with respect to an exogenous change in technology that lowers unit costs of all goods at any given vector of factor prices. They show this implies that real wages of at least one worker type must increase. Moreover the average real wage increases if the prices of investment goods fall faster than of consumption goods. *A fortiori*, these results hold also in our model, which additionally incorporates endogenous capital accumulation and technical progress. As shown in Proposition 2, automation and vanishing labor share can co-exist with real wages that rise without bound.

While our model has focused on predictions of long run labor share, it also provides a potential explanation of falling labor share which is distinct from other explanations in the existing literature. Aside from those based on automation and already discussed above, the remaining literature can be classified into the following two categories:[24]

(i) An argument based on *sustained human capital investment*, which causes effective labor to grow relative to effective capital. In Grossman et al (2020), human capital investments rise owing to a fall in the interest rate, driven in turn by an exogenous decline in rates of technical progress. A number of additional assumptions are needed for their result: capital-skill complementarity, a low intertemporal substitution elasticity in consumption, a closed economy, and an aggregate capital-labor elasticity of substitution below 1.

(ii) Theories based on *globalization, rising markups, rising market concentration, or fall in labor bargaining power*. Arguments include globalization, whereby labor in developed countries are displaced by competing cheap imports (Autor, Dorn and Hansen 2016), selection into more profitable, higher-markup firms (Autor et al 2017), or factors such as the rise of the gig economy or greater product differentiation, leading to a decline in firm competition and the bargaining power of labor (Neary 2003, Gutiérrez and Philippon 2017, Azar and Vives 2018, Eggertsson, Robbins, and Wold 2018, and Kaplan and Zoch 2020).

Our approach is distinct both in terms of underlying assumptions and detailed predictions. The relative growth of human capital and physical capital in efficiency units is inverted relative to Grossman et al (2020). While our model can be extended to incorporate market power of firms, our results would continue to apply in the absence of any changes in market power; moreover, the evidence presented by Karabarbounis and Neiman (2014) indicates the relevance of our approach, even if rising markups provide part of the explanation of falling labor share.

[24]We exclude explanations based on *sustained capital accumulation* in an aggregative model with capital-labor substitution elasticities exceeding one (e.g., Piketty 2014), because these are at odds with evidence from industry panel studies which show inelastic substitution in most industries (Chirinko and Mallick 2014).

6. Concluding Remarks

We study the possibility of long-term automation and decline in the labor share, driven by capital accumulation rather than biased technical progress or rising markups. Our argument relies on a fundamental asymmetry across physical and human capital in modern economies. While physical capital can be scaled up for the same activity and accumulates in natural units, human capital accumulates via education that alters choice into higher-skilled occupations, but — from the vantage point of a household or individual — cannot scale up the quantity of labor for a *given* occupation to an unlimited degree. Under a self-replication condition on the technology of the robot-producing sector, and some additional conditions made explicit in the paper, we show that the share of human labor in national income must dwindle to zero in the long run.

The self-replication condition plays an important role in the model. Though involving the technology of the robot sector *alone*, it turns out to have far reaching implications for long run growth and functional inequality. There is increasing recognition that the “production of robots by means of robots" is not merely a hypothetical possibility:

> “They are a dream of researchers but perhaps a nightmare for highly skilled computer programmers: artificially intelligent machines that can build other artificially intelligent machines . . . Jeff Dean, one of Google’s leading engineers, spotlighted a Google project called AutoML . . . [which] is a machine-learning algorithm that learns to build other machine-learning algorithms. With it, Google may soon find a way to create A.I. technology that can partly take the humans out of building the A.I. systems that many believe are the future of the technology industry." (*The New York Times*, November 5, 2017.)

The model therefore suggests that the implications of recent developments in AI for the future of inequality may well be fundamentally different from anything observed in the past.

On the other hand, our paper also provides a number of different reasons why the labor share need not vanish asymptotically: if the self-replication does not hold, non-homothetic demand that progressively favors sectors where humans are harder to displace, the existence of growing sectors where humans cannot be displaced at all, or technical progress biased in favor of humans. However, while any of these scenarios is possible, we do not see any reason why they should be inevitable. Our main purpose has been to identify, as clearly as possible, a set of minimal sufficient conditions for a zero asymptotic labor share. And that at the same time, automation can help generate growth in the long run, fueling an absolute increase in human wages even as it causes a relative decline in labor share.

Our emphasis throughout has been on the *functional* distribution of income. Whether a household’s income manages to keep step with the rest of the economy — the question of the personal distribution of income — will depend on whether they invest in financial wealth or human capital (or neither, or both). This is a question we have not yet addressed, though our model provides the means to study it, and is something we plan to undertake. It will become necessary to take closer account of both the heterogeneity of the population in their preference parameters, as well as to incorporate a detailed description of credit market constraints. Both these features are currently present in the model, but play no more than a background role. Finally, we note that despite its generality, the theory presented here is simple and tractable, which may also allow it to be useful in analyzing effects of fiscal policies such as capital taxes, education subsidies, universal basic income or other policy interventions to address the distributional consequences of automation.

References

D. Acemoglu and J. Robinson (2015), “The Rise and Decline of General Laws of Capitalism," Journal of Economic Perspectives, 29(1), 3-28.

D. Acemoglu and P. Restrepo (2018), “The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares and Employment," American Economic Review, 108(6): 1488-1542.

——————————— (2019), “Automation and New Tasks: How Technology Displaces and Reinstates Labor,” Journal of Economic Perspectives, 33(2): 3-30.

P. Aghion, B. Jones, and C. Jones (2019), “Artificial Intelligence and Economic Growth," in Ajay Agrawal, Joshua Gans and Avi Goldfarb (eds), *The Economics of Artificial Intelligence: An Agenda*, National Bureau of Economic Research.

K.J. Arrow (1951), “Alternative Proof of the Substitution Theorem for Leontief Models in the General Case". In Activity Analysis of Production and Allocation, ed. T.C. Koopmans. New York: John Wiley.

D. Autor, D. Dorn, L.F. Katz, C. Patterson and J. van Reenen (2017), “Concentrating on the Fall of the Labor Share," American Economic Review Papers and Proceedings, 107(5), 180-185.

J. Azar and X. Vives (2018), “Oligopoly, Macroeconomic Performance, and Competition Policy," mimeo., IISE Business School, Barcelona.

A. Banerjee and A. Newman (1993), “Occupational Choice and the Process of Development." *Journal of Political Economy* **101**, 274–298.

S. Benzell, L. Kotlikoff, G. LaGarda and J. Sachs (2018). “Robots Are Us: Some Economics of Human Replacement," NBER Working Paper 20941.

J.I. Bernstein and M.I. Nadiri (1988), “Interindustry R&D Spillovers, Rates of Return, and Production In High-Tech Industries," *American Economic Review - AEA Papers and Proceedings*, 78, No. 2, pp. 429-434, May 1988.

N. Bloom, Schankerman, M., and van Reenen, J. (2013). “Identifying Technology Spillovers and Product Market Rivalry". *Econometrica*, 81(4):1347-1393.

R. Chirinko and D. Mallick (2014), “The Substitution Elasticity, Factor Shares, Long Run Growth and the Low Frequency Panel Model," Working Paper, University of Illinois at Chicago.

D. Comin, A. Danieli and M. Mestieri (2019), “Income-Driven Labor Market Polarization," Working Paper.

G. Eggertsson, J. Robbins, and E.G. Wold (2018), “Kaldor and Piketty’s Facts: The Rise of Monopoly Power in the United States," mimeo.

O. Galor and J. Zeira (1993), “Income Distribution and Macroeconomics." *Review of Economic Studies* **60**, 35–52.

G. Grossman, E. Helpman, E. Oberfield and T. Sampson (2020), “Endogenous Education and Long-Run Factor Shares," working paper, Princeton University.

S. Grossman and O. Hart (1979), “A Theory of Competitive Equilibrium in Stock Market Economies," *Econometrica* **47**, 293–330.

G. Gutiérrez and T. Philippon (2017), ““Declining Competition and Investment in the U.S," mimeo., NYU Stern.

D. Hemous and M. Olsen (2020). “The Rise of the Machines: Automation, Horizontal Innovation and Income Inequality," forthcoming, American Economic Journal Macro.

J. Hubmer (2018), “The Race Between Preferences and Technology," working paper, Yale University Department of Economics.

N. Jaimovich, S Rebelo and A. Wong (2019), “Trading Down and the Business Cycle," *Journal of Monetary Economics*, 102, 96-121.

W. Johnson (2018), “Economic Growth and the Evolution of Comparative Advantage in an Occupation-Based Network of Industries," PhD dissertation, Department of Economics, Boston University.

G. Kaplan and P. Zoch (2020), “Markups, Labor Market Inequality and the Nature of Work," working paper, University of Chicago.

L. Karabarbounis and B. Neiman (2014), “The Global Decline of the Labor Share," *Quarterly Journal of Economics*, 129(1): 61-103.

L. Ljungqvist (1993), “Economic Underdevelopment: The Case of Missing Market for Human Capital," *Journal of Development Economics* **40**, 219–239.

L. Makowski (1980), “Perfect Competition, the Profit Criterion, and the Organization of Economic Activity," *Journal of Economic Theory* **22**, 222–42.

G. Mankiw (2015), “Yes $r > g$. So What?" *American Economic Review* (Papers and Proceedings) **105**, 43–47.

D. Mookherjee and D. Ray (2002), “Is Equality Stable?," *American Economic Review* **92** (Papers and Proceedings), 253–259.

D. Mookherjee and D. Ray (2003), “Persistent Inequality," *Review of Economic Studies* **70**, 369–394.

D. Mookherjee and D. Ray (2010), “Inequality and Markets: Some Implications of Occupational Diversity," *American Economic Journal: Microeconomics* **2**, 38–76.

J. P. Neary (2003), “Globalization and Market Structure," *Journal of the European Economic Association* **1**, 245–271.

T. Piketty (2014). *Capital in the Twenty First Century*. Belknap Press.

D. Ray (2015), “Nit-Piketty: A Comment on Thomas Piketty’s Capital in the Twenty First Century," *CESifo Forum* **16** (1), 19–25.

P. A. Samuelson (1951) “Abstract of a Theorem concerning Substitutability in Open Leontief Models". in *Activity Analysis of Production and Allocation*, ed. T.C. Koopmans. New York: John Wiley.

S. Santens (2016), “Robots Will Take Your Job," *The Boston Globe*, February 24.

H. Uzawa (1961-2), “On a Two-Sector Model of Economic Growth," *Review of Economic Studies* **29**, 40–47.

J. Zeira (1998), “Workers, Machines and Economic Growth," *Quarterly Journal of Economics*, 113(4): 1091-117.

Appendix: Proofs

Proof of Theorem 1. We begin with some preliminary observations.

LEMMA **2.** *For each j, there is $\underline{p}_j > 0$ such that in any equilibrium and at any date t,*

$$p_j(t) \geq \underline{p}_j > 0 \tag{37}$$

whenever $y_j(t) > 0$. If in addition, self-replication holds, then

$$p_j(t) \leq c_j\left(1, \{(\nu^o)^{-1} \sup P^*\}\right) < \infty \tag{38}$$

for all j and at every date t, where P^* is defined in Proposition 1 and o ranges over O_j.

Proof. See Supplementary Appendix. ■

Recall the automation index introduced in the main text (in a generic occupation o, sector subscript removed):

$$a^o(\zeta) \equiv \min_{(r^o,h^o)} \left\{ \frac{r^o}{h^o\zeta + r^o} \middle| (r^o, h^o) \text{ minimizes unit cost } (h^o\zeta + r^o) \right\}.$$

LEMMA **3.** *The automation index $a^o(\zeta) \to 1$ as $\zeta \to \infty$.*

Proof. Let $h^o(\zeta)$ be any selection from the set of unit cost minimizing choices of human labor at price ratio ζ. We claim that $\lim_{\zeta\to\infty} \zeta h^o(\zeta) = 0$. To prove this, pick any sequence where $\lim_{\zeta\to\infty} \zeta h^o(\zeta)$ is well-defined (possibly infinite), and a further subsequence (retain notation) so that the corresponding robot choice $r^o(\zeta)$ converges to some limit, call it r^*, as $\zeta \to \infty$. Because minimized unit cost cannot exceed that from the feasible method of producing one unit of λ using ν^o units of r^o alone, we have $h^o(\zeta) \to 0$. Because λ^o is continuous,

$$\lambda^o(h^o(\zeta), r^o(\zeta)) = \lambda^o(0, r^*) = 1.$$

Therefore r^* is always a feasible choice for unit production, and so

$$\zeta h^o(\zeta) + r^o(\zeta) \leq r^*,$$

Because $r^0(\zeta) \to r^*$, the claim follows. Finally, unit production of λ^o is maintained through the sequence, so $\liminf_\zeta r(\zeta) > 0$. It follows immediately that $a^o(\zeta) \to 1$ as $\zeta \to \infty$. ■

Now we prove part (i) of the Theorem. We first show that in an equilibrium and for any group m that satisfies (23), we have $z_m(t) \to \infty$. Consider the indirect utility functions $v_m(z(t), \mathbf{p}(t))$ for individual expenditure $z(t)$ at any date t. In any equilibrium, an individual in this group has

$F_0 > 0$ units of a financial asset at date 0, and thereafter makes educational and financial asset choices (and consumption choices), under fully anticipated prices, which includes a sequence of return factors $\{\gamma(t)\}$ on financial holdings. She has several necessary conditions that describe her behavior, but one set of these has to do with her choice of financial assets. Because her initial income can be strictly positive if she so pleases (there is a positive subsistence wage), her current expenditure $z_m(T)$ must be strictly positive at some date T, but then $z_m(t) > 0$ for all $t \geq T$, by the unbounded steepness of v_m in z at 0. For ease in writing set $T = 0$. It follows that the Euler equation on financial assets must hold with a particular inequality at every date $t \geq 0$:

$$v'_m(z(t), \mathbf{p}(t)) \geq \beta_m \gamma(t) v'_m(z(t+1), \mathbf{p}(t+1)). \tag{39}$$

If (39) fails, she could always transfer resources one period into the future and increase lifetime utility. (Equality may not hold because human capital could have a higher rate of return than financial assets, and the individual may not be able to marginally pull back funds from future to present, because of credit constraints.) Now we compound this Euler inequality just as in the main text to arrive at (26), reproduced here as:

$$v'_m(z_m(0), \mathbf{p}(0)) \geq \frac{\beta_m^{t-1}\left[(1-\delta) + \frac{1}{c_k(1,\{(\nu^o)^{-1}\sup P^*\})}\right]^{t-2}}{c_k\left(1, \{(\nu^o)^{-1}\sup P^*\}\right)} v'_m(z_m(t), \mathbf{p}(t)). \tag{40}$$

It follows from condition (23) that $v'_m(z_m(t), \mathbf{p}(t)) \to 0$ as $t \to \infty$. But v_m is strictly increasing and concave for every $\mathbf{p}$. Moreover, every active final goods price is bounded above and below by (38) of Lemma 2.[25] Therefore (40) can *only* hold if $z_m(t) \to \infty$ as $t \to \infty$. With a bounded credit limit on every other individual, we must conclude that per-capita income $Y(t)$ as defined in (17) must go to infinity.

For part (ii), we show that any sector j must have its automation index converge to 1 along any subsequence in which its output grows. To show this, we argue first that inputs from some occupation $o \in O_j$ in that sector must also grow. If this were false for every occupation in O_j, then $k_j(\tau) \to \infty$ as $\tau \to \infty$, and so by the unbounded steepness condition,

$$q^o(\tau) = p_j(\tau) \frac{\partial}{\partial \lambda^o} f_j(k_j(\tau), \boldsymbol{\lambda}_j(\tau)) \to \infty \text{ as } \tau \to \infty \tag{41}$$

for *some* occupation $o \in O_j$. But we know that

$$q^o(\tau) = c^o(w^o(\tau), p_r(\tau)) \leq (\nu^o)^{-1} p_r(\tau) \leq (\nu^o)^{-1} \sup P^* < \infty,$$

where the first inequality comes from the fact that automation is feasible and the second from self-replication and Proposition 1. But that contradicts (41). So $\lambda^o(\tau)$ must grow in some occupation $o \in O_j$. In any such occupation, $h^o(\tau) \leq 1$, so $r^o(\tau) \to \infty$. If $w^o(\tau)$ is bounded along some

[25] If a final good is inactive it has no effect on v_m anyway, as it is not consumed.

subsequence, then the automation index must converge to 1 along that subsequence. If $w^o(\tau)$ is unbounded along some subsequence, then — recalling that $p_r(\tau)$ is bounded — Lemma 3 applies and the automation index for occupation o also converges to 1 along that subsequence. Averaging the index over all growing occupations in sector j completes the proof.

Part (iii). For this part, we need the following

LEMMA **4.** *Let S be the set of all infinite-dimensional nonnegative vectors $\mathbf{s} \equiv (s_1, s_2, \ldots)$, with components in $[0,1]$ and $\sum_{j=1}^{\infty} s_j = 1$. Let $\mathbf{s}(t)$ be a sequence in S, and suppose that there is $\hat{\mathbf{s}} \in S$ such that $\mathbf{s}(t)$ converges pointwise to $\hat{\mathbf{s}} = (\hat{s}_j)$. Let $\mathbf{\Psi}(t)$ be a corresponding convergent sequence with components $(\Psi_1(t), \Psi_2(t), \ldots)$, where $\Psi_j(t) \in [0,1]$ for every j and t, with $\Psi_j(t) \to 0$ as $t \to \infty$ for every j with $\hat{s}_j > 0$. Then $\lim_{t\to\infty} \sum_{j=1}^{\infty} \Psi_j(t)\hat{s}_j(t) = 0$.*

Proof. See Supplementary Appendix. ∎

For any active sector j and date t, define

$$\Psi_j(t) = \frac{\sum_{o \in O_j} w^o(t)h^o(t)}{p_j(t)y_j(t)} \in [0,1],$$

and set $\Psi_j(t) = 0$ if $y_j(t) = 0$. This is well-defined: $p_j(t) > 0$ whenever $y_j(t) > 0$ (Lemma 2). We claim that if $y_j(t) \to \infty$ along a subsequence of dates, $\Psi_j(t) \to 0$. To see this, pick any limit point of $\Psi_j(t)$ along the subsequence in question. Choose further subsequences such that for every occupation $o \in O_j$, $w^o(t)$ is either bounded or diverges to infinity; retain the original index t. Now, if $w^o(t)$ is bounded for some $o \in O_j$, then certainly

$$\frac{w^o(t)h^o(t)}{p_j(t)y_j(t)} \to 0$$

as $t \to \infty$. (Because $p_j(t)$ is bounded below, $p_j(t)y_j(t) \to \infty$.) Otherwise, if $w^o(t) \to \infty$ for some $o \in O_j$, $\zeta^o(t) = w^o(t)/p_r(t) \to \infty$, given that $p_r(t)$ is bounded above (Proposition 1). By linear homogeneity of λ^o and Lemma 3,

$$\frac{w^o(t)h^o(t)}{p_j(t)y_j(t)} \leq \frac{w^o(t)h^o(t)}{w^o(t)h^o(t) + p_r(t)r^o(t)} = \frac{\zeta^o(t)h^o(t)}{\zeta^o(t)h^o(t) + r^o(t)} \leq 1 - a^o(\zeta^o(t)) \to 0.$$

Aggregating these observations over all the occupations proves the claim.

If $\ell(t)$ denotes the share of human labor in national income, it follows that

$$\ell(t) = \frac{\sum_{o \in O_j} w^o(t)h^o(t)}{Y(t)} = \frac{\sum_j \Psi_j(t)p_j(t)y_j(t)}{Y(t)}$$

$$= \left[\frac{\sum_{i=1}^{\infty} \Psi_i(t)p_i(t)y_i(t)}{Y(t)}\right] + \frac{\sum_{j=e,r,k} \Psi_j(t)p_j(t)y_j(t)}{Y(t)} \tag{42}$$

at every date t, where it is understood that any sector inactive at any date has an entry of 0 in the sum above. Write for every final good i active at date t:

$$\frac{p_i(t)y_i(t)}{Y(t)} = \sum_m \phi_m(t)s_{mi}(t), \tag{43}$$

where $\phi_m(t) \equiv Z_m(t)/Y(t)$ is the ratio of current aggregate *expenditure* of type m to total *income*, and $s_{mi}(t)$ is the corresponding expenditure share on good i by type m. Combining (42) and (43),

$$\ell(t) = \sum_{i=1}^{\infty} \Psi_i(t) \left[\sum_m \phi_m(t)s_{mi}(t)\right] + \frac{\sum_{j=e,r,k} \Psi_j(t)p_j(t)y_j(t)}{Y(t)} \tag{44}$$

for all t. We will show that the right hand side of (44) converges to 0 as $t \to \infty$. To this end, pick any subsequence of dates (but retain original notation) so that $\ell(t)$ converges. Exploiting the fact that the number of sectors is countable, use a diagonal argument to extract a further subsequence (again retain notation) so that *each* of the bounded sequences $\Psi_j(t)$, $\phi_m(t)$, $s_{mi}(t)$, $p_j(t)$, and $[p_j(t)y_j(t)]/Y(t)$ also converge.[26] The last finite sum in (44) pertains only to three sectors: e, r and k. For any of these sectors, call it j, $\Psi_j(t) \to 0$ along any subsequence for which j is consequential, and on any other subsequence $p_j(t)y_j(t)$ must be bounded, while $\Psi_j(t) \in [0, 1]$. Putting these observations together with $Y(t) \to \infty$, we must conclude that this last finite term in (44) converges to 0. The rest of the argument concerns the first set of terms in (44).

Let M be the set of all indices for which $\lim_t \phi_m(t) > 0$ for the subsequence under consideration. If M is empty, we are done, so assume it is nonempty. Then, using the fact that the interchange of a finite and infinite sum is always valid, we have

$$\begin{aligned}\sum_{i=1}^{\infty} \Psi_i(t) \left[\sum_m \phi_m(t)s_{mi}(t)\right] &= \sum_m \phi_m(t) \left[\sum_{i=1}^{\infty} \Psi_i(t)s_{mi}(t)\right] \\ &= \sum_{m\in M} \phi_m(t) \left[\sum_{i=1}^{\infty} \Psi_i(t)s_{mi}(t)\right] + \sum_{m\notin M} \phi_m(t) \left[\sum_{i=1}^{\infty} \Psi_i(t)s_{mi}(t)\right].\end{aligned} \tag{45}$$

Because $\phi_m(t) \to 0$ for all $m \notin M$, the second term on the right hand side of this equation converges to 0. It remains to show that same is true of the first term. It will suffice to show that for each $m \in M$,

$$\sum_{i=1}^{\infty} \Psi_i(t)s_{mi}(t) \to 0 \tag{46}$$

as $t \to \infty$ along our chosen subsequence. Because $\lim_t \phi_m(t) > 0$ for $m \in M$ and $Y(t) \to \infty$, it follows that expenditures diverge to infinity for a positive measure of individuals of each type m. Let $Z_m(t)$ be the aggregate expenditure of type m and $x_{mi}(t)$ the aggregate demand for good i by

[26] In particular, the ratio $\phi_m(t) = Z_m(t)/Y(t)$ is also bounded because of finite credit limits.

this type. By asymptotic homotheticity,

$$\hat{s}_{mi} \equiv \lim_t s_{mi}(t) = \lim_t \frac{p_i(t)x_{mi}(t)}{Z_m(t)} = \lim_t p_i(t)d_i^m(\mathbf{p}(t)),$$

We claim that each $p_i(t)$ is bounded above and below by strictly positive numbers. The upper bound is given by Lemma 2. For the lower bound, suppose by contradiction that I, the set of indices such that $p_j(t) \to 0$, is nonempty. Then, by assumption (ii) on the function d_m, we have $\liminf_t d_{mi}(\mathbf{p}(t)) > 0$ for some $i \in I$. But then that sector is active at all large dates, which means that its price is bounded below (see (37) of Lemma 2), a contradiction. Therefore the claim is true, and given assumption (i) on $\mathbf{d}_m$, it follows that $\hat{s}_{mi}$ forms a "bonafide share vector" with $\sum_i \hat{s}_{mi} = 1$. So the conditions in Lemma 4 are satisfied (ignore index m). Therefore this Lemma implies (46), and the income share of human labor must converge to zero. Recall (17) to write out income:

$$Y = \sum_i p_i y_i + p_e y_e + p_k y_k,$$

and express it as the sum of (machine) capital and human income:

$$\begin{aligned} Y &= \sum_i p_i y_i + p_e y_e + p_k y_k = \sum_{j \neq r} [k_j + p_r r_j + w_j h_j] \\ &= \sum_{j \neq r} [k_j + w_j h_j] + p_r [y_r - r_r] = \sum_{j \neq r} [k_j + w_j h_j] + [k_r + w_r h_r] \\ &= \sum_j [k_j + w_j h_j] = K + \sum_j w_j h_j. \end{aligned}$$

That means that the income share of capital converges to 1.

Proof of Proposition 2. By Theorem 1, per-capita income must grow without bound. Moreover, by Lemma 2, $0 < \inf_t p_j(t) \leq \sup_t p_j(t) < \infty$. Therefore by asymptotic homotheticity and the full support restriction on d_m for each m, every final goods sector must grow without bound. By the unbounded steepness of each sectoral production function in its inputs, the demand for every final goods occupation must also grow, and so each such occupation must be asymptotically automated. As a consequence, all occupations in the capital and robot sectors must also experience unbounded growth, and they too must be asymptotically automated.

Part (a). Consider any occupation $o \in O_{-e}$ and any subsequence of dates t with $h^o(t) > 0$ for all such t. Then, by the first-order necessary conditions for optimality,

$$(47) \qquad w^o(t) \leq \left[\frac{\partial \lambda^o / \partial r}{\partial \lambda^o / \partial h}(h^o(t), r^o(t))\right]^{-1} p_r(t) \leq \left[\frac{\partial \lambda^o / \partial r}{\partial \lambda^o / \partial h}(h^o(t), r^o(t))\right]^{-1} \sup P^*,$$

where the second inequality uses Proposition 1. Because o is asymptotically automated,

$$\frac{\partial\lambda^o/\partial r}{\partial\lambda^o/\partial h}(h^o(t), r^o(t)) \to \theta^0 \text{ as } t \to \infty. \tag{48}$$

Combining (47) and (48), we must conclude that

$$\lim \sup_{t:h^o(t)>0} w^o(t) \leq \frac{\sup P^*}{\theta^o}, \tag{49}$$

so that

$$\sup_{o\in O_{-e}} \left[\lim \sup_{t:h^o(t)>0} w^o(t)\right] \leq \sup_{o\in Q} \frac{\sup P^*}{\theta^o} = \frac{\sup P^*}{\inf_{o\in O_{-e}} \theta^o} < \infty. \tag{50}$$

We now consider the education sector, and claim that $\lim \sup_{o\in O_e,t} w^o(t) < \infty$. For if not, then $w^o(t) \to \infty$ for some subset of occupations $o \in O_e$ along some subsequence of dates. Now, $\sup_t p_e(t) < \infty$ by Lemma 2, and moreover, the education function is uniformly bounded. So the cost of all education is uniformly bounded. Therefore, given (50), all humans must ultimately be in these educational occupations with unboundedly rising wages, along the aforementioned subsequence of dates. Therefore the total input cost of providing education is unbounded along the same subsequence, so total revenue from education must also be unbounded. Using $\sup_t p_e(t) < \infty$ again, the total *output* of education must grow without bound along the same subsequence of dates. But now we have a contradiction, for only a bounded amount of education is produced at any date (the education function is uniformly bounded and there is a unit measure of humans). So the claim is true, and $\lim \sup_{o\in O_e,t} w^o(t) < \infty$.

With this result in hand, we can finally extend the bound (50) to include *all* occupations and dates, not just those for which $h^o(t) > 0$. For if $w^o(t)$ were to climb to infinity along some subsequence of occupations and dates for which $h^o(t) = 0$, then at large t, individuals must move to these sectors, which contradicts the presumption that $h^o(t) = 0$.

Therefore $\lim \sup_{o\in O,t} w^o(t) < \infty$, which establishes part (a).

Part (b). Recall that all occupations $o \in O_{-e}$ grow along with output, and so are asymptotically automated. Because $r^o(t) > 0$ for all large t, it must be the case that for such t,

$$w^o(t) \geq \left[\frac{\partial\lambda^o/\partial r}{\partial\lambda^o/\partial h}(h^o(t), r^o(t))\right]^{-1} p_r(t).$$

Passing to the limit with asymptotic automation, using (48), and invoking Lemma 2 to find a strictly positive lower bound $\underline{p}_r$ for the robot price, we conclude that

$$\lim \inf_t w^o(t) \geq \frac{\underline{p}_r}{\theta^o}.$$

for every occupation $o \in O_{-e}$. It follows that

$$\sup_{q \in O_{-e}} \liminf_{t} w^o(t) \geq \frac{\underline{p}_r}{\inf_{o \in O_{-e}} \theta^o} = \infty,$$

which implies that there is a sequence of occupations and dates along which the human wage climbs without bound. To complete the proof, we claim that every human wage $w_\iota(t)$ must climb to infinity as well. Suppose not, then $w_\iota(t)$ is bounded along a subsequence of t, while there is some occupation o with $w^o(t) \to \infty$ along that same subsequence. But the education function is uniformly bounded, and so is the price of a unit of education $p_e(t)$. So at some large t, person ι can profitably deviate by selecting occupation o for one period, and returning to her presumed optimal plan from date $t+1$ onwards, a contradiction. ■

Proof of Proposition 3. By the minimum subsistence bound on wages and (37) of Lemma 2, there is $q_- > 0$ such that in any equilibrium, $q(t) \geq q_-$ for all t. Recalling the definition of $\Lambda_j(q)$, we can easily use the linear homogeneity of f_j and invoke (27) to see that there is $\epsilon > 0$ such that the share of task costs in total cost in sector j satisfies:

$$\frac{q_j(t)\lambda_j(t)}{p_j(t)y_j(t)} = \Lambda_j(q(t)) \geq \epsilon > 0,$$

for every t and every active sector j. Therefore, if $\Lambda(t)$ denotes the overall share of task costs in total production costs at date t, then, because it is simply a convex combination of all the sector-specific shares,

$$\Lambda_j(t) \geq \epsilon > 0 \tag{51}$$

as well, for every date t.

Now consider any sequence of dates (retain original index t) along which the overall income share of human labor converges. Using a diagonal argument, extract a subsequence such that in every sector j, $\Lambda_j(t)$ converges — to a strictly positive limit, by (51), and the overall shares of human labor cost and robot cost in sectoral task costs converges as well. If the robot cost share converges to a number strictly smaller than one, then the proof is complete. Otherwise, the robot cost share converges to 1, and given that the latter has a positive limit, it follows that $\lim_j r_j(t) > 0$. In particular, for large dates, the robot sector is active, so that:

$$p_r(t) = c_r(1, q_r(t)) \leq c_r(1, \nu_r^{-1} p_r(t)). \tag{52}$$

where the latter inequality comes from the feasibility of automation in the robot sector.

Now, self-replication fails by assumption, so $\lim_{\eta \to 0} c_r(\eta, 1) \geq \nu_r$. Multiplying through by $p_r \nu_r^{-1}$, and using the concavity of the robot cost function (the first part of our regularity condition on f_r),

$p_r \leq c(1, \nu_r^{-1} p_r)$ for every $p_r > 0$. Indeed, using (37) of Lemma 2 and the unbounded steepness of c at $p_r = 0$ (inherited in turn from the unbounded steepness of f_r), we make a stronger claim: there is $\epsilon > 0$ such that

$$p_r(t) \leq c_r(1, \nu_r^{-1} p_r(t)) - \epsilon. \tag{53}$$

at every conceivable *equilibrum* price $p_r(t)$ at any date.[27] Combining (52) and (53),

$$c_r(1, q_r(t)) \leq c_r(1, \nu_r^{-1} p_r(t)) - \epsilon,$$

which in turn implies the existence of $\epsilon' > 0$ such that

$$q_r(t) \leq \nu_r^{-1} p_r(t) - \epsilon'$$

for all t large. So, because the unit task cost is bounded away from what it would have been with full automation, it follows that $a_r(t) = r_r(t)/h_r(t)$ must be bounded above. But then, because asks in the robot sector can be produced by humans alone (the second part of our regularity condition on f_r), it must be that the share of human labor income in the total value of robot production (equal to robot income) is bounded away from 0. Therefore in this case, too, the share of human income in task cost is bounded away from zero, and the proof of the proposition is complete. ■

Proof of Proposition 4. In any equilibrium, all prices are bounded below (pointwise) by strictly positive numbers, just as before; see (37) of Lemma 2. Under self-replication, Proposition 1 additionally applies and robot prices are also bounded above exactly as before, and *independent* of human productivity. In turn, this provides pointwise upper bounds on prices in all sectors, see (38) of Lemma 2. That includes the same bound on price of capital, so part (i) of Theorem 1 holds under the same conditions and following exactly the same proof.

The remainder of the proof consists in applying the following argument at more than one point:

Claim. Suppose that for some occupation $o \in \cup_j O_j$, the human wage per unit of productivity, $w^o(t)$, is bounded on the equilibrium path by some $\bar{w}^o < \infty$. Then human labor in efficiency units is also bounded along that same path.

To establish the Claim, pick some $S > 0$ such that

$$\underline{p}_e S > \frac{\bar{w}^o}{1 - \beta} + \bar{p}_e L^o \tag{54}$$

where β is the largest discount factor among all types. Next, using (H.1), pick $M < \infty$, *larger than initial productivity endowment and the cross-occupation bound*, such that $e(\mu, \mu + \Delta, o, o) \geq S\Delta$

[27] Note first that $p_r(t)$ is bounded below (Lemma 2. Now consult Panel B, Figure 1. Because $c_r(1, p_r)$ is concave and initially lies strictly above the diagonal, it cannot converge back to the diagonal without actually crossing it. So it must remain separated from the diagonal by some strictly positive number.

for all $\mu \geq M$. Suppose an individual contemplates a move beyond a productivity of M without changing occupation; i.e., there exists t such that she moves from $\mu^o(t-1) \geq M$ to $\mu^o(t) > \mu^o(t-1)$. Let $\Delta \equiv \mu^o(t) - \mu^o(t-1)$. Then the lifetime wage gain as a result of this move is bounded above by $\bar{w}^o \Delta/(1-\beta)$. Also, the higher productivity can lower the marginal cost of subsequent actions. By (H.2), these gains are bounded above by $\bar{p}_e L_j \Delta$, where $\bar{p}_e$ is an upper bound on the price of education. So total gains are bounded above by

$$\frac{\bar{w}^o \Delta}{1-\beta} + \bar{p}_e L_j \Delta \tag{55}$$

On the other hand, the cost of this move is given by

$$p_e(t)e(\mu^o(t-1), \mu^o(t), o, o) = p_e(t)e(\mu^o(t-1), \mu^o(t-1) + \Delta, o, o) \geq \underline{p}_e S\Delta.$$

Combining this expression with (54) and (55), we must conclude that the cost of the proposed move exceeds its benefits, so it will never be made. That proves the Claim.

For parts (ii) and (iii), minor adjustments are needed. In (ii), we prove that any sector j must be asymptotically fully automated along any subsequence in which its output grows. Just as in the proof of Theorem 1, we can first show that inputs in some occupation $o \in O_j$ in that sector must also grow. Now we consider two possibilities. If $w^o(t)$ grows along some further subsequence, then the share of human labor income in total income to occupation o must converge to zero along that subsequence; Lemma 3. The second possibility is that $w^o(t)$ is bounded. Then by the Claim, individual productivity is also bounded, and — given that this occupation grows — it must become asymptotically automated.

For part (iii), we need to show again that

$$\Psi_j(t) = \frac{\sum_{o \in O_j} w^o(t) h^o(t)}{p_j(t) y_j(t)} \in [0, 1],$$

converges to zero, as in the proof of Theorem 1. Very similar (and minor) changes need to be made as in the preceding paragraph, using the Claim. We omit the details. With this established, there is no change in the rest of the argument to establish Theorem 1. ■