ANNUAL REVIEWS

Annual Review of Economics

Directed Technical Change in Labor and Environmental Economics

David Hémous[1,2] and Morten Olsen[3]

[1]Department of Economics, University of Zurich, 8006 Zurich, Switzerland; email: david.hemous@econ.uzh.ch

[2]Centre for Economic Policy Research, London EC1V 0DX, United Kingdom

[3]Department of Economics, University of Copenhagen, 1165 Copenhagen, Denmark; email: mgo@econ.ku.dk

Annu. Rev. Econ. 2021. 13:571–97

First published as a Review in Advance on May 20, 2021

The *Annual Review of Economics* is online at economics.annualreviews.org

https://doi.org/10.1146/annurev-economics-092120-044327

JEL codes: O31, O33, O41, O44, E25, J24, Q55

Keywords

endogenous growth, automation, directed technical change, climate change, income inequality

Abstract

It is increasingly evident that the direction of technological change responds to economic incentives. We review the literature on directed technical change in the context of environmental economics and labor economics, and we show that these fields have much in common both theoretically and empirically. We emphasize the importance of a balanced growth path and show that the lack of such a path is closely related to the slow development of green technologies in environmental economics and to growing inequality in labor economics. We discuss whether the direction of innovation is efficient.

1. INTRODUCTION

Economists have long recognized that the direction of innovation responds to economic incentives (Hicks 1932, Kennedy 1964). Whereas early models of endogenous growth, such as those by Romer (1990) and Aghion & Howitt (1992), only featured one type of innovation, later models of directed technical change (DTC) were quickly developed that included several types of innovation. The earliest example is provided by Aghion & Howitt (1996), who model separately research and development and analyze researchers' incentives to allocate their efforts to each one of the two.[1] Closer to the questions posed by Hicks (1932) and Kennedy (1964), Acemoglu (1998) develops the canonical DTC model, in which innovation can augment either low- or high-skill labor. Since then, the insights of DTC have been incorporated into several areas of economics, and here we focus on two of them: environmental economics and labor economics.

Despite some differences between these two strands of literature, we show that they have much in common both theoretically and empirically, as demonstrated by frequent cross-fertilization between the two. On the theory side, we emphasize two aspects: First, we ask whether a given model features a balanced growth path (BGP), that is, whether there exists an equilibrium path in which relevant variables grow at equal rates. The lack of such a feature is closely related to the slow development of green technologies in environmental economics and to rising inequality in labor economics. Second, we discuss whether the direction of innovation is efficient: Are clean research subsidies necessary to address climate change in the presence of carbon taxation? Is there too much automation? Further, we show that there is overwhelming empirical evidence that the direction of technology responds strongly to economic incentives in the environmental context, and similar evidence is emerging in the labor context.

Section 2 briefly presents a version of the DTC models of Acemoglu (1998, 2002). Section 3 shows how environmental economics has used this framework. Section 4 continues with more recent DTC models that depart from the usual assumption of factor-augmenting technical change to study automation. Finally, Section 5 presents empirical evidence.

2. THE CANONICAL DIRECTED TECHNICAL CHANGE MODEL

The last two decades of the twentieth century saw a concurrent increase in both the skill ratio and the skill premium. A common explanation for this is that, because of skill-biased technical change, the relative demand for skill outpaced the relative supply (Goldin & Katz 2008). Acemoglu (1998) notes that this simultaneity is in need of an explanation and seeks to endogenize the rise in skill demand as a technological response to the increase in the skill ratio. Acemoglu (1998, 2002) describes two competitive markets for intermediate goods. These goods are combined using an aggregate constant elasticity of substitution (CES) production function,

$$Y = \left(Y_{\mathrm{L}}^{\frac{\varepsilon-1}{\varepsilon}} + Y_{\mathrm{H}}^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}}, \qquad 1.$$

where $\varepsilon > 0$ is the elasticity of substitution between the two (we omit time subscripts when they are not necessary). In the original setting, Y_{L} is an intermediate input produced using low-skill labor, whereas Y_{H} is produced using high-skill labor. However, as we show below, this framework is well suited for other applications. The intermediate inputs are each produced using a combination of labor and a unique set of machines of mass 1. These machines are distinct for each sector, and

[1] In their model, research corresponds to the development of a new potential line of products, and development corresponds to secondary innovations that introduce one of these products to the market.

their productivity evolves endogenously. The level of technology of the most advanced machine, the one employed in equilibrium, is denoted by $A_{ji} > 0$ for $j \in \{\mathrm{L}, \mathrm{H}\}$ and $i \in [0, 1]$. The production functions for the two sectors are

$$Y_{\mathrm{L}} = \frac{1}{1-\beta} L_{\mathrm{L}}^{\beta} \int_0^1 A_{\mathrm{L}i}^{\beta} x_{\mathrm{L}i}^{1-\beta} di \quad \text{and} \quad Y_{\mathrm{H}} = \frac{1}{1-\beta} L_{\mathrm{H}}^{\beta} \int_0^1 A_{\mathrm{H}i}^{\beta} x_{\mathrm{H}i}^{1-\beta} di, \qquad 2.$$

where L_j is the supply of workers of type j. Machines are produced monopolistically with $1 - \beta$ units of the final good, which is the numeraire. With a demand elasticity of $1/\beta$, the price of a machine is 1.

Let p_j denote the price of the intermediate goods. We can use the monopolist's solution to find that output for intermediate input $j = L, H$ obeys

$$Y_j = \frac{1}{1-\beta} p_j^{(1-\beta)/\beta} A_j L_j, \qquad 3.$$

where $A_j \equiv \int_0^1 A_{ji} di$ is the aggregate technology in sector j. The profits of a monopolist are given by

$$\pi_{ji} = \beta p_j^{1/\beta} L_j A_{ji}. \qquad 4.$$

Combining the final good producer's problem with labor market–clearing conditions gives the skill premium

$$\frac{w_{\mathrm{H}}}{w_{\mathrm{L}}} = \left(\frac{L_{\mathrm{H}}}{L_{\mathrm{L}}}\right)^{-\frac{1}{\sigma}} \left(\frac{A_{\mathrm{H}}}{A_{\mathrm{L}}}\right)^{\frac{\sigma-1}{\sigma}}, \qquad 5.$$

where $\sigma \equiv 1 + \beta(\varepsilon - 1) > 0$ is the derived elasticity of substitution between L and H, and $\sigma > 1$ if and only if $\varepsilon > 1$. This mirrors the framework of Goldin & Katz (2008) (building on Katz & Murphy 1992). They focus on the college skill premium and reconcile the large rise in college attainment with the substantial increase in the skill premium since the 1980s by arguing that low-skill and high-skill workers are gross substitutes—i.e., $\sigma > 1$—and by inferring a positive secular trend in $A_{\mathrm{H}}/A_{\mathrm{L}}$. They take this trend in skill-biased technical change as exogenous, whereas Acemoglu (1998) argues that when technology is endogenous, the growth in $A_{\mathrm{H}}/A_{\mathrm{L}}$ can be driven by the increase in the skill ratio, $L_{\mathrm{H}}/L_{\mathrm{L}}$.

To demonstrate this, we model innovation in a quality ladder fashion (Aghion & Howitt 1992). The literature generally models the cost of innovation in terms of the final good or of a limited factor of production, that is, scientists (Acemoglu 2002). To facilitate comparison with Section 3.1, we implement the latter case. Time is discrete, and the usual Ramsey setup gives the interest rate r_t. At the beginning of every period, scientists of mass $S = 1$ can work to innovate either in the low-skill-intensive sector or in the high-skill-intensive sector. Given this choice, each scientist is randomly allocated to one machine in their target sector without congestion (this can be rationalized using within-sector spillovers).

Following Acemoglu (2002), the probabilities of successful innovation for scientists in the low-skill and the high-skill sector are given by $\eta_{\mathrm{L}}(A_{\mathrm{H}t}/A_{\mathrm{L}t})^{(1-\delta)/2}$ and $\eta_{\mathrm{H}}(A_{\mathrm{L}t}/A_{\mathrm{H}t})^{(1-\delta)/2}$, respectively.[2] δ is inversely related to the complementarity of technologies in the innovation functions. When

[2] Acemoglu (2002) has an expanding variety framework, and the probability of innovation for each scientist in the low-skill sector obeys $\eta_{\mathrm{L}} N_{\mathrm{L}}^{(1+\delta)/2} N_{\mathrm{H}}^{(1-\delta)/2}$, where N_{L} (N_{H}) is the mass of low-skill (high-skill) products. This formulation is equivalent to ours, since in the expanding variety model, profits for each firm are mechanically diluted with the number of products: They are proportional to $p_{\mathrm{L}} Y_{\mathrm{L}}/N_{\mathrm{L}}$ in the expanding variety model but to $p_{\mathrm{L}} Y_{\mathrm{L}}$ in the quality ladder model.

$\delta = 1$, the innovation possibility frontier is independent of the technology levels. When $\delta < 1$, the productivity of innovation declines with the level of technology in the same sector, but knowledge spillovers from the other sector compensate for this in a way that permits a BGP. Once innovation is complete, the scientist increases the quality of their targeted machine by a factor $1 + \gamma$ and obtains monopoly rights until they are replaced by a future innovator. We impose the inconsequential assumption that $(1 + \gamma) > (1 - \beta)^{\frac{\beta-1}{\beta}}$, which ensures that the technological leader charges the unconstrained monopoly price.

We focus on a BGP in which the two technologies grow at the same rate, and the probability ρ that an incumbent is replaced by an entrant is constant and is the same in both sectors. Moreover, the interest rate r and the profits for a given technology are also constant. Therefore the value of a firm obeys

$$V_{ji} = \frac{\pi_{ji}(1+r)}{r+\rho}. \qquad 6.$$

Because scientists are randomly allocated within a sector, the expected technology obtained by an innovator in sector j is given by $(1 + \gamma)A_{j(t-1)}$. Using Equation 4, Equation 6, and the BGP condition that the two technologies grow at the same rate [such that $A_{\mathrm{L}t}/A_{\mathrm{H}t} = A_{\mathrm{L}(t-1)}/A_{\mathrm{H}(t-1)}$], we obtain the relative value of innovating in the low-skill versus the high-skill sector, Ω, as

$$\Omega_t = \frac{\eta_{\mathrm{L}}}{\eta_{\mathrm{H}}}\left(\frac{A_{\mathrm{H}t}}{A_{\mathrm{L}t}}\right)^{1-\delta}\frac{p_{\mathrm{L}t}Y_{\mathrm{L}t}}{p_{\mathrm{H}t}Y_{\mathrm{H}t}} = \frac{\eta_{\mathrm{L}}}{\eta_{\mathrm{H}}}\underbrace{\left(\frac{p_{\mathrm{L}t}}{p_{\mathrm{H}t}}\right)^{\frac{1}{\beta}}}_{\text{price effect}}\underbrace{\frac{L_{\mathrm{L}}}{L_{\mathrm{H}}}}_{\text{market size effect}}\underbrace{\left(\frac{A_{\mathrm{L}t}}{A_{\mathrm{H}t}}\right)^{\delta}}_{\text{technology effects}}.$$

The first equality emphasizes Kennedy's (1964) finding that the relative incentive to innovate combines the innovation possibility frontier and the relative factor shares (more specifically, intermediate input shares). The second equality emphasizes Acemoglu's (2002) decomposition between a price effect, a market size effect, and technology effects. Innovation has higher value in the sector with the more expensive good and the larger labor market. Technology also directly increases the value of innovation, but this effect is diminished by the presence of knowledge spillovers (when $\delta < 1$) across the two types of technology. Solving for the relative price $p_{\mathrm{L}}/p_{\mathrm{H}}$ returns

$$\Omega_t = \frac{\eta_{\mathrm{L}}}{\eta_{\mathrm{H}}}\left(\frac{L_{\mathrm{L}}}{L_{\mathrm{H}}}\right)^{\frac{\sigma-1}{\sigma}}\left(\frac{A_{\mathrm{L}t}}{A_{\mathrm{H}t}}\right)^{\frac{\delta\sigma-1}{\sigma}}. \qquad 7.$$

Innovation can only occur in both sectors when $\Omega = 1$. If $\delta\sigma > 1$, the relative incentive to innovate in low-skill products is increasing in $A_{\mathrm{L}t}/A_{\mathrm{H}t}$, and a BGP is not stable. Therefore, except for knife-edge cases, the economy eventually features innovation in only one sector. Intuitively, the sector with a technological advantage commands a larger revenue share when the elasticity of substitution σ is larger, and the knowledge spillovers are lower when δ is higher, both of which make a BGP less likely.

In contrast, if $\delta\sigma < 1$, a stable BGP with innovation in both sectors is possible. Solving for $A_{\mathrm{L}t}/A_{\mathrm{H}t}$ and using the expression for the skill premium in Equation 5 we obtain, on a BGP,

$$\frac{w_{\mathrm{H}}}{w_{\mathrm{L}}} = \left(\frac{\eta_{\mathrm{H}}}{\eta_{\mathrm{L}}}\right)^{\frac{\sigma-1}{1-\sigma\delta}}\left(\frac{L_{\mathrm{H}}}{L_{\mathrm{L}}}\right)^{\frac{\sigma-2+\delta}{1-\delta\sigma}}.$$

This replicates the strong induced-bias hypothesis of Acemoglu (1998): If $\sigma > 2 - \delta$ (and $\delta\sigma < 1$), an increase in the skill ratio increases the skill premium. Intuitively, an increase in the skill ratio leads to skill-biased technical change: An increase in $L_{\mathrm{H}}/L_{\mathrm{L}}$ decreases Ω, which pushes innovation toward the high-skill sector if and only if $\sigma > 1$ (see Equation 7), and a decline in A_L/A_H is high-skill biased if and only if $\sigma > 1$ (see Equation 5). When the two inputs are sufficient substitutes the technological response is sufficient to overturn the direct supply effect.[3]

Acemoglu (2003) uses an analogous framework to demonstrate that when capital is a reproducible factor, and capital and labor are complements, innovation is labor augmenting. This endogenizes one of the assumptions underlying the Uzawa theorem (Uzawa 1961) and ensures stable factor shares. An extensive literature has emerged building on the framework of Acemoglu (1998), including work by Acemoglu & Zilibotti (2001) and Acemoglu et al. (2012b). In the following, we focus on applications to environmental economics (Section 3) and on new DTC models that depart from factor-augmenting technologies (Section 4).[4]

3. DIRECTED TECHNICAL CHANGE AND THE ENVIRONMENT

While policy makers and climate scientists have long argued that overcoming the challenges of climate change requires the development of clean technologies, the economics literature initially focused on models with exogenous technological change (see, e.g., Nordhaus 1994). Meanwhile, a growing empirical literature has shown that innovation responds to energy prices (see Section 5). Several papers have added induced technical change to computable general equilibrium (CGE) models; still, they do not build on modern growth theory and therefore either ignore knowledge externalities or model them in an ad-hoc way. For instance, in work by Nordhaus (2002) and Popp (2004, 2006), technological progress results from the accumulation of an R&D stock similar to capital.[5] Bovenberg & Smulders (1995, 1996) present the first model of modern endogenous growth theory in an environmental context, but they only model one type of innovation.

We focus here on DTC models, which build on modern endogenous growth theory and feature two different types of innovations. In the environmental context, these models come in two varieties. Some focus on energy-saving innovation and model either energy or a resource as an input that is complementary to capital or labor (the first example is in Smulders & de Nooij 2003). Other models analyze DTC between two substitute inputs, one of which is cleaner than the other (Acemoglu et al. 2012).[6] We start with the substitute case in Section 3.1, move to the complement case in Section 3.2, and present further applications of the DTC framework in Section 3.3. Our review is not exhaustive, and we focus primarily on recent work.[7]

3.1. The Substitute Case: Clean and Dirty Energy

Acemoglu et al. (2012) build on the framework of Section 2, but the two inputs differ in whether they generate greenhouse gas emissions (the dirty input, Y_{dt}) or not (the clean input, Y_{ct}). The two

[3] Acemoglu (2007) demonstrates that this result holds very generally in models with factor-augmenting technology. Loebbing (2021) demonstrates that an increase in the skill ratio increases the skill premium if and only if the production function is quasi-convex once one takes into account the technology response.

[4] We focus on models of imperfect competition where profits drive innovation effort. There is a small literature on DTC models with perfect competition, but these models are too different in their setup to be covered here (see Irmen 2017, Irmen & Tabaković 2017, Casey & Horii 2019, and references therein).

[5] Readers are referred to Goulder & Schneider (1999), Sue Wing (2003), and Massetti et al. (2009). Gerlagh & Lise (2005) and Grimaud & Rouge (2008) microfound innovation but still impose ad-hoc relationships between its social and private values.

[6] Instead of building on Acemoglu's (1998) DTC framework, Hart (2004) and Ricci (2007) present models in which innovation either increases only the productivity of an intermediate or increases it by a lower amount while making it cleaner.

[7] For other literature reviews, readers are referred to Popp et al. (2010) and Fischer & Heutel (2013).

inputs are assumed to be substitutes ($\varepsilon > 1$; see empirical evidence in Papageorgiou et al. 2017), so that this framework can be used to analyze the choice between renewable (or nuclear) and fossil fuel energy, or the choice between electric and fossil fuel vehicles.[8] Production occurs as described in Section 2, except that the labor allocation between the two sectors is endogenous.

CO_2 emissions are directly proportional to the use of the dirty input. Implicitly, using the dirty input requires consuming a freely available fossil fuel with a Leontief technology. As a result, Acemoglu et al. (2012) do not model improvements in the energy efficiency or resource productivity (i.e., thermal efficiency) of power plants or fossil fuel vehicles, but they rather focus on other innovations that reduce their effective costs.[9]

Innovation is modeled as in the previous section, except that patents only last for one period and $\delta = 1$, so that the innovation possibility frontier is independent of the technology levels.[10] The law of motion of input $j \in \{\text{c}, \text{d}\}$ technology is

$$A_{jt} = \left(1 + \gamma \eta_j s_{jt}\right) A_{jt-1},$$

where s_{jt} is the mass of scientists in sector j, η_j is their productivity, and γ is the innovation size. This innovation setup features a building-on-the-shoulders-of-giants externality, since innovators not only improve the current technology but also enable future innovators to build on their innovations.

As profits still obey Equation 4, the expected profits of a scientist working for sector j are given by

$$\Pi_{jt} = \eta_j \left(1 + \gamma\right) \beta p_{jt}^{\frac{1}{\beta}} L_{jt} A_{j(t-1)} = \frac{\eta_j \beta p_{jt} Y_{jt}}{1 + \gamma \eta_j s_{jt}}.$$

Scientists target the sector with the highest expected profits, which is the clean sector if the following ratio is greater than 1:

$$\frac{\Pi_{\text{c}t}}{\Pi_{\text{d}t}} = \frac{\eta_\text{c} \left(1 + \gamma \eta_\text{d} s_{\text{d}t}\right)}{\eta_\text{d} \left(1 + \gamma \eta_\text{c} s_{\text{c}t}\right)} \frac{p_{\text{c}t} Y_{\text{c}t}}{p_{\text{d}t} Y_{\text{d}t}} = \frac{\eta_\text{c}}{\eta_\text{d}} \underbrace{\left(\frac{p_{\text{c}t}}{p_{\text{d}t}}\right)^{\frac{1}{\beta}}}_{\text{price effect}} \underbrace{\frac{L_{\text{c}t}}{L_{\text{d}t}}}_{\text{market size effect}} \underbrace{\frac{A_{\text{c}t-1}}{A_{\text{d}t-1}}}_{\text{direct productivity effect}}. \qquad 8.$$

Therefore, scientists target the sector with the largest revenue (adjusted with the productivity of the innovation technology). Relative revenues depend on the same forces as above. Yet, there are no cross-sectoral knowledge spillovers and the labor allocation is now endogenous, with the more advanced sector attracting relatively more labor when the inputs are substitute.

We can then express the relative expected profits from innovation as

$$\frac{\Pi_{\text{c}t}}{\Pi_{\text{d}t}} = \frac{\eta_\text{c}}{\eta_\text{d}} \left(\frac{1 + \gamma \eta_\text{c} s_{\text{c}t}}{1 + \gamma \eta_\text{d} s_{\text{d}t}}\right)^{\sigma - 2} \left(\frac{A_{\text{c}t-1}}{A_{\text{d}t-1}}\right)^{\sigma - 1}. \qquad 9.$$

When the two inputs are substitutes, the price effect is weaker and innovation tends to be directed toward the most advanced sector: It exhibits path dependence, which is the first lesson of the

---

[8]Aghion & Howitt (2009, chapter 16) preempt some of the results of Acemoglu et al. (2012) in the case of perfect substitutes.

[9]Such innovation could be included if pollution were proportional to the use of dirty machines, or $x_{\text{d}it}$ [similar to Gans's (2012) model]. This would not change any of the following results.

[10]The supply of R&D resources is fixed, so that clean R&D fully crowds out dirty R&D. This is not an innocuous assumption, as a policy that aims at increasing clean innovation also depresses dirty innovation and output growth (see Popp 2004).

framework.[11] In fact, this is a so-called bang-bang solution, and for a sufficiently low ratio A_{c0}/A_{d0}, all innovation at time 1 occurs in the dirty sector. A_{ct}/A_{dt} further declines and innovation remains locked in dirty technologies. Intuitively, we should not expect much clean innovation in a laissez-faire equilibrium, because an innovation that aims at improving a component in a solar panel would have a much smaller market than an innovation aimed at improving a component in a fossil fuel power plant. Therefore, whereas the canonical model of Section 2 focuses on a BGP, we focus here on unbalanced trajectories.[12]

As a result, when fossil fuel technologies are initially ahead, the production of dirty inputs in the laissez-faire equilibrium grows without bound, and so do CO_2 emissions. To prevent this, a social planner could implement a carbon tax or research subsidies for clean innovation. A carbon tax imposes a wedge between the producer price of the dirty input and its marginal product in the final good production, and it decreases the producer price p_{dt} for given technologies in Equation 8. A clean research subsidy directly multiplies the right-hand side of Equation 8.[13] With a sufficiently strong policy intervention, the social planner can redirect innovation away from dirty technologies and toward clean ones. If this intervention is maintained for a sufficiently long time, clean technologies will catch up, and market forces will favor clean innovation. When the two inputs are sufficient substitutes ($\varepsilon > 1/\beta$), a temporary intervention is enough to ensure that emissions will decline in the long run. This intervention, however, has the cost of lower productivity growth during the catch-up phase while innovation is improving the less productive input. Yet, the longer the social planner waits, the larger the gap between clean and dirty technologies before the intervention, and the longer the intervention and the larger the costs. This is the second lesson from the framework: Taking endogenous technical change into account calls for earlier intervention. Gerglagh et al. (2009) similarly find that endogenous innovation in abatement technology calls for a front-loaded policy.

Finally, Acemoglu et al. (2012) study the optimal policy when the representative agent values consumption and is hurt by environmental degradation. They show that this policy can be decentralized using a Pigovian carbon tax and research subsidies to clean innovation (plus a subsidy to remove the monopoly distortion). This is the third lesson from the framework: A carbon tax is not enough to obtain the first best. In the optimum, innovation is allocated to the sector with the highest social value. The ratio of social values can be expressed as

$$\frac{SV_{ct}}{SV_{dt}} = \frac{\eta_c\,(1+\gamma\,\eta_d s_{dt})\sum_{\tau\geq t}\lambda_{t,\tau}\,p_{c\tau}^{\frac{1}{\beta}}L_{c\tau}A_{c\tau}}{\eta_d\,(1+\gamma\,\eta_c s_{ct})\sum_{\tau\geq t}\lambda_{t,\tau}\,p_{d\tau}^{\frac{1}{\beta}}L_{d\tau}A_{d\tau}}, \tag{10}$$

where $\lambda_{t,\tau}$ is the discount factor between t and τ. This ratio reflects the environmental value, as a higher carbon tax decreases $p_{d\tau}$. However, even with a carbon tax, the market still allocates innovation according to the ratio of Equation 8, and in general it will not implement the first best without a research subsidy. Intuitively, the social planner allocates innovation according to the discounted benefits that a higher technology brings in every period, while the market only cares about immediate profits.

[11] For more discussion on path dependence, readers are referred to the review by Aghion et al. (2019a).

[12] With cross-sectoral knowledge spillovers, as in the innovation function of Section 2 where scientists' productivity obeys $\eta_j(A_{(-j)t}/A_{jt})^{(1-\delta)/2}$, there is still path dependence when $\sigma > 2-\delta$. The difference with the threshold $\sigma > 1/\delta$ given above comes from the endogeneity of the labor allocation.

[13] That is, Equation 9 becomes $\frac{\Pi_{ct}}{\Pi_{dt}} = (1+q_t)\,(1+\tau_t)^{\epsilon}\,\frac{\eta_c}{\eta_d}\left(\frac{1+\gamma\,\eta_c s_{ct}}{1+\gamma\,\eta_d s_{dt}}\right)^{\sigma-2}\left(\frac{A_{ct-1}}{A_{dt-1}}\right)^{\sigma-1}$, where q_t is a clean research subsidy and τ_t is an (ad valorem) carbon tax.

This intuition extends to the case of patents lasting more than one period, or of patents lasting until the following innovation [as mentioned by Acemoglu et al. (2012); see also Greaker et al. 2018]. Moreover, the private value does not internalize the building-on-the-shoulders-of-giants externality. To see this, consider an extreme case with perpetual patents, such that future innovators would have to pay royalties to the incumbents to compensate them for their profit losses once the new technologies have arrived. In that setup, the ratio of private values of innovation would obey

$$\frac{\Pi_{\mathrm{c}t}}{\Pi_{\mathrm{d}t}} = \frac{\eta_{\mathrm{c}}\left(1+\gamma\eta_{\mathrm{d}}s_{\mathrm{d}t}\right)\sum_{\tau\geq t}\lambda_{t,\tau}p_{\mathrm{c}\tau}^{\frac{1}{\beta}}L_{\mathrm{c}\tau}A_{\mathrm{c}t}}{\eta_{\mathrm{d}}\left(1+\gamma\eta_{\mathrm{c}}s_{\mathrm{c}t}\right)\sum_{\tau\geq t}\lambda_{t,\tau}p_{\mathrm{d}\tau}^{\frac{1}{\beta}}L_{\mathrm{d}\tau}A_{\mathrm{d}t}}. \tag{11}$$

In this setting, only the building-on-the-shoulders-of-giants externality is active. The only difference between Equations 10 and 11 is that the sum in Equation 11 only considers the expected technology of the current innovation A_{jt}, whereas the sum in Equation 10 considers the full path of expected future technologies $A_{j\tau}$. This corresponds directly to the building-on-the-shoulders-of-giants externality: An innovator improves not only current technology but also all future technologies, since innovators build on each other's work.[14] Therefore, finite-lived patents, creative destruction, imitation, and the building-on-the-shoulders-of-giants externality all imply that the private value of an innovation tends to be more short-sighted than its social value. The key is that this short-termism does not affect clean and dirty technologies equally. Consider a setting in which dirty technologies are initially more advanced, but the clean technologies must dominate in the future in the social planner's allocation. A larger fraction of the social value of dirty innovation is realized in the short run than is the case for clean innovation. In other words, a high share of the social value from improving a solar panel today comes from the benefits of getting better solar panels in the future, while most of the benefits from improving coal power plants are realized today. Then, the short-termism in the market innovation allocation implies inefficiently low clean innovation relative to dirty innovation, even with Pigovian taxation.[15]

To summarize, Acemoglu et al. (2012) provide three lessons. First, there is path dependence in the development of clean versus dirty technologies. Second, taking into account the endogeneity of innovation calls for earlier action. Third, in addition to Pigovian carbon taxation, the optimal policy includes research subsidies specifically devoted to clean innovation.

Acemoglu et al. (2016) further build on Acemoglu et al.'s (2012) work by calibrating a firm dynamics model with clean and dirty innovation. A final good is produced as a Cobb–Douglas aggregate of intermediates. Each intermediate can be produced with a clean or a dirty technology; the two evolve on their own ladder and are perfect substitutes within a line. A firm is a collection of leading clean or dirty technologies in different lines. Innovation can be incremental, building on each technology separately, or radical, building on the leading technology whether it is clean or dirty. As a result, their model features cross-sectoral spillovers that were absent in Acemoglu et al.'s (2012) model and mitigate (without eliminating) path dependence in innovation. As in the new DTC models of Section 4, technical change in each line is microfounded (similar to the tasks below) instead of immediately taking a factor-augmenting form at the aggregate level. Acemoglu et al. (2016) calibrate their model to the US energy sector. Their conclusions are in line with those

[14]In contrast, the optimal policy in a model with horizontal innovation and DTC need not feature research subsidies in addition to Pigovian taxation.

[15]Gerlagh et al. (2014) make a related point in a model with clean innovation only. This contrasts with an earlier literature of integrated assessment models with a constant ratio of social to private value of innovation (Nordhaus 2002; Popp 2004, 2006; Gerlagh & Lise 2005).

of Acemoglu et al. (2012): The optimal policy requires both (large) clean research subsidies and a carbon tax, and it features a rapid switch from dirty to clean innovation.

3.2. The Complementarity Case: Energy-Saving Innovation

Whereas Acemoglu et al. (2012) focus on the development of clean substitutes to dirty inputs, others have focused on energy- or resource-saving innovations. In our framework, the final good is produced according to Equation 1, but the two inputs Y_{L} and Y_{H} are replaced by a production input Y_{P} and an energy-services input Y_{E}, with $\epsilon < 1$. Both inputs are produced as in Equation 2, but with a capital-labor aggregate instead of low-skill labor for Y_{P} and energy (or a fossil fuel resource) E for Y_{E}. The papers differ in whether there is a fixed resource flow (Smulders & de Nooij 2003), a constant resource price (Shanker & Stern 2018), or an exhaustible resource stock (André & Smulders 2014, Hassler et al. 2019). This literature seeks to account for stylized facts regarding energy consumption and growth.

In particular, Hassler et al. (2019) build a quantitative macroeconomic model. They estimate an elasticity of substitution between energy and other inputs close to 0 and show that energy-saving technical change took off in the 1970s with the oil shocks, in line with the DTC theory. Their model further predicts that thanks to the innovation response, resource scarcity will only lead to a slight increase in the energy share. One of their conclusions is that "subsidies may not be necessary for regulating the direction of technical change." Why are their conclusions so different from Acemoglu et al.'s (2012)? This is because energy is a complement to other inputs, and consequently a BGP arises more easily.

Within the framework we have sketched and with one-period patents, the relative expected profits from labor-augmenting over energy-augmenting innovation obey

$$\frac{\Pi_{\mathrm{P}t}}{\Pi_{\mathrm{E}t}} = \frac{\eta_{\mathrm{P}}\left(1+\gamma\eta_{\mathrm{E}}s_{\mathrm{E}t}\right)}{\eta_{\mathrm{E}}\left(1+\gamma\eta_{\mathrm{P}}s_{\mathrm{P}t}\right)}\frac{p_{\mathrm{P}t}Y_{\mathrm{P}t}}{p_{\mathrm{E}t}Y_{\mathrm{E}t}} = \frac{\eta_{\mathrm{P}}}{\eta_{\mathrm{E}}}\left(\frac{p_{\mathrm{P}t}}{p_{\mathrm{E}t}}\right)^{\frac{1}{\beta}}\frac{L}{E}\frac{A_{\mathrm{P}t-1}}{A_{\mathrm{E}t-1}} = \frac{\eta_{\mathrm{P}}\left(1+\gamma\eta_{\mathrm{E}}s_{\mathrm{E}t}\right)}{\eta_{\mathrm{E}}\left(1+\gamma\eta_{\mathrm{P}}s_{\mathrm{P}t}\right)}\left(\frac{LA_{\mathrm{P}t}}{EA_{\mathrm{E}t}}\right)^{\frac{\sigma-1}{\sigma}}. \qquad 12.$$

Since the two inputs are complementary, the price effect dominates. When the resource flow is constant (as in Smulders & de Nooij 2003), innovation tends to favor the least advanced sector (following the last expression in Equation 12), and in the long run, the economy converges to a BGP with innovation in both sectors ($\Pi_{\mathrm{P}t} = \Pi_{\mathrm{E}t}$), equal growth in the two sectors, and constant factor shares (following the first equality in Equation 12). When the resource flow decreases (because of resource exhaustion or a growing carbon tax), the resulting increase in the energy share favors energy-saving technical change (following again the last expression in Equation 12).[16] Here as well, the economy converges in the long run toward a BGP that has a constant interior energy share but in which energy-saving technical change $A_{\mathrm{E}t}$ grows faster than labor-saving technical change $A_{\mathrm{P}t}$ to compensate for the reduction in the resource flow. This follows the same logic proposed by Acemoglu (2003), in which labor scarcity leads to labor-augmenting technical change. With a constant resource price, the logic is reversed and innovation in the long run is entirely labor augmenting.

The social planner solution also features balanced growth and converges toward the same innovation allocation as in the equilibrium, provided that energy is optimally priced through a carbon tax. Research subsidies may be necessary in the transition, but their importance is greatly reduced—Hassler et al. (2019) show a case in which they are not necessary in the transition either (see also Hart 2008). The short-termism of the market innovation allocation now favors the least

[16]A tax on energy, E, moves innovation toward A_{E} when $\varepsilon < 1$; a tax on energy services Y_{E} moves innovation toward A_{P} regardless of the value of ε.

advanced technology, adjusting for resource availability, which ensures that the economy moves toward a BGP, as called for by the social planner. While public intervention is crucial to the development of clean alternatives to fossil fuel energy, carbon pricing can do the heavy lifting for the development of energy-saving technologies.[17]

A consequence of DTC is that while the short-run elasticity between energy and other inputs is very low, the long-run energy share is constant. One may be tempted to conclude that climate models are not missing much by ignoring energy-saving technical change and simply assuming that energy enters final good production in a Cobb–Douglas way. Casey (2019), however, shows that this would be misguided. He builds a model similar to the one by Hassler et al. (2019), in which energy and the capital-labor aggregate are combined in a Leontief production function for given technologies, but the long-run elasticity is 1 for the same reason as above. He calibrates both his DTC model and a Cobb–Douglas economy to US data, and he shows that a given carbon tax is less effective at reducing cumulative emissions in the DTC model. Intuitively, technological adjustment is sluggish and with a Leontief production function, emissions do not decline as rapidly as in a Cobb–Douglas setting. Because climate damages depend on the stock of emissions, this transition period matters quantitatively.

3.3. Applying Directed Technical Change to Environmental Questions

In the following, we review papers that use these two DTC frameworks in the context of energy shocks, historical energy transitions, and carbon leakage.

3.3.1. Energy market shocks. Fried (2018) uses the oil shocks of the 1970s to calibrate a DTC model that combines elements of both Acemoglu et al.'s (2012) and Hassler et al.'s (2019) models but features a more detailed representation of the economy. A final good is produced with a production input and energy services; the latter are themselves an aggregate of local fossil fuel energy, oil imports, and green energy. The production input and energy services are highly complementary, while the different types of energy are substitutes. Innovation can be targeted at local fossil fuel energy, green energy, or the production input. As in Acemoglu et al.'s (2012) model, emissions are proportional to the quantity of fossil fuel energy. Fried (2018) studies the implementation of a carbon tax, which cuts emissions by 30% in 20 years. Such a carbon tax redirects innovation away from fossil fuel energy toward mostly green energy. DTC reduces the size of the necessary carbon tax by 19.2% compared to a model with exogenous technical change.[18]

Acemoglu et al. (2019) build on Acemoglu et al.'s (2012) model to study the shale gas boom, which started in 2009. They show that since then, the ratio of renewable patents relative to fossil fuel patents in the electricity sector has declined sharply. To analyze the consequences on emissions, they build a DTC model in which electricity can be green or can be produced with coal or natural gas. Innovation can be targeted at improving the productivity of fossil fuel power plants or green power plants. Following a drop in natural gas prices (as the one deriving from the shale gas boom), electricity production shifts toward natural gas. Because natural gas is much cleaner than coal, emissions decrease in the short run. However, the price decline also increases the market for innovations in fossil fuel power plants, and as a result green innovation declines. Calibrating their

[17]This conclusion may not hold in the presence of multiple equilibria, as shown by van der Meijden & Smulders (2017).

[18]Hart (2019) also calibrates an integrated assessment model with features à la Acemoglu et al. (2012). The optimal policy includes both a carbon tax and clean research subsidies, but the relative importance of research subsidies is diminished, particularly because of intersectoral knowledge spillovers.

model to the US electricity sector, Acemoglu et al. (2019) find that this innovation effect eventually dominates, so that emissions increase in the medium term following the shale gas boom. They argue that policy makers should react to the shale gas boom by raising subsidies to green innovation.[19]

3.3.2. Historical energy transitions. DTC can also be used to explain historical energy transitions. Stern et al. (2020) build on Acemoglu's (2002) model to explain the Industrial Revolution as resulting from the transition from a wood-powered to a coal-powered economy. In their model, the final good is produced with two substitute intermediate inputs, one wood intensive and the other coal intensive. Innovation may be directed at either. Wood is in fixed supply each period, whereas coal is supplied at a fixed extraction cost. Constant long-run growth is only possible in a coal-based economy, and their model can generate transitional dynamics akin to those of the British Industrial Revolution: Initially, the economy relies mostly on wood and grows slowly, but with economic development it progressively shifts toward coal, which spurs innovation in coal technologies through a market size effect. This leads to a takeoff in economic growth.[20]

Lemoine (2018) builds a DTC model in which different energy services are produced using two complementary inputs, machines and natural resources (as in Acemoglu et al. 2019), and in which natural resources are isoelastically supplied. Even though the model generates endogenous energy transitions, a calibration shows that the optimal climate policy still relies on clean research subsidies to accelerate the transition to renewables.

3.3.3. Carbon leakage. The models we have studied so far all consider either a country in isolation or a global solution. In practice, international climate negotiations have stalled, and countries have largely conducted climate policy unilaterally. International trade, however, may reduce the scope for unilateral actions as it may lead to carbon leakage (i.e., a move of the production of polluting goods from regulated to unregulated countries). The DTC literature shows that the elasticity of substitution between traded goods and the pattern of innovation/imitation across countries play a crucial role in determining carbon leakage. Di Maria & Smulders (2005) consider a two-country (North, South), two-good (energy-intensive, non-energy-intensive) trade model in which the North innovates while the South imitates exogenously. The implementation of a carbon tax in the North and the ensuing reallocation of energy-intensive production to the South leads to an increase in innovation in the non-energy-intensive sector. This reduces carbon leakage when the goods are substitutes and amplifies it when they are complements (because innovation in the energy-intensive sector is resource augmenting, and it is therefore resource saving in the complementary case). Di Maria & Valente (2008) start from the same setup but allow both countries to innovate on a global market. They find that carbon leakage is always reduced by the innovation response to a unilateral cut in emission. Acemoglu et al. (2014) and van den Bijgaart (2017) focus on endogenous imitation or innovation by the South (the unregulated country) by extending

[19]In a similar spirit, Acemoglu & Rafey (2019) look at the effect of an exogenous shock to geoengineering technology. They find that when environmental policy is endogenous and commitment is impossible, such a shock may decrease clean innovation, as it reduces future environmental taxes. Progress in geoengineering technology may then backfire, leading to an increase in emissions.

[20]Similarly, Gars & Olovsson (2019) build a DTC model to explain the nineteenth-century Great Divergence. In their model, a switch from wood-powered to fossil fuel–powered innovation leads to faster economic growth. However, when one country switches to fossil fuels, the world price of the latter increases, which reduces innovation in fossil fuel technologies elsewhere.

Acemoglu et al.'s (2012) model to a two-country setup. In both cases, the technological response by the South following a unilateral carbon tax by the North amplifies carbon leakage.

Hémous (2016) analyzes which unilateral policy can be successful with endogenous innovation. He also considers a two-country, two-good (energy-intensive, non-energy-intensive) trade model in which there is unit elasticity between the two goods, but the energy-intensive good can be produced in a clean or a dirty way (as in Acemoglu et al. 2012). In each country, innovation can be targeted at the non-energy-intensive sector or, within the energy-intensive sector, at clean or dirty technologies. The innovation response from the unregulated country amplifies carbon leakage: The implementation of a unilateral carbon tax displaces the production of the energy-intensive good toward the unregulated country, which increases dirty innovation in that sector when the dirty technology is more advanced than the clean one. A unilateral carbon tax may then backfire and lead to an increase in global emissions. Instead, a green industrial policy, consisting of green research subsidies and possibly carbon tariffs, can reduce emissions in both countries by directing innovation within the regulated country toward the clean sector and innovation in the unregulated country toward the non-energy-intensive sector and (with strong knowledge spillovers) toward clean energy. Overall, trade acts as a double-edged sword: Unilateral carbon taxes are less effective, but an appropriate policy can decrease emissions globally.[21]

3.3.4. Other applications and future research. DTC theory often provides policy answers that differ from those of models with exogenous technology, and it accounts well for historical trends. This calls for further integrating DTC in climate change economics. In particular, microfounded DTC should be more systematically incorporated in integrated assessment models. Dietz & Lanz (2019) offer a recent example in a detailed multisectoral model with endogenous population dynamics. Kruse-Andersen (2020) also includes population dynamics into a DTC model. Another important avenue for future research is the expansion of the two-country setups discussed above into more realistic models of international environmental agreements, building on game-theoretic contributions such as those by Barrett (2006) and Harstad et al. (2019). Finally, climate change is a problem riddled with uncertainties about climate dynamics and climate damages but also technological prospects. The models reviewed here are all deterministic, but the interaction between technology and uncertainty is a promising avenue for future research.[22]

4. AUTOMATION AND NEW DIRECTED TECHNICAL CHANGE MODELS

We have argued that the canonical DTC model has provided insights for the study of both income inequality and environmental issues. Nevertheless, a few papers have criticized this framework for being too imprecise in describing the effects of technology on work. Among these, Autor et al. (2003) postulate a "routinization hypothesis" by introducing the notion of tasks, that is, the work inputs required for producing a given output. They argue that because computers are

[21]Witajewski-Baltvilks & Fischer (2019) also build on Acemoglu et al.'s (2012) work, but they include trade in machines, so that innovation incentives reflect market conditions in both countries. A unilateral clean research subsidy can redirect innovation toward clean technologies in both countries if the regulated country is large enough. Moreover, it may induce the government of the unregulated country to introduce its own clean research subsidy, as long-run growth is higher when the two countries innovate in the same sector.
[22]Readers are also referred to Heutel et al. (2018), who show that geoengineering can be used as an insurance mechanism against climate uncertainty.

highly capable of performing tasks that can be codified in a computer program—labeled routine tasks—they are disproportionately substitutable for workers performing such tasks. Specifically, they model output in an industry *i* as

$$y(i) = (l_{\mathrm{R}}(i) + x(i))^{\beta_i}\, l_{\mathrm{N}}^{1-\beta_i}(i), \qquad 13.$$

where l_{R} denotes the input of routine labor, l_{N} the input of nonroutine labor, x the use of computer capital, and $\beta_i \in (0, 1)$ the importance of routine tasks in a given industry. Autor et al. (2003) also formalize technical progress as the continuously declining price of computer capital. When both computer capital and routine labor are employed, Equation 13 implies that low-skill wages equal the price of computer capital and decline correspondingly. Autor and colleagues show empirically that the implementation of computer capital correlates strongly with changes in the use of routine tasks across industries.

Acemoglu & Autor (2011) argue that the canonical model based on factor-augmenting technical change cannot account for several features of the evolution of the income distribution. These include a continuous increase in labor income inequality and absolute declines in low-skill wages (see also Acemoglu & Restrepo 2020c). Furthermore, the canonical model does not microfound automation as the replacement of workers with machines in the execution of certain tasks. To do so, they extend Autor et al.'s (2003) model with exogenous technology and endogenous assignment of skills to tasks.

Although Habakkuk (1962) had already postulated that labor scarcity encouraged innovation in the United States, Zeira (1998) has been the first to model automation in a growth model. Output is produced as an aggregate of intermediates, each of which can be produced with either a manual technology or a more capital-intensive automated technology. Exogenous technological progress in total factor productivity (TFP) raises the wage so that a growing number of intermediate producers adopt the industrial technology over time. Zeira (1998) then focuses on the role of automation in amplifying productivity differences across countries. Acemoglu (2010) shows that Habbakuk's (1962) hypothesis only holds if innovation is labor saving. Peretto & Seater (2013) build a dynamic model in which innovation in automation changes the exponent of an aggregate Cobb–Douglas production function. Yet, none of these papers features DTC because they all consider only one type of innovation.

4.1. Automation and Nonbalanced Growth

A recent literature, starting with Hémous & Olsen (2021), has explicitly built task models into DTC frameworks. Hémous & Olsen's (2021) model endogenizes several aspects of the automation process described by Zeira (1998) and provides some answers to Acemoglu & Autor's (2011) critique of the DTC literature. Hémous & Olsen (2021) build on the expanding variety model (Romer 1990) and consider an economy in which a final good is produced as a CES aggregate of a mass N of products. This can be written as

$$Y = \left(\int_0^N y(i)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}},$$

with $\sigma > 1$ and with $y(i)$ being the use of product i. Each product is produced monopolistically using a generalization of Equation 13,

$$y(i) = \left(l(i)^{\frac{\epsilon-1}{\epsilon}} + \alpha(i)x(i)^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon\beta}{\epsilon-1}} h(i)^{1-\beta}, \qquad 14.$$

where $l(i)$ denotes low-skill workers and $h(i)$ denotes high-skill workers. These labor inputs correspond to different tasks, so that each product comes with its own tasks. Automation occurs when machines can be used to (partly) substitute for low-skill labor in a task and $\alpha(i)$ switches from 0 to 1. In the baseline model, machines are produced one-for-one with the final good.[23] Technology is characterized by both the number of products, N_t, and the share of automated products (or low-skill tasks), G_t. Automation is therefore a secondary innovation that occurs in product lines developed through horizontal innovation. This makes Hémous & Olsen (2021) closer in spirit to Aghion & Howitt (1996).

Aggregate output can be represented as

$$Y = N^{\frac{1}{\sigma-1}} \left(\{1-G\}^{\frac{1}{\sigma}} \Big\{ \underbrace{(L^{\mathrm{N}})^{\beta} (H^{\mathrm{P,N}})^{1-\beta}}_{T_1} \Big\}^{\frac{\sigma-1}{\sigma}} + G^{\frac{1}{\sigma}} \Big\{ \underbrace{[(L^{\mathrm{A}})^{\frac{\epsilon-1}{\epsilon}} + (X)^{\frac{\epsilon-1}{\epsilon}}]^{\frac{\epsilon\beta}{\epsilon-1}} (H^{\mathrm{P,A}})^{1-\beta}}_{T_2} \Big\}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{15}$$

where L^{A} (L^{N}) denotes the total mass of low-skill workers in automated (nonautomated) firms, $H^{\mathrm{P,A}}$ ($H^{\mathrm{P,N}}$) denotes the total mass of high-skill workers hired in production in automated (nonautomated) firms, and $X = \int_0^N x(i)di$ denotes the total use of machines. The first term, T_1, captures the case in which production is nonautomated. The second term, T_2, represents the factors used within automated products and features substitutability between low-skill labor and machines. Equation 15 shows that G is not a factor-augmenting technology but rather the share parameter of the automated products nest. $N^{\frac{1}{\sigma-1}}$ plays the role of a TFP parameter.

Hémous & Olsen (2021) assume that $\mu = \beta(\sigma-1)/(\epsilon-1) < 1$, which ensures that automation is low-skill-labor saving at the firm level.[24] Yet, this does not necessarily imply that automation is low-skill-labor saving for the aggregate economy. For given N and G, the static equilibrium can be described as the intersection of two equations, illustrated in **Figure 1**, in which w_{L} and w_{H} denote the wages of low- and high-skill workers, respectively. The unit isocost curve represents the cost of producing one unit of final good and is a downward sloping curve in the (w_{L}, w_{H}) space. The relative demand curve is upward sloping.[25] When there is no automation ($G = 0$), the aggregate economy inherits the Cobb–Douglas structure and the relative demand curve is linear. With $G > 0$, the curve bends upwards: Automated firms can substitute more toward machines, and nonautomated firms lose market size when w_{L} rises. An increase in automation pivots this curve counterclockwise, which reduces low-skill wages, a negative aggregate substitution effect. It also increases the productive capabilities of the economy, which pushes the isocost curve toward the upper right, a positive aggregate scale effect. Consequently, while high-skill wages and the skill premium are always increasing in G, the effect on low-skill wages is ambiguous. In particular, for low G, the aggregate scale effect dominates and low-skill wages rise in G, but for higher levels of automation, the effect is negative (an analogous point is made informally in Autor 2014). Horizontal innovation, an inflow of nonautomated products, raises both low- and high-skill wages and, for high-enough N, lowers the skill premium. **Figure 1** illustrates a central feature of Hémous & Olsen's (2021) model. For any path of technology $[N_t, G_t]_{t=0}^{\infty} \in (0, \mathbb{R}^{+}) \times (0, 1)$, where $\lim_{t\to\infty} N_t = \infty$ and G_t has a strictly positive limit, low-skill wages must grow at a positive but

[23]In contrast to Autor et al.'s (2003) model, in this case the price of an existing machine is fixed, but there is technological progress insofar as machines can be used in a growing number of tasks.

[24]Automation both lowers the cost of production, which increases demand for low-skill labor, and allows for substitution with machines, which lowers it. The latter effect dominates when $\mu < 1$.

[25]The unit isocost is given by $\frac{\sigma N^{\frac{1}{1-\sigma}} \left(G(1+w_{\mathrm{L}}^{1-\epsilon})^{\mu} + (1-G) w_{\mathrm{L}}^{\beta(1-\sigma)} \right)^{\frac{1}{1-\sigma}} w_{\mathrm{H}}^{1-\beta}}{(\sigma-1)\beta^{\beta}(1-\beta)^{1-\beta}} = 1$, and the relative demand curve is given by $\frac{w_{\mathrm{H}} H}{w_{\mathrm{L}} L} = \frac{1-\beta}{\beta} \frac{G(1+w_{\mathrm{L}}^{\epsilon-1})^{\mu} + (1-G)}{G(1+w_{\mathrm{L}}^{\epsilon-1})^{\mu-1} + (1-G)}$.

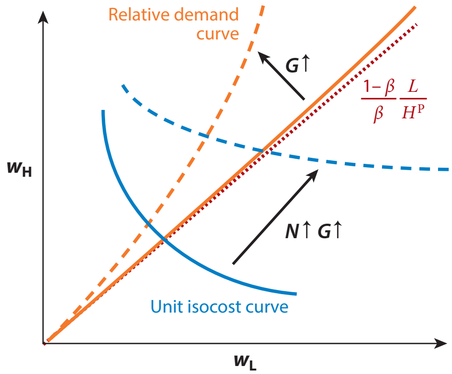

Figure 1

The static equilibrium in Hémous & Olsen's (2021) model. In the graph, w_H denotes the high-skill wage, w_L the low-skill wage, N the number of products, G the share of automated products, L the number of low-skill workers, H^P the number of high-skill workers in production, and $1-\beta$ the factor share of high-skill labor in production. An increase in the number of products raises both wages and the skill premium when $G > 0$. An increase in the share of automated products increases high-skill wages and the skill premium but has an ambiguous effect on low-skill wages. Figure adapted with permission from Hémous & Olsen (2021); copyright 2021 American Economic Association.

lower rate than high-skill wages. This will make a BGP with equal growth in low- and high-skill wages impossible.

Hémous & Olsen (2021) endogenize both N_t and G_t. Horizontal innovation uses high-skill workers; that is, $\dot{N}_t = \gamma N_t H_t^D$. Moreover, a monopolist of a nonautomated firm can hire high-skill workers, h_t^A, to automate the firm's production technology with Poisson rate $\eta G_t^{\tilde{\kappa}}\left(N h_t^A\right)^{\kappa}$, where $\kappa \in (0,1)$ controls the curvature of the innovation function and $\tilde{\kappa} \in [0,\kappa]$ parameterizes a knowledge externality in automation innovation. This gives the law of motion

$$\dot{G}_t = \eta G_t^{\tilde{\kappa}}(h_t^A N_t)^{\kappa}(1-G_t) - G_t g_t^N, \qquad 16.$$

where $g_t^N = \dot{N}_t/N_t$. Equation 16 has strong similarities with a capital accumulation function, and the stock of automated tasks can be seen analogously: Automation of existing tasks accumulates automation stock, and the inflow of (not-yet-)automated tasks depreciates the existing automation stock. Both of these innovation processes respond to economic incentives. The resulting state of technology, (N_t, G_t), then determines wages in the economy.

Let V_t^A and V_t^N denote the value functions of automated and nonautomated firms, respectively. Nonautomated firms employ high-skill workers to automate the production process, implying the following first-order condition for automation innovation:

$$\eta\kappa G_t^{\tilde{\kappa}}\left(N_t h_t^A\right)^{\kappa-1}\left(V_t^A - V_t^N\right) = w_{Ht}/N_t. \qquad 17.$$

The number of automation innovations therefore depends on the ratio between the increase in firm value associated with automation and its effective cost, or

$$\frac{V_t^A - V_t^N}{w_{Ht}/N_t} \tilde{\propto} \frac{\pi_t^A - \pi_t^N}{G_t\pi_t^A + (1-G_t)\pi_t^N} = \frac{1-\left(1+w_{Lt}^{\epsilon-1}\right)^{-\mu}}{G_t + (1-G_t)\left(1+w_{Lt}^{\epsilon-1}\right)^{-\mu}}, \qquad 18.$$

where $\tilde{\propto}$ means approximately proportional and π_t^A and π_t^N are the profits of automated and nonautomated firms. Intuitively, with a positive discount rate, $V_t^A - V_t^N$ moves with $\pi_t^A - \pi_t^N$ to a first approximation. Furthermore, since both aggregate profits and high-skill labor compensations are proportional to output, w_{Ht}/N_t is proportional to average profits. The second half of the equation follows from $\pi_t^{N}/\pi_t^{A} = (1 + w_L^{\epsilon-1})^{-\mu}$. This highlights low-skill wages as the key determinant of automation innovations. When $w_{Lt} \approx 0$, there is little advantage to being automated and $\pi_t^{A} \approx \pi_t^{N}$, which implies little automation innovation. When $w_{Lt} \to \infty$, the right-hand side of Equation 18 approaches a constant and, with it, the share of innovation in automation technologies. In contrast to the classic DTC model with factor-augmenting technical change, in this model the direction of technical change is entirely determined by a price effect with no market size effect (as in Acemoglu & Restrepo 2018, discussed below). Intuitively, horizontal innovation and automation affect the same market. This price effect bears similarity to Zeira's (1998) model, in which the adoption of a labor-saving technology also depends on the price of labor.

Hémous & Olsen (2021) show that this economy cannot feature a BGP with equal growth in low- and high-skill wages. An economy starting with low N_t, and consequently low w_{Lt} and automation, is initially close to a BGP, growing just through horizontal innovation. As low-skill wages grow, so does the incentive to automate, and the economy endogenously shifts toward automation innovation. This leads to an endogenous rise in the skill premium and a decline in the labor share, as experienced in the United States since the 1980s. Eventually, the model features an asymptotic steady state in which G_t is constant and all wages grow, but the skill premium continues to grow. Hémous & Olsen (2021) calibrate an extended version of the model in which machines belong to a capital stock to the data. The model can replicate quantitatively (and endogenously) the evolution of the skill premium, the labor share, and automation and productivity growth in the United States from 1963 to 2012.

4.2. Automation and Balanced Growth

Acemoglu & Restrepo (2018) also consider DTC in a task model but reach sharply different conclusions. In contrast to Hémous & Olsen (2021), they include in their model a unit measure of task, and output obeys

$$Y = \left(\int_{N-1}^{N} y(i)^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}. \tag{19}$$

For this review, we restrict $\sigma > 1$. Therefore, a new task N replaces an old, now obsolete task, $N - 1$. Tasks are produced monopolistically using the production function

$$y(i) = \alpha(i)k(i) + \gamma(i)l(i), \tag{20}$$

where $k(i)$ is the use of capital in the production of task i, $l(i)$ is the use of labor, and $\alpha(i) \in \{0, 1\}$ is the automation index. In this function, $\gamma(i) = e^{Ai}$ is the labor-augmenting productivity of task i, where $A > 0$: This means that new tasks feature higher labor productivity, though, once automated, all tasks have the same (capital) productivity.

In the DTC case, automation is costly, so that all automated firms use machines. With $\gamma(i)$ increasing in i, there is a threshold I such that tasks below I are automated $[\alpha(i) = 1]$ and sold at price $p(i) = \sigma/(\sigma - 1)R$, where R is the gross return on capital. In contrast, tasks in $[I, N]$ are nonautomated $[\alpha(i) = 0]$ and sold at price $p(i) = \sigma/(\sigma - 1) \times W/\gamma(i)$, where W is the real wage. Using Equations 19 and 20, and factor-clearing conditions for K and L, one gets the aggregate production function

$$Y = \left\{ [I - (N-1)]^{\frac{1}{\sigma}} K^{\frac{\sigma-1}{\sigma}} + \left(\int_{I}^{N} \gamma(i)^{\sigma-1} di \right)^{\frac{1}{\sigma}} L^{\frac{\sigma-1}{\sigma}} \right\}^{\frac{\sigma}{\sigma-1}}. \tag{21}$$

Equation 21 demonstrates that technology, characterized by N and I, determines the factor shares. Hence, as in Hémous & Olsen's (2021) model, technology is not factor augmenting.[26]

For given factors K and L, W/R depends positively on the introduction of new nonautomated products, N, and negatively on the automation of existing products, I. An increase in N always increases the absolute level of wages. As in Hémous & Olsen's (2021) model, an increase in automation has an ambiguous effect on wages due to the combination of a scale effect and a substitution effect—which are called productivity effect and displacement effect, respectively, by Acemoglu & Restrepo (2018). The latter effect may dominate and leave automation to be labor saving. This occurs in particular when $R \approx W/\gamma(i)$, i.e., when the cost savings of automation are relatively low, a situation deemed a "so-so automation" by Acemoglu & Restrepo (2019a). Acemoglu & Restrepo (2019a, 2020a) argue that many modern innovations in automation have this feature and correspondingly are labor saving.[27]

Following this, Acemoglu & Restrepo (2018) endogenize capital through a standard Ramsey setting. Innovation is undertaken by scientists who are in fixed supply S and either develop new tasks or automate existing tasks, so that $\dot{N}(t) = \kappa_N S_N(t)$ and $\dot{I}(t) = \kappa_I S_I(t)$, where S_N and S_I are the respective numbers of scientists. Acemoglu & Restrepo assume that innovators must compensate the previous incumbent. Here we focus on BGPs in which both types of innovation are active. This gives the following value functions of automating a task and introducing a new one, respectively,

$$V_I(t) = Y(t) \int_t^{\infty} e^{-\int_t^{\tau} \left(R(s)-\delta-g_Y(s)\right)ds} \left(\pi_I(\tau) - \pi_N(\tau, I(t))\right) d\tau \text{ and} \tag{22.}$$

$$V_N(t) = Y(t) \int_t^{\infty} e^{-\int_t^{\tau} \left(R(s)-\delta-g_Y(s)\right)ds} \left(\pi_N(\tau, N(t)) - \pi_I(\tau)\right) d\tau, \tag{23.}$$

where $g_Y(t) = \dot{Y}(t)/Y(t)$, $R(s) - \delta$ is the return to capital net of depreciation, $\pi_N(t, i) = \sigma^{-1}(W(t)/\gamma(i))^{1-\sigma} Y(t)$ denotes the profits of nonautomated tasks/products, and $\pi_I(t) = \sigma^{-1} R(t)^{1-\sigma} Y(t)$ denotes the profits of automated tasks/products. Equation 22 reflects the total discounted value of being a monopolist of an automated task minus the compensation to the previous nonautomated monopolist who operated task $I(t)$ at the time of automation, t. Equation 23 reflects the analogous expression for a monopolist developing the new nonautomated task $N(t)$. Therefore, the value function of a new automated task, $V_I(t)$, depends positively on (the path of) $W(t)/R(t)$, and the value function of a new nonautomated task, $V_N(t)$, depends negatively on it.

Combining this insight with the fact that W/R depends positively on $N - I$, Acemoglu & Restrepo (2018) demonstrate that under appropriate regularity conditions (on ρ and κ_I/κ_N), a BGP exists in which $\kappa_I V_I = \kappa_N V_N$, and all variables—$W(t)$, $K(t)$, $Y(t)$, $V_I(t)$, and $V_N(t)$—grow at the same rate while $R(t)$ is constant.[28] Further, N and I grow at the same rate, such that the share of products

[26]In fact, one can write the Equation 21 as $Y = \left[(AK)^{\frac{\sigma-1}{\sigma}} + (BL)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$, where $A \equiv (I - N + 1)^{\frac{1}{\sigma-1}}$ and $B \equiv (\int_I^N \gamma(i)^{\sigma-1} di)^{\frac{1}{\sigma-1}}$. When $\sigma > 1$, automation, I, increases A and decreases B and can therefore be seen as a combination of capital-augmenting and labor-depleting technical change (see also Aghion et al. 2019b).

[27]The aggregate substitution effect of Hémous & Olsen (2021) is derived for an endogenous use of machines as intermediate inputs, whereas Acemoglu & Restrepo (2018) hold the stock of capital constant. This distinction is made explicit by Acemoglu & Restrepo (2019a), who refer to the endogenous response of capital as a capital-deepening effect. Acemoglu & Restrepo (2020a) find that robotization leads to a decline in employment, which suggests that the aggregate substitution effect dominates the scale effect.

[28]The incentive for automation innovation in Acemoglu & Restrepo's (2018) model depends on the relative cost of labor and capital, whereas it only depends on low-skill wages in Hémous & Olsen's (2021) model. This difference arises from the assumption that machines are intermediate inputs in the baseline model by Hémous & Olsen but is immaterial. In fact, when Hémous & Olsen take the model to the data, machines are part of the capital stock, and the incentive to automate similarly depends on the relative cost of low-skill labor to capital.

that are automated, $1 - N + I$, as well as the factor shares, is constant. Intuitively, new tasks come with a higher labor productivity, so that the effective amount of labor grows at the same rate as capital. In their full-fledged model, Acemoglu & Restrepo (2018) include heterogeneous effort costs for scientists working in each type of innovation, which ensures that the allocation of scientists varies smoothly. They demonstrate that this BGP is locally stable, such that a small perturbation in the stock of automation, I, above its BGP path reduces $W(t)/R(t)$, relatively discourages automation, and correspondingly brings the economy back to the BGP level of automation.

Therefore, Acemoglu & Restrepo (2018) demonstrate that one can build an economic model with tasks and automation that replicates the features of a model with purely labor-augmenting technical change. In their model, workers' role in the economy remains undiminished despite the continued presence of automation. This is in sharp contrast to Hémous & Olsen's (2021) model, which shows that no BGP is possible and the role of low-skill workers in the economy must diminish as the economy grows. Therefore, the two models present distinct views on the development of the US economy over the past decades. Seen through Acemoglu & Restrepo's (2018) lenses, the recent decline in the labor share in the United States must come from factors outside of the model. Acemoglu & Restrepo (2019b) compare the US economy before and after 1987 and argue that the latter period features lower overall productivity growth. They ascribe this shift to tax advantages for equipment compared to labor, to a popular focus on automation, and to declining government support for innovation, which otherwise tends to favor the creation of new tasks. In contrast, seen through Hémous & Olsen's (2021) lenses, the recent development of the US economy is simply consistent with the fact that automation endogenously and gradually increases as an economy matures. The extent to which the increase in automation innovation reflects the endogenous development of an economy or rather shocks and policy changes is an important issue for future research.

4.3. Other Models of Automation

The distinct theoretical predictions of Acemoglu & Restrepo (2018) and Hémous & Olsen (2021) arise because of different assumptions on labor-augmenting technical progress in new products. To demonstrate this, consider a combination of the two approaches by replacing Equation 20 with

$$y(i) = \left[b(i)l(i) + \alpha(i)b(i)^{\varsigma}x(i)\right]^{\beta}\left[b(i)h(i)\right]^{1-\beta}, \qquad 24.$$

where variables are as defined by Hémous & Olsen (2021), and $b(i)$ represents a technology level. We are interested in whether low- and high-skill wages can grow at the same rate for an economy with an asymptotic BGP. With this in mind, we consider exogenous technical change in which $b(i) = \exp(Bi)$ and $N_t = nt$ for some $n > 0$, and in which products are automated at a constant Poisson rate. The parameter $\varsigma \in [0, 1]$ reflects factor-augmenting technical change for machines, where $\varsigma = 0$ ensures only labor-augmenting technical change, in the spirit of Acemoglu & Restrepo's model. Aggregate output continues to be given by a standard CES production function using all products (N_t). The stocks of low-skill and high-skill labor are exogenously given, whereas machines are produced one-for-one with the final good. In the **Supplemental Appendix**, we demonstrate that only for $\varsigma = 0$ will low- and high-skill wages grow at the same rate asymptotically. When $\varsigma > 0$, low-skill wages must grow slower. This result is in the spirit of the Uzawa theorem but differs insofar as it refers to technological progress from one product to another rather than to aggregate technological progress.

Supplemental Material >

A related result is demonstrated by Ray & Mookherjee (2020). In a general framework with both capital (complementary to labor) and robots (substitutes for labor), they demonstrate that under general conditions, an economy that grows through capital accumulation must eventually have a labor share going toward zero, although wages might be growing asymptotically as low-skill wages do in Hémous & Olsen's (2021) model. They extend their model to include DTC, which permits but does not require technological change to be labor augmenting as in Acemoglu & Restrepo's (2018) model. They show that, asymptotically, capital-augmenting technical change will be at least as rapid as labor-augmenting technical change, and consequently the growth in labor income will still be lower than the growth of the overall economy.

Neither Hémous & Olsen (2021) nor Acemoglu & Restrepo (2018) draw clear inferences about whether the direction of technology is efficient. More interestingly, Acemoglu & Restrepo (2019b) argue that innovation is inefficiently directed toward automation because the benefits from automation may be more front-loaded (analogously to dirty innovation in Acemoglu et al. 2012). Acemoglu & Restrepo (2020b) argue that this is particularly the case for artificial intelligence. This could be easily microfounded in Hémous & Olsen's (2021) framework by assuming that automation can also be undertaken by entrants. Since the profits realized once a good is automated partly motivate the creation of new intermediates, the incentive for horizontal innovation will be lower if there is a risk that another firm will reap those profits. In this sense, the returns to horizontal innovation are back-loaded relative to the returns to automation, a situation akin to that of clean versus dirty innovation in Acemoglu et al.'s (2012) model.

In contrast with the models described above, Aghion et al. (2019b) model automation in a task framework when the different tasks are complementary but ignore other innovations (in Equation 19, this means that $\sigma < 1$ and N is fixed). Automation still allows for the use of machines in place of workers in a given task, but with $\sigma < 1$, it is now equivalent to labor-augmenting technical change combined with capital-depleting technical change (see footnote 26). They show that there is a path to automation that is nearly consistent with balanced growth. It would be interesting to see whether this proposition can be reconciled with endogenous innovation. Prettner & Strulik (2020) model automation as the creation of additional varieties of machines that are imperfect substitutes for labor (so that automation does not directly lead to the replacement of workers in existing tasks).[29]

Further, whereas both Acemoglu & Restrepo (2018) and Hémous & Olsen (2021) focus on the automation of goods production and have models in which the asymptotic growth rate is finite and constant, in a second model Aghion et al. (2019b) explicitly focus on the automation of the idea production function. They show that whether explosive growth happens depends on whether a model features increasing returns to accumulable factors; that is, the product of the extent to which each of the output and idea functions scales with the reproducible factors of technology and capital. For instance, in either Acemoglu & Restrepo's (2018) or Hémous & Olsen's (2021) model, if any of the tasks currently performed by high-skill workers/scientists in the development of new products/tasks was automatable, the models would feature explosive growth.

Nakamura & Zeira (2018) study the effects of automation on unemployment in a model related to that of Hémous & Olsen, including general processes of automation and horizontal innovation. Automated tasks can be produced solely by capital, but to allow for a BGP, they require the productivity of capital to be lower for newer tasks. They consider a small open economy with free capital movement and assume that workers who are displaced by automation stay unemployed for a fixed period of time before finding a new job. On their (asymptotic) BGP, the fraction of tasks performed by humans that is automated in every period declines toward zero, and with it

[29]In a model with exogenous technical change, Martinez (2019) also derives an aggregate CES production function from a microfoundation of automation. In a cross-industry analysis, he finds evidence that automation was a driving force behind the recent decline of the US labor share.

unemployment. Casey (2018) also develops a model that features technological unemployment in equilibrium. In his model with DTC, innovation might speed up both labor productivity growth and unemployment.

This new DTC literature is still in its infancy, and more research can and should be done, particularly to analyze the policy implications of DTC.

5. EMPIRICAL EVIDENCE

5.1. Environmental Economics

A large empirical literature has looked for evidence of induced technical change in environmental economics. Popp et al. (2010) and Popp (2019) provide extensive literature reviews, and here we mainly focus on a few recent papers.

Newell et al. (1999) show that technical change in air conditioners was biased against energy efficiency in the 1960s, when energy prices were low, but this bias reversed after the energy shocks of the 1970s. Most of the early literature uses macro data in contexts in which identification is difficult. In a seminal paper, Popp (2002) uses time-series data on US patents and finds a long-run elasticity of energy efficiency innovation with respect to energy prices of 0.35. In a panel of US manufacturing industries, Brunnermeier & Cohen (2003) find that environmental patents increase following an increase in pollution abatement expenditures. In a panel of Organisation for Economic Co-operation and Development (OECD) countries, Jonstone et al. (2010) find that public policies affect innovation in renewable energy, with broad policies (such as a renewable mandate) being more effective for technologies that are closer competitors of fossil fuels (namely wind, in their sample). Technologies farther from the market (e.g., solar) require more targeted subsidies. Such results are consistent with Acemoglu et al.'s (2012) framework.[30]

Aghion et al. (2016) go further by presenting firm-level evidence. They focus on the car industry from 1978 to 2005 and distinguish between clean patents (associated with electric, hybrid, and hydrogen engines) and dirty patents (associated with fossil fuel engines). To measure the effect of fuel prices on innovation at the firm level, they take advantage of the fact that innovators in the car industry sell in several national markets to build a firm-specific fuel price. This fuel price is computed as a weighted average of country-level fuel prices, where the firm-specific weights are computed using a firm's patent history pre-sample (as a proxy for a firm's market shares).[31] In the spirit of a shift-share instrument, the effect of fuel price on firms' innovation is identified through the differential impact across firms of cross-country variations in fuel prices (or taxes) depending on firms' exposure to international markets. Aghion et al. (2016) estimate a large positive effect of fuel prices on clean innovation, with an elasticity close to 1, and a negative effect on dirty innovation, with an elasticity close to -0.5.[32] Furthermore, they find evidence of path dependence. Through simulations, they show that, in line with Acemoglu et al.'s (2012) model, path dependence exacerbates the gap between clean and dirty knowledge in business-as-usual conditions, but

[30] Readers are also referred to Verdolini & Galeotti (2011), who include knowledge spillovers, and Dechezleprêtre & Glachant (2011), who separate domestic and foreign policies.

[31] Since a patent only protects an invention in the country in which it is applied for, whether a firm decides to apply for a patent in a given country or not reflects how important this country is for the firm. Coelli et al. (2020) show empirically that this is a good proxy for market share.

[32] In line with these results, Knittel (2011) finds that there is a trade-off between improved fuel efficiency and other vehicle attributes, and that technical progress has responded to the implementation of regulatory standards.

it actually reduces the increase in fuel prices necessary to induce clean technologies to catch up with dirty ones.

Several papers have used the same method to obtain variation at the firm level. Noailly & Smeets (2015) study how clean and dirty innovations in electricity production respond to both fuel prices and market size. Overall, their results support the DTC hypothesis: An increase in renewable market size or fossil fuel prices increases renewable innovation, and a larger fossil fuel market leads to more fossil fuel innovation. Surprisingly, an increase in fossil fuel prices also leads to a large increase in fossil fuel innovation, but it is an increase in energy-efficiency innovations that drives this. Their results also support path dependency (see also Lazkano et al. 2017, Lööf et al. 2018).

Using different identification strategies, other recent papers measure the direct effect of environmental policies on innovation using micro data. Calel & Dechezleprêtre (2016) show that the European Union Emission Trading System (EU ETS) has increased low-carbon innovation by 10% in regulated firms. To establish this result, they take advantage of the existence of regulatory thresholds at the plant level and follow a matched difference-in-differences strategy whereby they compare regulated firms with unregulated firms of the same size. Calel (2020) finds similar results. Dugoua (2020) evaluates the effect of international environmental agreements on innovation. She focuses on the Montreal Protocol, which has regulated the use of chlorofluorocarbons (CFCs) since 1989, and finds that it led to an increase of 4,000% in patents pertaining to CFC substitutes relative to similar molecules. Howell (2017) exploits the fact that the US Department of Energy allocates R&D grants to small businesses through a grading scheme. Using a regression discontinuity analysis, she finds that receiving a grant increases patenting, survival rate, and venture capital, with stronger effects for firms likely to be more financially constrained.

Having established the empirical existence of DTC from price and market size effects, the literature is moving to study other factors driving technical change as well as interaction effects. For instance, Aghion et al. (2020) extend the setup of Aghion et al. (2016) to study the roles of both consumer value and competition in driving innovation in the car industry. They find that when consumers value the environment more, clean innovation in the car industry increases, particularly when competition is more intense. They estimate that the simultaneous increase in environmental valuation and competition that happened in 1998–2002 and 2008–2012 had the same effect on innovation as a 40% increase in fuel prices.

5.2. Labor Economics

The empirical literature on DTC in labor economics is comparatively smaller. A few papers show that labor market conditions affect labor-saving technology adoption in health care (Acemoglu & Finkelstein 2008), agriculture (Hornbeck & Naidu 2014, Clemens et al. 2018), and manufacturing (Lewis 2011). Both Lordan & Neumark (2018) and Aaronson & Phelan (2019) look at the consequences of minimum wage hikes on routine jobs. Acemoglu & Restrepo (2019c) find that aging of the population is associated with greater adoption of robots and other automation technologies in cross-country regressions. Further, this effect is stronger in industries relying more on middle-skill workers in industry $\times$ country-level regressions. More importantly for the purpose of this review, they also find a positive correlation between population aging and patenting in robotics. Alesina et al. (2018) find that across countries, labor market regulations are positively correlated with innovation in low-skill-intensive sectors.

A few recent papers use micro evidence. Dechezleprêtre et al. (2019) develop a new classification of patents in the machinery sector as automation or nonautomation by combining information on patent texts and technological classes. They build on the empirical strategy of Aghion et al. (2016), taking advantage of the market structure for most innovation in automation technology:

Innovation is highly concentrated in a few large companies, which sell their technology to other (typically manufacturing) firms around the world. Consequently, the demand for automation, and with it the incentive to innovate, is determined by the wages paid by these potential customers. They compute a proxy for the low- and high-skill wages paid by these customers by taking a weighted average of country-level wages, where the weights are calculated using the geographical dispersion of patents pre-sample. They find a large positive effect of low-skill wages on automation innovations, with an elasticity between 2 and 4. In line with capital-skill complementarity, high-skill wages tend to reduce automation innovations. In contrast, wages do not have a significant effect on nonautomation innovations in machinery. Moreover, they show that the Hartz reforms, a series of reforms implemented in Germany between 2003 and 2005 to increase labor market flexibility, led to a relative decrease in automation innovations in non-German firms more exposed to Germany. Relatedly, Bena & Simintzi (2019) attempt to distinguish between process and product innovations in patent data and find that firms with better access to the Chinese labor market decreased their share of process innovations after the 1999 US–China trade agreement. Note that process and automation innovations may overlap but are distinct concepts.

Several papers use immigration to relate labor scarcity to innovation. San (2019) shows that following the exclusion of Mexican seasonal agricultural workers, patenting increased for crops that rely more on agricultural labor. Danzer et al. (2020) rely on the regional allocation in Germany of ethnic German migrants from the collapsing Soviet Union. They also classify patents as automation or nonautomation and find that regions receiving more immigrants developed fewer automation patents. Andersson et al. (2020) use an instrumental variable strategy to show that Swedish emigration to the United States led to higher wages and innovation in the most affected municipalities. They do not, however, look at the direction of innovation. In contrast, Doran & Yoon (2020) find that innovation decreased in the US cities most affected by the 1920 immigration quotas, which reduced immigration from Southern and Eastern Europe. These perhaps contradictory results highlight the fact that analyzing the effect of labor scarcity on innovation requires a distinction between different forms of innovation.

6. CONCLUSIONS AND FUTURE RESEARCH

The literature has established that innovation responds strongly to market incentives and that its endogeneity matters for macroeconomic outcomes. The development of several COVID-19 vaccines in less than a year provides another example in a different context. The original DTC framework of Acemoglu (1998) has been successfully applied in various contexts. A recent literature has developed new DTC task models to analyze automation. A potential avenue for future research is to use these new models in other contexts, notably environmental economics.

Our review has identified two important questions to be asked. First, is the economy on a BGP? Should this not be the case, DTC can account for path dependence in energy technologies in environmental models and for growing income inequality in labor models. In contrast, on a BGP, an economy would revert to the same path after a shock. Testing for the existence of a BGP would be a complex but rewarding empirical endeavor.

Second, is the gap between the private and social returns of innovation the same for all technologies? The answer to this question determines whether industrial innovation policies are called for. In the environmental context, Acemoglu et al. (2012) and the literature that followed provide a strong case for a green innovation policy: Climate policy should be designed with innovation at the forefront. The question is more open in the labor context: Should automation be encouraged or hindered? Future research should delve deeper into this important issue, in particular because the DTC labor literature has paradoxically sidelined the distributional aspects of innovation policy.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement 805007). M.O. received funding from the Rockwool Foundation (Globalization, New Technologies and the Labor Market project).

LITERATURE CITED

Aaronson D, Phelan B. 2019. Wage shocks and the technological substitution of low-wage jobs. *Econ. J.* 129:1–34

Acemoglu D. 1998. Why do new technologies complement skills? Directed technical change and wage inequality. *Q. J. Econ.* 113:1055–89

Acemoglu D. 2002. Directed technical change. *Rev. Econ. Stud.* 69:781–809

Acemoglu D. 2003. Labor- and capital-augmenting technical change. *J. Eur. Econ. Assoc.* 1:1–37

Acemoglu D. 2007. Equilibrium bias of technology. *Econometrica* 75:1371–409

Acemoglu D. 2010. When does labor scarcity encourage innovation? *J. Political Econ.* 118:1037–78

Acemoglu D, Aghion P, Barrage L, Hémous D. 2019. *Climate change, directed innovation, and energy transition: the long-run consequences of the shale gas revolution.* Work. Pap., Mass. Inst. Technol., Cambridge

Acemoglu D, Aghion P, Bursztyn L, Hémous D. 2012a. The environment and directed technical change. *Am. Econ. Rev.* 102:131–66

Acemoglu D, Aghion P, Hémous D. 2014. The environment and directed technical change in a North–South model. *Oxf. Rev. Econ. Policy* 30:513–30

Acemoglu D, Akcigit U, Hanley D, Kerr W. 2016. The transition to clean technology. *J. Political Econ.* 124:52–104

Acemoglu D, Autor D. 2011. Skills, tasks, and technologies: implications for employment and earnings. In *Handbook of Labor Economics*, Vol. 4, ed. O Ashenfelter, D Card, pp. 1043–171. Amsterdam: Elsevier

Acemoglu D, Finkelstein A. 2008. Input and technology choices in regulated industries: evidence from the health care sector. *J. Political Econ.* 116:837–80

Acemoglu D, Gancia G, Zilibotti F. 2012b. Competing engines of growth: innovation and stardardization. *J. Econ. Theory* 147:570–601.e3

Acemoglu D, Rafey W. 2019. *Mirage on the horizon: geoengineering and carbon taxation without commitment*. NBER Work. Pap. 24411

Acemoglu D, Restrepo P. 2018. The race between machine and man: implications of technology for growth, factor shares and employment. *Am. Econ. Rev.* 108:1488–542

Acemoglu D, Restrepo P. 2019a. Artificial intelligence, automation, and work. In *The Economics of Artificial Intelligence: An Agenda*, ed. A Agrawal, J Gans, A Goldfarb, pp. 197–236. Chicago: Univ. Chicago Press

Acemoglu D, Restrepo P. 2019b. Automation and new tasks: how technology displaces and reinstates labor. *J. Econ. Perspect.* 33:3–30

Acemoglu D, Restrepo P. 2019c. *Demographics and automation*. Work. Pap., Mass. Inst. Technol., Cambridge

Acemoglu D, Restrepo P. 2020a. Robots and jobs: evidence from US labor markets. *J. Political Econ.* 128:6. **https://doi.org/10.1086/705716**

Acemoglu D, Restrepo P. 2020b. The wrong kind of AI? Artificial intelligence and the future of labour demand. *Camb. J. Reg. Econ. Soc.* 13:25–35

Acemoglu D, Restrepo P. 2020c. Unpacking skill bias: automation and new tasks. *Am. Econ. Assoc. Pap. Proc.* 110:356–61

Acemoglu D, Zilibotti F. 2001. Productivity differences. *Q. J. Econ.* 116:563–606

Aghion P, Bénabou R, Martin R, Roulet A. 2020. *Environmental preferences and technology choices: Is market competition clean or dirty*? NBER Work. Pap. 26921

Aghion P, Dechezleprêtre A, Hémous D, Martin R, Van Reenen J. 2016. Carbon taxes, path dependency, and directed technical change: evidence from the auto industry. *J. Political Econ.* 214:1. **https://doi.org/10.1086/684581**

Aghion P, Hepburn C, Teytelboym A, Zenghelis D. 2019a. Path dependence, innovation and the economics of climate change. In *Handbook on Green Growth*, ed. R Fouquet, pp. 67–83. Cheltenham, UK: Edward Elgar Publ.

Aghion P, Howitt P. 1992. A model of growth through creative destruction. *Econometrica* 60:323–51

Aghion P, Howitt P. 1996. Research and development in the growth process. *J. Econ. Growth* 1:49–73

Aghion P, Howitt P. 2009. *The Economics of Growth*. Cambridge, MA: MIT Press

Aghion P, Jones BF, Jones CI. 2019b. Artificial intelligence and economic growth. In *The Economics of Artificial Intelligence: An Agenda*, ed. A Agrawal, J Gans, A Goldfarb, pp. 237–90. Chicago: Univ. Chicago Press

Alesina A, Battisti M, Zeira J. 2018. Technology and labor regulations: theory and evidence. *J. Econ. Growth* 23:41–78

Andersson D, Karadja M, Prawitz E. 2020. Mass migration and technological change. SocArXiv, March. **https://doi.org/10.31235/osf.io/74ub8**

André F, Smulders S. 2014. Fueling growth when oil peaks: directed technological change and the limits to efficiency. *Eur. Econ. Rev.* 69:18–39

Autor D. 2014. Polanyi's paradox and the shape of employment growth. In *Economic Policy Proceedings: Reevaluating Labor Market Dynamics*, pp. 129–77. Kansas City, MO: Fed. Reserve Bank Kansas City

Autor D, Levy F, Murnane R. 2003. The skill content of recent technological change: an empirical exploration. *Q. J. Econ.* 118:1279–333

Barrett S. 2006. Climate treaties and "breakthrough" technologies. *Am. Econ. Rev.* 96:22–25

Bena J, Simintzi E. 2019. *Machines could not compete with Chinese labor: evidence from U.S. firms' innovation*. Work. Pap., Univ. B. C., Vancouver, Can.

Bovenberg AL, Smulders S. 1995. Environmental quality and pollution-augmenting technological change in a two-sector endogenous growth model. *J. Public Econ.* 57:369–91

Bovenberg AL, Smulders S. 1996. Transitional impacts of environmental policy in an endogenous growth model. *Int. Econ. Rev.* 37:861–93

Brunnermeier S, Cohen M. 2003. Determinants of environmental innovation in US manufacturing industries. *J. Environ. Econ. Manag.* 45:278–93

Calel R. 2020. Adopt or innovate: understanding technological responses to cap-and-trade. *Am. Econ. J. Econ. Policy* 12:170–201

Calel R, Dechezleprêtre A. 2016. Environmental policy and directed technical change: evidence from the European carbon market. *Rev. Econ. Stat.* 98:173–91

Casey G. 2018. *Growth, unemployment, and labor-saving technical change*. Unpublished manuscript, Williams Coll., Williamstown, MA

Casey G. 2019. *Energy efficiency and directed technical change: implications for climate change mitigation*. Dep. Econ. Work. Pap. 2019-17, Williams Coll., Williamstown, MA

Casey G, Horii R. 2019. *A multi-factor Uzawa growth theorem and endogenous capital-augmenting technological change*. ISER Discuss. Pap. 1051, Inst. Soc. Econ. Res., Osaka Univ., Osaka, Jpn.

Clemens M, Lewis E, Postel H. 2018. Immigration restrictions as active labor market policy: evidence from the Mexican bracero exclusion. *Am. Econ. Rev.* 108:1468–87

Coelli F, Moxnes A, Ulltveit-Moe KH. 2020. Better, faster, stronger: global innovation and trade liberalization. *Rev. Econ. Stat.* In press. **https://doi.org/10.1162/rest_a_00951**

Danzer A, Feuerbaum C, Gaessler F. 2020. *Labor supply and automation innovation*. IZA Discuss. Pap. 13429, Inst. Labor Econ., Bonn, Ger.

Dechezleprêtre A, Glachant M. 2011. Does foreign environmental policy influence domestic innovation? Evidence from the wind industry. *Environ. Resour. Econ.* 58:391–413

Dechezleprêtre A, Hémous D, Olsen M, Zanella C. 2019. *Automating labor: evidence from firm-level patent data*. CEPR Discuss. Pap. 14249, Cent. Econ. Policy Res., London

Di Maria C, Smulders SA. 2005. Trade pessimists versus technology optimists: induced technical change and pollution havens. *B.E. J. Econ. Anal. Policy* 3:1–27

Di Maria C, Valente S. 2008. Hicks meets Hotelling: the direction of technical change in capital-resource economies. *Environ. Dev. Econ.* 13:691–717

Dietz S, Lanz B. 2019. *Can a growing world be fed when the climate is changing?* IRENE Work. Pap. 19-09, IRENE Inst. Econ. Res., Cergy, Fr.

Doran K, Yoon C. 2020. *Immigration and invention: evidence from the quota acts*. Unpublished manuscript, Univ. Notre Dame, Notre Dame, IN

Dugoua E. 2020. *Induced innovation and international environmental agreements: evidence from the ozone regime*. Work. Pap., London Sch. Econ., London

Fischer C, Heutel G. 2013. Environmental macroeconomics: environmental policy, business cycles, and directed technical change. *Annu. Rev. Resour. Econ.* 5:197–210

Fried S. 2018. Climate policy and innovation: a quantitative macroeconomic analysis. *Am. Econ. J. Macroecon.* 10:9–118

Gans J. 2012. Innovation and climate change policy. *Am. Econ. J. Econ. Policy* 4:125–45

Gars J, Olovsson C. 2019. Fuel for economic growth? *J. Econ. Theory* 184:104941

Gerglagh R, Kverndokk S, Rosendahl KE. 2009. Optimal timing of climate change policy: interaction between carbon taxes and innovation externalities. *Environ. Resour. Econ.* 43:369–90

Gerlagh R, Kverndokk S, Rosendahl KE. 2014. The optimal time path of clean energy R&D policy when patents have finite lifetime. *J. Environ. Econ. Manag.* 67:2–19

Gerlagh R, Lise W. 2005. Carbon taxes: a drop in the ocean, or a drop that erodes the stone? The effect of carbon taxes on technological change. *Ecol. Econ.* 54:241–60

Goldin C, Katz L. 2008. *The Race Between Education and Technology*. Cambridge, MA: Harvard Univ. Press

Goulder LH, Schneider SH. 1999. Induced technological change and the attractiveness of CO2 abatement policies. *Resour. Energy Econ.* 21:211–53

Greaker M, Heggedal TR, Rosendahl KE. 2018. Environmental policy and the direction of technical change. *Scand. J. Econ.* 120:1100–38

Grimaud A, Rouge L. 2008. Environment, directed technical change and economic policy. *Environ. Resour. Econ.* 41:439–63

Habakkuk J. 1962. *American and British Technology in the Nineteenth Century*. Cambridge, UK: Cambridge Univ. Press

Harstad B, Lancia F, Russo A. 2019. Compliance technology and self-enforcing agreements. *J. Eur. Econ. Assoc.* 17:1–29

Hart R. 2004. Growth, environment and innovation—a model with production vintages and environmentally oriented research. *J. Environ. Econ. Manag.* 48:1078–98

Hart R. 2008. The timing of taxes on CO_2 emissions when technological change is endogenous. *J. Environ. Econ. Manag.* 55:194–212

Hart R. 2019. To everything there is a season: carbon pricing, research subsidies, and the transition to fossil-free energy. *J. Assoc. Environ. Resour. Econ.* 6:349–89

Hassler J, Krusell P, Olovsson C. 2019. *Directed technical change as a response to natural-resource scarcity*. Work. Pap. Ser. 375, Cent. Bank Swed., Stockholm

Hémous D. 2016. The dynamic impact of unilateral environmental policies. *J. Int. Econ.* 103:80–95

Hémous D, Olsen M. 2021. The rise of the machines: automation, horizontal innovation and income inequality. *Am. Econ. J. Macroecon.* In press

Heutel G, Moreno-Cruz J, Shayegh S. 2018. Solar geoengineering, uncertainty, and the price of carbon. *J. Environ. Econ. Manag.* 87:24–41

Hicks J. 1932. *The Theory of Wages*. London: Macmillan

Hornbeck R, Naidu S. 2014. When the levee breaks: black migration and economic development in the American South. *Am. Econ. Rev.* 104:963–90

Howell S. 2017. Financing innovation: evidence from R&D grants. *Am. Econ. Rev.* 107:1136–64

Irmen A. 2017. Capital- and labor-saving technical change in an aging economy. *Int. Econ. Rev.* 58:261–85

Irmen A, Tabakovic A. 2017. Endogenous capital-and labor-augmenting technical change in the neoclassical growth model. *J. Econ. Theory* 170:346–84

Jonstone N, Hascic I, Popp D. 2010. Renewable energy policies and technological innovation: evidence based on patent counts. *Environ. Resour. Econ.* 45:133–55
Katz L, Murphy K. 1992. Changes in relative wages, 1963–1987: supply and demand factors. *Q. J. Econ.* 107:35–78
Kennedy C. 1964. Induced bias in innovation and the theory of distribution. *Econ. J.* 74:541–47
Knittel C. 2011. Automobiles on steroids: product attribute trade-offs and technological progress in the automobile sector. *Am. Econ. Rev.* 101:3368–99
Kruse-Andersen P. 2020. *Directed technical change, environmental sustainability, and population growth*. Discuss. Pap. 19-12, Univ. Copenhagen, Copenhagen, Den.
Lazkano I, Nøstbakken L, Pelli M. 2017. From fossil fuels to renewables: the role of electricity storage. *Eur. Econ. Rev.* 99:113–29
Lewis E. 2011. Immigration, skill mix, and capital skill complementarity. *Q. J. Econ.* 126:1029–69
Loebbing J. 2021. An elementary theory of directed technical change and wage inequality. *Rev. Econ. Stud.* **https://doi.org/10.1093/restud/rdab025**
Lööf H, Baum C, Perez L. 2018. *Directed technical change in clean energy: evidence from the solar industry*. Work. Pap. Ser. Econ. Inst. Innov. 470, R. Inst. Technol., CESIS, Stockholm, Swed.
Lordan G, Neumark D. 2018. People versus machines: the impact of minimum wages on automatable jobs. *Labour Econ.* 52:40–53
Martinez J. 2019. *Automation, growth and factor shares.* Work. Pap., London Bus. Sch., London
Massetti E, Carraro C, Nicita L. 2009. How does climate policy affect technical change? An analysis of the direction and pace of technical progress in a climate-economy model. *Energy J.* 30:7–38
Nakamura H, Zeira J. 2018. *Automation and unemployment: Help is on the way*. Work. Pap., Osaka City Univ., Osaka, Jpn.
Newell RG, Jaffe AB, Stavins RN. 1999. The induced innovation hypothesis and energy-saving technological change. *Q. J. Econ.* 114:941–75
Noailly J, Smeets R. 2015. Directing technical change from fossil-fuel to renewable energy innovation: an application using firm-level patent data. *J. Environ. Econ. Manag.* 72:15–37
Nordhaus WD. 2002. Modeling induced innovation in climate-change policy. In *Technological Change and the Environment*, ed. A Grübler, N Nakicenovic, WD Nordhaus, pp. 182–209. Washington, DC: Resour. Future
Papageorgiou C, Saam M, Schulte P. 2017. Substitution between clean and dirty energy inputs: a macroeconomic perspective. *Rev. Econ. Stat.* 99:281–90
Peretto P, Seater J. 2013. Factor-eliminating technical change. *J. Monet. Econ.* 60:459–73
Popp D. 2002. Induced innovation and energy prices. *Am. Econ. Rev.* 92:160–80
Popp D. 2004. ENTICE: endogenous technological change in the DICE model of global warming. *J. Environ. Econ. Manag.* 24:742–68
Popp D. 2006. ENTICE-BR: the effects of backstop technology R&D on climate policy models. *Energy Econ.* 28:188–222
Popp D. 2019. Environmental policy and innovation: a decade of research. *Int. Rev. Environ. Resour. Econ.* 13:265–337
Popp D, Newell R, Jaffe A. 2010. Energy, the environment and technological change. In *Handbook of the Economics of Innovation*, ed. B Hall, N Rosenberg, pp. 873–937. Amsterdam: Elsevier
Prettner K, Strulik P. 2020. Innovation, automation and inequality: policy challenges in the race against the machine. *J. Monet. Econ.* 116:249–65
Ray D, Mookherjee D. 2020. *Growth, automation and the long run share of labor*. NBER Work. Pap. 26658
Ricci F. 2007. Environmental policy and growth when inputs are differentiated in pollution intensity. *Environ. Resour. Econ.* 38:285–310
Romer P. 1990. Endogenous technological change. *J. Political Econ.* 98:S71–S102
San S. 2019. *Labor supply and directed technical change: evidence from the abrogation of the Bracero Program in 1964*. Work. Pap., New York Univ., New York
Shanker A, Stern D. 2018. *Energy intensity, growth and technical change*. CAMA Work. Pap. 46/2018, Aust. Natl. Univ., Canberra, Aust.

Smulders S, de Nooij M. 2003. The impact of energy conservation on technology and economic growth. *Resour. Energy Econ.* 25:59–79

Stern D, Pezzey J, Lu Y. 2020. *Directed technical change and the British Industrial Revolution.* Dep. Work. Pap., Aust. Natl. Univ., Canberra, Aust.

Sue Wing I. 2003. *Induced technical change and the cost of climate policy*. Rep. 102, MIT Jt. Progr. Sci. Policy Glob. Change, Mass. Inst. Technol., Cambridge

Uzawa H. 1961. Neutral inventions and the stability of growth equilibrium. *Rev. Econ. Stud.* 28:117–24

van den Bijgaart I. 2017. The unilateral implementation of a sustainable growth path with directed technical change. *Eur. Econ. Rev.* 91:305–27

van der Meijden G, Smulders S. 2017. Carbon lock-in: the role of expectations. *Int. Econ. Rev.* 58:1371–415

Verdolini E, Galeotti M. 2011. At home and abroad: an empirical analysis of innovation and diffusion in energy technologies. *J. Environ. Econ. Manag.* 61:119–34

Witajewski-Baltvilks J, Fischer C. 2019. *Green innovation and economic growth in a North-South model.* Work. Pap. 19-04, Resour. Future, Washington, DC

Zeira J. 1998. Workers, machines, and economic growth. *Q. J. Econ.* 113:1091–117