The Simple Macroeconomics of AI[*]

Daron Acemoglu
Massachusetts Institute of Technology

May 12, 2024

Abstract

This paper evaluates claims about large macroeconomic implications of new advances in AI. It starts from a task-based model of AI's effects, working through automation and task complementarities. So long as AI's microeconomic effects are driven by cost savings/productivity improvements at the task level, its macroeconomic consequences will be given by a version of Hulten's theorem: GDP and aggregate productivity gains can be estimated by what fraction of tasks are impacted and average task-level cost savings. Using existing estimates on exposure to AI and productivity improvements at the task level, these macroeconomic effects appear nontrivial but modest—no more than a 0.66% increase in total factor productivity (TFP) over 10 years. The paper then argues that even these estimates could be exaggerated, because early evidence is from easy-to-learn tasks, whereas some of the future effects will come from hard-to-learn tasks, where there are many context-dependent factors affecting decision-making and no objective outcome measures from which to learn successful performance. Consequently, predicted TFP gains over the next 10 years are even more modest and are predicted to be less than 0.53%. I also explore AI's wage and inequality effects. I show theoretically that even when AI improves the productivity of low-skill workers in certain tasks (without creating new tasks for them), this may increase rather than reduce inequality. Empirically, I find that AI advances are unlikely to increase inequality as much as previous automation technologies because their impact is more equally distributed across demographic groups, but there is also no evidence that AI will reduce labor income inequality. Instead, AI is predicted to widen the gap between capital and labor income. Finally, some of the new tasks created by AI may have negative social value (such as design of algorithms for online manipulation), and I discuss how to incorporate the macroeconomic effects of new tasks that may have negative social value.

JEL Classification: E24, J24, O30, O33.

Keywords: Artificial Intelligence, automation, ChatGPT, inequality, productivity, technology adoption, wage.

---

1 Introduction

Artificial intelligence (AI) has captured imaginations. Promises of rapid, even unparalleled, productivity growth as well as new pathways for complementing humans have become commonplace. There is no doubt that recent developments in generative AI and large language models that produce text, information and images—and Shakespearean sonnets—in response to simple user prompts are impressive and even spellbinding. ChatGPT, originally released on November 30, 2022, soon became the fastest spreading tech platform in history, with an estimated 100 million monthly users only two months after launch.

AI will have implications for the macroeconomy, productivity, wages and inequality, but all of them are very hard to predict. This has not stopped a series of forecasts over the last year, often centering on the productivity gains that AI will trigger. Some experts believe that truly transformative implications, including artificial general intelligence (AGI) enabling AI to perform essentially all human tasks, could be around the corner.[1] Other forecasters are more grounded, but still predict big effects on output. Goldman Sachs (2023) predicts a 7% increase in global GDP, equivalent to $7 trillion, and a 1.5% per annum increase in US productivity growth over a 10-year period. Recent McKinsey Global Institute (2023) forecasts suggest that generative AI could offer a boost as large as $17.1 to $25.6 trillion to the global economy, on top of the earlier estimates of economic growth from increased work automation. They reckon that the overall impact of AI and other automation technologies could produce up to a $1.5 - 3.4$ percentage point rise in average annual GDP growth in advanced economies over the coming decade.[2]

Are such large effects plausible? And if there are going to be productivity gains, who will

---

[1]Korinek and Suh (2024) predict a "baseline" GDP growth of 100% over the next 10 years, and also entertain the possibility of much higher "aggressive" AGI growth rates, such as a 300% increase in GDP. Many others are seeing recent developments as a confirmation of the forecasts in Kurzweil (2005) about the impending arrival of "singularity" and "explosive" economic growth (Davidson, 2021).

[2]Three caveats are in order. First, although most recent advances are in generative artificial intelligence, the economic forces explored here apply to other types of AI, and estimates of exposed tasks I use come on the basis of anticipated improvements in a range of AI-related technologies, including computer vision and software building on large language models. Hence, I consider the numbers here to apply to all of artificial intelligence and thus typically refer to "AI", unless there is a reason to emphasize generative AI.

Second, I focus on the US economy because much of the existing evidence on microeconomic effects of AI and prevalence of exposed tasks is from the United States. The impact on other industrialized nations should be similar, whereas the consequences for the developing world are harder to ascertain and require much more in-depth research.

Third, some commentators use "productivity" to refer to output per worker (or average labor productivity), while others mean total factor productivity (TFP). Throughout, I distinguish between aggregate TFP and GDP (per capita/worker) effects, and I use productivity improvement at the micro/task level as synonymous to cost savings.

be their beneficiary? With previous automation technologies, such as robotics, most gains accrued to firm owners and managers, while workers in impacted occupations experienced negative outcomes (e.g., Acemoglu and Restrepo, 2020a). Could it be different this time? Some experts and commentators are more optimistic. A few "proof-of-concept" experimental studies document nontrivial productivity gains from generative AI, largely driven by improvements for less productive or lower-performing workers (e.g., Peng et al., 2023; Noy and Zhang, 2023; Brynjolfsson et al., 2023), and this has prompted some experts to be cautiously optimistic (Autor, 2024), while others are forecasting a "blue-collar bonanza" (The Economist, 2023).

This paper uses the framework from Acemoglu and Restrepo (2018, 2019b, 2022) to provide some insights for these debates, especially relevant for the medium-term (about 10-year) macroeconomic effects of AI. I build a task-based model, where the production of a unique final good requires a series of tasks to be performed, and these tasks can be allocated to either capital or labor, which have different comparative advantages. Automation corresponds to the expansion of the set of tasks that are produced by capital (including digital tools and algorithms). In this framework, AI-based productivity gains—measured either as growth of average output per worker or as total factor productivity (TFP) growth—can come from a number of distinct channels (see Acemoglu and Restrepo, 2019a):

- *Automation* (or more precisely extensive-margin automation) involves AI models taking over and reducing costs in certain tasks. In the case of generative AI, various mid-level clerical functions, text summary, data classification, advanced pattern recognition, and computer vision tasks are among those that can be profitably automated.

- *Task complementarity* can increase the productivity in tasks that are not fully automated and may even raise the marginal product of labor. For example, workers performing certain tasks may have better information or access to other complementary inputs. Alternately, AI may automate some subtasks, while at the same time enabling workers to specialize and raise their productivity in other aspects of their job.

- *Deepening of automation* can take place, increasing the productivity of capital in tasks that have already been automated. For example, an already-automated IT security task may be performed more successfully by generative AI.

- *New tasks* may be created thanks to AI and these tasks may impact the productivity

of the whole production process.[3]

In this paper, I focus on the first two channels, though I also discuss how new tasks enabled by AI can have positive or negative effects. I do not dwell on deepening of automation, because the tasks impacted by (generative) AI are different than those automated by the previous wave of digital technologies, such as robotics, advanced manufacturing equipment and software systems.[4] I also do not discuss how AI can have revolutionary effects by changing the process of science (a possibility illustrated by neural network-enabled advances in protein folding and new crystal structures discovered by the Google subsidiary DeepMind), because large-scale advances of this sort do not seem likely within the 10-year time frame and many current discussions focus on automation and task complementarities.

I show that when AI's microeconomic effects are driven by cost savings (equivalently, productivity improvements) at the task level—due to either automation or task complementarities—its macroeconomic consequences will be given by a version of Hulten's theorem: GDP and aggregate productivity gains can be estimated by what fraction of tasks are impacted and average task-level cost savings. This equation disciplines any GDP and productivity effects from AI. Despite its simplicity, applying this equation is far from trivial, because there is huge uncertainty about which tasks will be automated or complemented, and what the cost savings will be.

Nevertheless, as an illustrative exercise, I use data from a number of recent studies, in particular, Eloundou et al. (2023) and Svanberg et al. (2024), as well as the experimental studies mentioned above, to obtain some back-of-the-envelope numbers. Eloundou et al. (2023) provide the first systematic estimates of what tasks will be impacted by generative AI and computer vision technologies. Their methodology does not fully distinguish whether the impact will take the form of automation or task complementarities, and does not provide information on when we should expect these impacts to be realized and how large their cost savings will be.[5] For computer vision technologies, Svanberg et al. (2024) provide estimates

---

[3]New tasks in this framework also capture the possibility of productivity-enhancing reorganizing production. The role of AI in enabling such reorganization is emphasized by, among others, Bresnahan (2019) and Agrawal et al. (2023).

[4]Eloundou et al. (2023) report negative statistical associations between their measure of exposure to AI, which I use below, and measures of exposure to robots and manual routine tasks.

[5]More specifically, I use the most granular information that Eloundou et al. (2023) present, which is their "automation index", coded with help from GPT-4. This index provides information on how much of the activities involved in a task/occupation can be performed by AI. Although this index has somewhat greater emphasis on automation, it does not systematically distinguish between automation and task complementarities. As I discuss further below and Eloundou et al. (2023) themselves note, their exposure measure often captures the possibility that generative AI and related digital technologies can perform some of the

of what fraction of tasks that are potentially exposed to AI can be feasibly automated in different time frames.

I take Eloundou et al.'s estimates of tasks that are exposed to AI (without distinguishing automation vs. task complementarities). I then aggregate this to the occupational level and weight the importance of each occupation by its wage bill share in the US economy. This calculation implies that 20% of US labor tasks are exposed to AI. I then use Svanberg et al.'s estimate for computer vision tasks that, among all exposed tasks, 23% can be profitably performed by AI (for the rest, the authors estimate that the costs would exceed the benefits). I take the average labor cost savings to be 27%—the average of the estimates in Noy and Zhang (2023) and Brynjolfsson et al. (2023)—and turn this into overall cost savings using industry labor shares, which imply average overall cost savings of 14.4%.

This calculation implies that total factor productivity (TFP) effects within the next 10 years should be no more than 0.66% in total—or approximately a 0.064% increase in TFP growth annually. If we add bigger productivity gains from Peng et al. (2023), which are less likely to be broadly applicable, or incorporate further declines in GPU costs, this number still remains around 0.9%.

To turn these numbers into GDP estimates, we need to know how much the capital stock will increase due to AI. I start with the benchmark of a rise in the capital stock proportional to the increase in TFP. This benchmark is consistent with the fact that generative AI does not seem to require huge investments by users (beyond those made by designers and trainers of the models). With these investment effects incorporated, GDP is also estimated to grow by $0.93\% - 1.16\%$ over the next 10 years. When I assume that the investment response will be similar to those for earlier automation technologies and use the full framework from Acemoglu and Restrepo (2022) to estimate the increase in the capital stock, the upper bound on GDP effects rises to around $1.4\% - 1.56\%$. Nevertheless, my framework also clarifies that if the capital-output ratio increases in response to the TFP rise, this may increase GDP by more than TFP, but does not additionally contribute to welfare, because the extra investment comes out of consumption.[6]

I then argue that the numbers above may be overestimates of the aggregate productivity benefits from AI, because existing estimates of productivity gains and cost savings are in tasks that are "easy-to-learn", which then makes them easy for AI. In contrast, some of the

---

subtasks in an occupation, enabling workers to focus on and specialize in other activities, and thus contains both automation and task complementarity elements.

[6]In addition, if AI models continue to increase their energy requirements, this would contribute to measured GDP, but would not be a beneficial change for welfare.

future effects will come from "hard-to-learn" tasks, where there are many context-dependent factors affecting decision-making, and most learning is based on the behavior of human agents performing similar tasks (rather than objective outcome measures). Productivity gains from AI in these hard tasks will be less—though, of course, it is challenging to determine exactly how much less. Using a range of (speculative) assumptions, I estimate an upper bound of 73% easy tasks among Eloundou et al.'s exposed tasks. I suppose that productivity gains in hard tasks will be approximately one quarter of the easy ones. This leads to an updated, more modest increase in TFP and GDP in the next 10 years that can be upper bounded by 0.53% and 0.90%, respectively.

New tasks created with AI can more significantly boost productivity. However, some of the new AI-generated tasks are manipulative and may have negative social value, such as deepfakes, misleading digital advertisements, addictive social media or AI-powered malicious computer attacks. While it is difficult to put numbers on good and bad new tasks, based on recent research I suggest that the negative effects from new bad tasks could be sizable. I make a very speculative attempt using numbers on the negative welfare effects of social media from a recent paper by Bursztyn et al. (2023). These authors find that consumers have positive willingness to pay for using social media (in particular Instagram and TikTok) when others are using it, but they would prefer that neither themselves nor others use it. Roughly speaking, their estimates imply that revenue can increase by about $53 per user-month, but this has a negative impact on total GDP/welfare equivalent to $19 per user-month. Combining these numbers with an estimate of the fraction of activities that may generate negative social value (in practice, revenues from social media and spending on attack-defense arms races in IT security), I suggest that with more intensive use of AI, it is possible to have nontrivial increases in GDP. For example, AI may appear to increase GDP by 2%, while in reality reducing welfare by -0.72% (in consumption equivalent units).

Finally, I explore AI's wage and inequality effects. My framework implies that productivity gains from AI are unlikely to lead to sizable wage rises. Moreover, even if AI improves the productivity of low- and middle-performing workers (or workers with limited expertise in complex tasks), I argue that, theoretically, this may not translate into lower inequality. In fact, I show by means of a simple example how an increase in the productivity of low-skill workers in certain tasks can lead to *higher* rather than lower inequality. Adapting the general equilibrium estimates from Acemoglu and Restrepo (2022) to the setting of AI, I find that the more intensive use of AI is unlikely to lead to substantial wage declines for

affected groups, because AI-exposed tasks are more evenly distributed across demographic groups than were the tasks exposed to earlier waves of automation. Nevertheless, I estimate that AI will not reduce inequality and is likely to have a negative effect on the real earnings of low-education women (especially white, native-born low-education women). My findings also suggest that AI will further expand the gap between capital and labor income as a whole.

In the conclusion, I argue that as originally suggested in Acemoglu and Restrepo (2018), more favorable wage and inequality effects, as well as more sizable productivity benefits, will likely depend on the creation of new tasks for workers in general and for middle- and low-pay workers in particular. While this is feasible in theory and I have argued elsewhere how it could be achieved (Acemoglu, 2021 and Acemoglu et al., 2023), I also discuss why this does not seem to be the focus of artificial intelligence research at the moment.

In sum, it should be clear that forecasting AI's effects on the macroeconomy is extremely difficult and will have to be based on a number of speculative assumptions. Nevertheless, the gist of this paper is that a simple framework can discipline our thinking and forecasts, and if we take this framework and existing estimates seriously, it is difficult to arrive at very large macroeconomic gains.

The rest of the paper is organized as follows. The next section outlines the conceptual framework I use throughout the paper and derives a number of theoretical insights on aggregate productivity gains, investment responses, and wage and inequality effects. It also discusses the crucial distinction between easy-to-learn and hard-to-learn tasks and their productivity implications, and introduces the contrast between good and bad new tasks. Section 3 provides a preliminary quantitative analysis of new AI breakthroughs within this framework. It first presents a baseline (upper bound) estimate on the basis of the fraction of existing tasks that are likely to be impacted by AI within the next 10 years and existing estimates of cost savings (productivity improvements) from AI. It then refines this estimate by introducing the distinction between easy-to-learn and hard-to-learn tasks and undertakes a preliminary classification of AI-exposed tasks into the easy and hard categories. I also make an even more speculative attempt at incorporating the macroeconomic implications of bad new tasks into this framework. Finally, I report estimates on the wage and inequality implications of recent AI advances. Section 4 concludes with a general discussion, while the Appendix includes additional information on how tasks are classified into exposed and non-exposed and easy-to-learn and hard-to-learn categories.

2 Conceptual Framework

The model here builds on Acemoglu and Autor (2011) and Acemoglu and Restrepo (2018, 2019b, 2022), and I focus on the main elements of the framework, referring the reader to these papers for further details and refinements. The economy is static and involves the production of a unique final good, and all markets are competitive.[7]

The production of a unique final good takes place by combining a set of tasks, with measure N, using the following production function

$$Y = B(N) \left(\int_0^N y(z)^{\frac{\sigma-1}{\sigma}} dz \right)^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

where $Y(z)$ denotes the output of task z for $z \in [0, N]$, $\sigma \geq 0$ is the elasticity of substitution between tasks and the parameter $B(N)$ depends on N to capture the possible system-wide effects of new tasks, though in what follows I will suppress this dependence to simplify the notation. For now, the elasticity σ can take any value, but it is reasonable to presume $\sigma \leq 1$, so that tasks are gross complements. I later set the elasticity of substitution between tasks to $\sigma \simeq 0.5$, as estimated by Humlum (2021) and also imposed in Acemoglu and Restrepo (2022).

Tasks can be produced using capital or labor according to the production function

$$y(z) = A_L \gamma_L(z) l(z) + A_K \gamma_K(z) k(z) \text{ for any } z \in [0, N],$$

where A_L and A_K are labor-augmenting and capital-augmenting productivity terms, $\gamma_L(z)$ and $\gamma_K(z)$ are labor's and capital's task-specific productivity schedules, and $l(z)$ and $k(z)$ denote labor and capital allocated to performing task z. This task production function implies that capital and labor have different productivities in different tasks, but within a task they are perfect substitutes.[8]

---

[7]Acemoglu and Restrepo (2018) provide a dynamic version of this economy with capital accumulation and endogenous technological choices, while Acemoglu and Restrepo (2022) provide a generalization with multiple types of labor and multiple sectors, and Acemoglu and Restrepo (2023) consider a non-competitive version of this economy. Extending the framework in any of these directions does not materially affect the results I discuss here.

[8]One important simplification is to assume that tasks assigned to labor do not require any capital or tools, which is clearly unrealistic. The online Appendix of Acemoglu and Restrepo (2018) shows that the results are very similar if the task production function is modified such that:

$$y(z) = A_L \gamma_L(z) \left[l(z)^{1-\kappa} k_C(z)^\kappa \right] + A_K \gamma_K(z) k(z),$$

where $\kappa \in (0, 1)$ and $k_C(z)$ is labor-complementary capital in task z (while $k(z)$ denotes capital used for automating task z). Because $\kappa < 1$, tasks assigned to labor are still less intensive in capital than are fully-automated tasks.

Throughout, I assume that $\gamma_L(z)/\gamma_K(z)$ is increasing in z, so that labor has a *comparative advantage* in higher-indexed tasks. This implies that there exists a threshold I such that tasks $z \leq I$ are produced with capital and those above this threshold are produced with labor.

I normalize the total population to 1 and assume that different workers have different units of effective labor. To simplify the discussion, I assume that there are two types of labor, high-skill and low-skill, and there is no comparative advantage difference between these two types of labor (I generalize this later). The only difference is that high-skill workers, which make up a fraction ϕ^H of the population, have λ^H units of effective labor, while the remaining $\phi^U = 1 - \phi^H$ low-skill workers have only $\lambda^U < \lambda^H$ units of effective labor. This specification ensures that both high-skill and low-skill workers could be performing some of the same tasks. It also implies that wage inequality is pinned down by λ^H/λ^U—a feature I relax later.

I also assume that all labor is supplied inelastically, so I write the total supply of labor as

$$\phi^U \lambda^U + \phi^H \lambda^H = L.$$

The labor market-clearing condition is

$$L = \int_0^N l(z)dz, \tag{2}$$

and I denote the wage rate by w.

Capital is specialized for the tasks in which it is used, and I assume that capital of type z is produced linearly from the final good with unit cost

$$R(z) = R(K)\rho(z), \tag{3}$$

where

$$K = \int_0^N k(z)dz$$

is the total capital stock of the economy. All firms take the cost of capital for task z, $R(z)$, as given. The first term in (3) implies that the required rate of return on capital can increase when the capital stock of the economy is larger and the second term is task-specific, representing the possibility that different types of capital could have different costs. For tasks that are not yet technologically automated—meaning that they cannot be produced by capital—we can either set $\gamma_K(z) = 0$ or take $\rho(z)$ to be very large.

Finally, I assume that there exists a (non-satiated) representative household that consumes the final good (net of capital expenditures) and I denote the consumption of this household by C.

2.1 Equilibrium

I focus on a competitive equilibrium, which satisfies the following usual conditions:

- The allocation of tasks $z \in [0, N]$ is cost-minimizing. That is, task $z \in [0, N]$ is produced by labor if and only if

$$\frac{w}{A_L \gamma_L(z)} < \frac{R(z)}{A_K \gamma_K(z)}.$$

- The amount of capital $k(z)$ is chosen to maximize $Y - R(z)k(z)$, where Y is given as in (1).

- The labor market clears. That is, (2) holds.

Notice that the first condition imposes an innocuous tie-breaking rule that when indifferent, firms use capital for performing a task. Given this tie-breaking rule, all tasks $z > I$ will be performed by labor (i.e., $l(z) = 0$ for all $z \leq I$ and $k(z) = 0$ for all $z > I$). Whether this is high- or low-skill labor is indeterminate in the baseline model, so I focus on the overall amount of effective labor units.

In a competitive equilibrium, all tasks performed by labor must have

$$B^{\frac{\sigma-1}{\sigma}} A_L^{\frac{\sigma-1}{\sigma}} \gamma_L(z)^{\frac{\sigma-1}{\sigma}} l(z)^{-\frac{1}{\sigma}} Y^{\frac{1}{\sigma}} = w. \tag{4}$$

This implies that for any two tasks $z > I$ and $z' > I$,

$$\frac{l(z)}{l(z')} = \frac{\gamma_L(z)^{\sigma-1}}{\gamma_L(z')^{\sigma-1}}. \tag{5}$$

Notice that when $\sigma < 1$, less labor is allocated to tasks in which labor's productivity is higher—a feature whose implications I will emphasize later. Equation (5), combined with the labor market-clearing condition (2), implies

$$l(z) = \frac{\gamma_L(z)^{\sigma-1}}{\int_I^N \gamma_L(z)^{\sigma-1} dz} L. \tag{6}$$

Moreover, with a similar reasoning for any task $z < I$, only capital is used, and the first-order condition for capital intensity is simply

$$B^{\frac{\sigma-1}{\sigma}} A_K^{\frac{\sigma-1}{\sigma}} \gamma_K(z)^{\frac{\sigma-1}{\sigma}} k(z)^{-\frac{1}{\sigma}} Y^{\frac{1}{\sigma}} = R(K)\rho(z). \tag{7}$$

Combining (6) and (7) with (1), GDP or total output can be written as

$$Y = \left(\frac{\left(\int_I^N \gamma_L(z)^{\sigma-1} dz \right)^{\frac{1}{\sigma}} (BA_L L)^{\frac{\sigma-1}{\sigma}}}{1 - \left(\int_0^I \left(\frac{\gamma_K(z)}{R(K)\rho(z)} \right)^{\sigma-1} dz \right) A_K^{\sigma-1} B^{\frac{\sigma^2-1}{\sigma}}} \right)^{\frac{\sigma}{\sigma-1}}. \tag{8}$$

The denominator here is due to the roundabout nature of production, and I assume that

$$\left(\int_0^I \left(\frac{\gamma_K(z)}{R(K)\rho(z)} \right)^{\sigma-1} dz \right) A_K^{\sigma-1} B^{\frac{\sigma^2-1}{\sigma}} < 1 \tag{9}$$

to ensure that output is finite in this economy. (Otherwise, because output linearly produces machines, which then produce output, overall output can reach infinity). With an identical argument to that in Acemoglu and Restrepo (2022), an equilibrium exists and is unique, provided that (9) is satisfied.

2.2 How AI Could Affect Production

Before completing the characterization of equilibrium, I discuss how AI could affect production in this economy.

1. AI enables further (extensive-margin) *automation*, increasing I. Such automation could be triggered either because AI reduces the cost of capital for some marginal tasks (i.e., tasks slightly above I) or increases the effectiveness of machinery or algorithms performing some marginal tasks, thus raising $\gamma_K(z)$ for some z above I. Obvious examples of this type of automation include generative AI tools such as large language models taking over simple writing, translation and classification tasks as well as somewhat more complex tasks related to customer service and information provision, or computer vision technologies taking over image recognition and classification tasks.

2. AI can generate new *task complementarities*, raising the productivity of labor in tasks it is performing.[9] For example, AI could provide better information to workers, directly increasing their productivity. This possibility could be modeled as AI reducing the cost of complementary capital $k_C(z)$ in some tasks $z > I$ in the more general formulation in footnote 8. Alternatively, AI could automate some subtasks (such as providing ready-made subroutines to computer programmers) and simultaneously enable humans to

---

[9]I refer to this channel as "tasks complementarities," rather than "labor augmentation," because it can alter the distribution of subtasks workers perform and complement worker productivity in some of these subtasks, and also because "augmentation" is sometimes used to refer to the introduction of complementary activities for labor, such as new tasks, which have very distinct effects, as I explain.

specialize in other subtasks, where their performance improves. This channel would require the explicit modeling of the range of subtasks making up each task. In this case, new AI technologies would perform some of these subtasks and do so with sufficiently high productivity, so that the subtask-level displacement would be weaker than the productivity gains, expanding the demand for labor and the marginal productivity of labor in these tasks. The possibility that the productivity effect could be larger than the displacement effect and expand labor demand is the same as in the basic models of automation, as exposited in Acemoglu and Restrepo (2018, 2019b). Even more interestingly, AI may enable workers to specialize in the non-automated subtasks and raise their expertise in these activities (e.g., when humans spend less time in writing standard subroutines, they can become better at other parts of programming). I represent task complementarities by an increase in $\gamma_L(z)$ in some tasks $z \leq I$, or when they happen in all tasks, by an increase in A_L.

3. AI could induce *deepening of automation*—meaning improving performance, $\gamma_K(z)$, or reducing costs, $\rho(z)$, in some previously capital-intensive tasks (tasks $z \leq I$). Examples include IT security, automated control of inventories, and better automated quality control (see Acemoglu and Restrepo, 2019a).

4. AI can generate *new labor-intensive products or tasks*, which corresponds to an increase in N. As argued in Acemoglu and Restrepo (2020b), Acemoglu (2021) and Acemoglu et al. (2023), there are many pathways for such new tasks. Later I also discuss the case where some of these new products and tasks can be manipulative and have negative social value.

The effects of new AI tools will depend on the extent of each one of these effects, and I will try to provide more specificity on these possibilities later. In the rest of this section, I derive the consequences of different effects of AI.

2.3   Equilibrium Wages and Comparative Statics

As a first step, let us combine (4) and (6), so that the equilibrium wage can be expressed as

$$w = \left(\frac{Y}{L}\right)^{\frac{1}{\sigma}} (BA_L)^{\frac{\sigma-1}{\sigma}} \left(\int_I^N \gamma_L(z)^{\sigma-1} dz\right)^{\frac{1}{\sigma}}. \tag{10}$$

This equation is intuitive. The first term shows that the wage is proportional to labor productivity (raised to the power $1/\sigma$), and the second term captures the contribution to

the marginal productivity of labor coming from Hicks-neutral and labor-augmenting technologies, while the third term represents the contribution of the allocation of tasks to the marginal productivity of labor. The effect of any small technological change (potentially altering multiple dimensions of the production technology, such as B; A_L and A_K; $\gamma_L(z)$ and $\gamma_K(z)$; and I and N) can then be written as:

$$d \ln w = \frac{1}{\sigma} d \ln \left(\frac{Y}{L} \right) + \frac{\sigma - 1}{\sigma} \left(d \ln B + d \ln A_L \right) + \frac{1}{\sigma} d \ln \left(\int_I^N \gamma_L(z)^{\sigma-1} dz \right). \quad (11)$$

The effect of further extensive-margin automation—an increase in I—is given by

$$\frac{d \ln w}{dI} = \frac{1}{\sigma} \frac{d \ln Y}{dI} - \frac{1}{\sigma} \frac{\gamma_L(I)^{\sigma-1}}{\left(\int_I^N \gamma_L(z)^{\sigma-1} dz \right)}.$$

In general, this expression has ambiguous sign, so automation can increase or reduce wages. More specifically, there are two opposing effects (Acemoglu and Restrepo, 2018, 2019b): (a) automation always produces a positive effect on wages (and labor demand) because it increases productivity (or equivalently, reduces costs). This positive productivity effect is represented by the first term; (b) simultaneously, automation displaces workers from the tasks they used to perform. The negative displacement effect is represented by the second term. In the special case where $R(K)$ is constant, it can be verified that automation increases wages. This is not the case, in general, when $R(K)$ is increasing, as shown in Acemoglu and Restrepo (2018), because the displacement effect can be larger than the productivity gains.[10]

Overall, the impact of (extensive-margin) automation on the equilibrium wage is closely tied to its productivity effect, to which I next turn.

Before doing so, I also note that the effects of task complementarities may be a little more complex than typically assumed. Even though an increase in $\gamma_L(z)$ raises the marginal physical product of labor, the equilibrium wage is determined by the *value* of the marginal product of labor, which depends on the adjustment of task prices. As tasks produced by labor become more abundant/easier to perform, these task prices are reduced, and in the empirically relevant case where $\sigma < 1$, task prices decline more than the increase in physical productivity. The equilibrium wage may still increase because of productivity gains, but

---

[10]In particular, for any differentiable constant returns to scale production function $F(K, L)$, Euler's theorem implies: $F(K, L) = wL + RK$. Suppose R is fixed. Then any change in technology that increases $F(K, L)$ at the initial factor supplies must increase $w = \frac{F(K,L) - RK}{L}$. This is not the case when K is fixed, because R can increase more than $F(K, L)$. In Acemoglu and Restrepo (2018, 2019b), the capital stock is taken as given, so a negative impact is possible.

In neoclassical growth models with exponential discounting and time-separable preferences, R is fixed in the long run. However, Acemoglu et al. (2024) show that with more general preferences, R tends to increase following automation and other technological changes, so a negative wage effect is possible.

the benefits to labor may be limited overall. For example, holding I and N constant, an increase in A_L will leave wages constant when $\sigma = s_K$ where s_K denotes the capital share in national income (Acemoglu and Restrepo, 2018). When $\sigma < s_K$, higher A_L can actually reduce real wages. Since $s_K \simeq 0.4$ currently in the US economy, Humlum's estimates of σ mentioned above, of about 0.5, imply that task complementarities or labor-augmenting technological improvements will not raise wages much. Even when they increase wages, these technological shifts reduce the labor share, just like automation does (Acemoglu and Restrepo, 2018, 2019b).

Finally, I note that deepening of automation (which is less likely to be relevant in the case of AI for the reasons discussed in the Introduction) and new tasks always increase wages and the latter always raises the labor share of national income as well (and thus tends to narrow the gap between capital and labor income). I return to a discussion of new tasks later.

2.4 Hulten's Theorem

The nexus of the many effects of AI on the economy is its impact on productivity. To gauge the extent of this impact, we can appeal to Hulten's theorem, which provides a simple formula for competitive economies with constant returns to scale, specifying how micro-level productivity improvements translate into macro changes. Since the economy here is competitive, this theorem applies and disciplines the productivity effects. I now explain this theorem and for ease of exposition, I first take the capital stock of the economy, K, as given. Consider an arbitrary small change in all/any aspects of technologies (B; A_L and A_K; $\gamma_L(z)$ and $\gamma_K(z)$; and I and N), so that the potential effects of AI based both on automation and complementarities in existing tasks are taken into account. Recall that GDP can be equivalently expressed as the price-weighted aggregate of task outputs:

$$Y = \int_0^N p(z)y(z)dz.$$

Since I am considering small changes in technology and the competitive equilibrium is efficient, the impact of all reallocations of factors across tasks and indirect effects via prices

will be second-order and can be ignored in computing GDP changes.[11] Then:

$$dY = \int_0^N p(z)dy(z)dz, \text{ and thus}$$

$$d\ln Y = \frac{dY}{Y} = \int_0^N \frac{p(z)y(z)}{Y}\frac{dy(z)}{y(z)}dz = \int_0^N \chi(z)d\ln y(z)dz,$$

where the last equality defines $\chi(z) = p(z)y(z)/Y$ as the GDP share of task z. Recalling that since capital and labor are fixed, this is also the total factor productivity (TFP) change, I write

$$d\ln \text{TFP} = d\ln Y|_K = \int_0^N \chi(z)\pi_L(z)dz, \tag{12}$$

where $\pi_L(z) = d\ln y(z)$ is the productivity improvement or cost savings in task z driven by AI. I emphasize once more that these cost savings could be rooted in automation or task complementarities. Total effect on GDP is given by

$$d\ln Y = d\ln \text{TFP} + s_K d\ln K, \tag{13}$$

where s_K is the capital share of GDP.[12] This derivation also clarifies that (12) and (13) apply regardless of the task-level production functions, and thus are more general than the wage and inequality results reported above.

If the main impact of AI on productivity is via automation, then $\pi_L(z)$'s are cost savings as capital is substituted for labor. These should not be huge. In the case of robots, expert reports put cost savings at about 30% of labor costs (see Acemoglu and Restrepo, 2020a), though there are reasons to expect that cost savings from AI could be smaller than this, as I discuss later. Estimates from Noy and Zhang (2023) and Brynjolfsson et al. (2023), which I review in greater detail below, suggest numbers in the same ballpark—about 27% of labor costs.[13] Although it is not clear whether these come from extensive-margin automation

---

[11]Namely, change in real GDP can be expressed as the change in the value of the objective function of the social planner (e.g., the utility function of the representative household if one exists). Then the Lagrange multipliers on task-level resource constraints are equivalent to the prices and can be taken as given in evaluating the effects of small changes in parameters/technology on the maximized value of this objective function, by the envelope theorem. See Hulten (1978).

[12]The contribution of capital to output growth can also be derived in the general case, just like Hulten's theorem. In particular, for any constant returns to scale production function $F(K, L|A)$, where A is a technology shifter, we have $d\ln Y/dA = d\ln F(K, L|A)/dA + (d\ln F(K, L)/dK) \cdot dK/dA$. The first term $d\ln F(K, L|A)/dA$ is $d\ln\text{TFP}$. Competitive factor markets imply $d\ln Y = (RK/Y)d\ln K$, and thus $d\ln Y = d\ln\text{TFP} + s_K d\ln K$. Moreover, $d\ln s_K = d\ln K - d\ln Y$, and hence $d\ln Y = (d\ln\text{TFP} + s_K d\ln s_K)/(1 - s_K)$. When the capital-output ratio (or equivalently the capital share) remains constant in response to changes in technology, then $d\ln Y = d\ln\text{TFP}/(1 - s_K)$, which is the first approximation I use. Later, I estimate by how much the capital-output ratio is predicted to increase due to automation and update this estimate.

[13]Peng et al. (2023) estimate even larger effects on how quickly certain programming tasks can be completed. But these refer to a very narrow set of tasks—subroutines that GitHub Copilot can write in common programming languages—and are thus less broadly applicable, and I consider them in robustness checks.

or task complementarities, this distinction is less important for the productivity effects, as noted above. In practice, labor cost savings need to be translated into overall cost savings, which can be done using the share of labor in overall costs for each industry. For example, if the share of labor in industry i is s_L^i, then labor cost savings of π_L^i will translate into an overall cost savings of $\pi^i = s_L^i \pi_L^i$. Taking the average of these across industries, and denoting this economy-wide cost savings by $\bar{\pi}$, I can write

$$d \ln \text{TFP} = \bar{\pi} \times \text{ GDP share of tasks impacted by AI.} \tag{14}$$

I discuss below how this share can be estimated from recent studies on which occupations and tasks will be impacted by AI.

2.5 Easy Tasks and Hard Tasks

As noted above, the proof-of-concept studies Peng et al. (2023), Noy and Zhang (2023), and Brynjolfsson et al. (2023) have focused on tasks where AI could have clear benefits and where there were already sophisticated applications (such as GitHub Copilot for the first study) or where some businesses were already making an effort to use generative AI (as in the customer service application of the third study). I now argue that estimates from tasks that are relatively easy for current AI technologies, even if reliable within their chosen context, cannot be directly extrapolated to the rest of the economy. More generally, it is useful to distinguish between "easy-to-learn tasks" where outside learning (and thus learning by AI models) is easy from those where gaining expertise on the basis of outside observation is hard. Easy-to-learn tasks, which are relatively straightforward for (generative) AI to learn and implement, are defined by two characteristics:

- there is a reliable, observable outcome metric, and

- there is a simple (low-dimensional) mapping between action and the outcome metric.

How to boil an egg (or providing instructions for boiling an egg), the verification of the identity of somebody locked out of a system or the composition of some well-known programming subroutines are easy tasks. The desired outcome—an egg that is boiled to the desired level, allowing only authorized people to access the system, or whether the subroutine works or not—is clear. In none of these cases do the successful outcomes depend on the complex interaction of many dimensions of actions. With reliable, objective measures of success (well boiled egg, no security breach given the ground truth of authorized people,

15

or a subroutine that does not crash), AI models can learn to perform well in a relatively straightforward manner. Beyond this, AI models can also learn from human actions, because there are expert humans who perform well in these tasks, such as expert programmers, and because objective measures of success are available, experts can be identified.[14]

"Hard tasks" typically do not have a simple mapping between action and desired outcome. In hard problems, what leads to the desired outcome in a given problem is typically not known and strongly depends on contextual factors, or the number of relevant contexts may be vast, or new problem-solving may be required. Additionally, there is typically not enough information for the AI system to learn or it is unclear exactly what needs to be learned. Diagnosing the cause of a persistent cough and proposing a course of treatment is a hard problem. There are many complex interactions between past events that may be the cause of the lingering cough and many rare conditions that should be considered. Moreover, there is no large, well-curated data set of successful diagnoses and cures. In hard tasks, AI models can still learn from human decision-makers, but because there is no clear metric of success, identifying and learning from workers with the highest level of expertise will not be straightforward either. As a result, there will be a tendency for the performance of AI models to be similar to the average performance of human decision-makers, limiting the potential for large productivity improvements and cost savings.

AI productivity gains observed so far are from easy tasks. It is reasonable to expect that productivity gains in hard tasks are more limited, at least at first. Productivity gains in easy tasks result from AI models performing these tasks more or less at the same level as expert workers and/or at lower cost than humans. For example, in the Noy and Zhang (2023) study, expert workers are those that are able to summarize and write reasonably well. The desired outcome in this case is straightforward to determine and does not require new problem-solving efforts (this may have been different if the writing tasks were more complex

---

[14]The hard vs. easy distinction is different from the routine vs. non-routine distinction introduced in Autor et al. (2003). Routine tasks are those that involve repeated performance of the same activities (such as knitting or switchboard operations) in stable, predictable environments. The use of digital technologies in these tasks involves step-by-step programming of relevant software and hardware. A significant number of routine tasks have already been automated, whereas AI's promise is its ability to perform non-routine tasks, as also emphasized by Susskind (2020). Some non-routine tasks such as parts of computer programming or writing of simple text are easy-to-learn, while others that require more context-specific decisions and where metrics of successful performance are scarce would be much harder to learn by outside observation.

Crucially, "easy tasks" in this sense need not be easy for humans. In fact, cost savings from easy tasks will be most pronounced when they are expensive for humans to perform. Conversely, some hard tasks, such as those that are based on intuition, experience and judgment, may be relatively straightforward for humans because they do not learn to perform these tasks on the basis of outside observation alone. This is the reason why cost savings relative to human performance are likely to be limited in hard tasks.

or required "creativity"). Generative AI models, such as GPT-4, are trained on vast amounts of this type of writing, making this a clear example of an easy task. Productivity gains in this case come from the fact that lower-expertise or less-skilled workers can be brought up to the level of more expert workers at little cost and the majority of the workers are helped to perform their assigned tasks more quickly. In easy problems, there may even be additional productivity gains from AI models discovering action combinations that were not typically tried or known by expert humans.[15]

There are several barriers to productivity gains in hard tasks. First, the lack of a simple mapping between action and desired outcome will make it much more difficult to train AI models, and early automation efforts and even human-complementary use of AI may be hampered by this slow learning, delaying any productivity gains. Second, human complementarity may be much harder to achieve. When there is no good information on what the desired outcome is (e.g., what was the right diagnosis for the persistent cough?), most of the training data of AI models will come from how humans act in similar circumstances. As a result, learning from humans will not lead to better than human performance and is unlikely to generate new complementarities and reveal different insights than what humans are already doing.[16]

If the premises of this section—that productivity gains will be slower in hard tasks than in easy tasks in the next 10 years and tasks impacted by AI in the near future will be harder than those today—are true, then (14) needs to be adjusted in an obvious way to account for

---

[15]For instance, suppose humans typically start boiling the water with the egg inside, and it reduces variability if the egg is placed inside the saucepan after the water has boiled, then this is something some AI models may discover. AlphaZero's discovery of new effective chess moves can be viewed as an example of the same phenomenon, even if it is in the context of a much more complex and interesting problem.

Looked at it this way, AlphaFold's big success in protein folding can be considered to have both easy and hard aspects. On the hard side, protein folding is a highly multidimensional problem, with no simple mapping between action and desired outcome. On the easy side, however, the desired outcome is observed, so with large amount of computing power, it is possible for an AI model to do even better than humans, as AlphaFold managed to do.

[16]Mathematically, the general learning problem can be formulated as follows. There is a mapping for outcomes specified as $f : X \times Z \to Y$, where X is the space of actions and Z denotes the space of "contexts", and Y is the set of possible outcomes. Observed outcomes are noisy versions of the true outcomes in Y, with some noise process which I do not need to specify here. The mapping f is unknown.

In principle, there are two ways in which AI can be trained in such a problem. First, with sufficient observations on Y and combinations of inputs from $X \times Z$, the AI model may learn the entire function f or some restricted version thereof. Second, if Y is not observable, then the AI model may be trained from human choices, which can be summarized by a correspondence $G : Z \rightrightarrows X$, specifying the possible actions that humans take when confronted with context Z. Either problem can be challenging if $X \times Z$ is very high dimensional. However, the latter type of learning is less likely to lead to something different and better than what humans were doing already. It can sometimes ensure performance at the level of expert humans, if these can be identified, but lack of reliable outcome data may also make it difficult to identify expertise.

which fraction of affected tasks are easy and hard. Specifically, denoting the fraction of easy tasks among those that are impacted by AI as μ^E, and the average cost savings for easy and hard tasks, respectively, as $\bar{\pi}^E$ and $\bar{\pi}^H$, (14) becomes

$$d\ln\text{TFP} = \left(\mu^E\bar{\pi}^E + (1-\mu^E)\bar{\pi}^H\right) \times \text{GDP share of tasks impacted by AI.} \quad (15)$$

I also attempt to provide estimates using this updated TFP equation below, which will lead to gains that are about 25% less than those that do not take the distinction between easy and hard tasks into account.

2.6 Investment Responses

To go from TFP responses to total GDP responses, we need to see how much capital increases. This is also relevant for understanding what magnitude of an investment boom generative AI may trigger.

With this aim, consider a change in technology, for example due to AI. Its effects on capital and investment can be gleaned from (7), which implies

$$k(z) = \frac{B^{\sigma-1}A_K^{\sigma-1}\gamma_K(z)^{\sigma-1}Y}{(R(K)\rho(z))^\sigma},$$

and hence

$$d\ln k(z) = (\sigma-1)d\ln B + (\sigma-1)d\ln A_K + (\sigma-1)\,d\ln\gamma_K(z) + d\ln Y - \sigma d\ln R(K) - \sigma d\ln\rho(z). \quad (16)$$

In what follows, I suppose that B and A_K do not change. If AI significantly improves the practice of science and invention and/or creates new high-productivity tasks, such changes may occur in the future. But I assume that they are constant within the next 10 years.

First, consider the use of AI in some task $z < I$, meaning that AI adds to and improves the performance of existing capital equipment in tasks already performed by capital.[17] Then:

$$d\ln k(z) = (\sigma-1)\,d\ln\left[A_K \cdot \gamma_K(z)\right] + d\ln Y - \sigma d\ln R(K).$$

Notice, however, that when $\sigma < 1$, and in particular for the empirically-plausible value of $\sigma = 0.5$, the first term is negative—an increase in the productivity of AI-augmented capital in capital-intensive tasks reduces investment and the equilibrium capital stock. The last term is also nonpositive, provided that total investment increases. Hence, a natural upper

---

[17]The expression for the case in which labor-intensive tasks also use capital, as in footnote 8, is similar and leads to the same conclusion.

bound for the proportional increase in the capital stock of tasks already performed by capital is the proportional increase in output.

Additionally, if I increases, investment will jump from zero to a positive amount in the newly-automated tasks.[18] Because in the estimates I report below, only a modest fraction of tasks will be automated with AI, this increase may be small as well. Hence, I start with a first estimate that takes the proportional increase in the capital stock to be the same as the increase in aggregate productivity. I then provide another estimate incorporating the full structure of between-industry and between-task substitution, based on the framework and results of Acemoglu and Restrepo (2022). This will enable me to estimate implied changes in the capital share and then use the expressions in footnote 12 to derive updated investment response and GDP estimates.

I finally note that it is straightforward to show that consumer welfare (and consumption) in this economy is proportional to TFP divided by one minus the capital share (i.e., $d \ln C = d \ln \text{TFP}/(1 - s_K)$). In particular, if the capital-output ratio rises and the increase in GDP is more pronounced than that implied by equation (13), then the GDP response would overstate the increase in consumer welfare, as investment comes at the expense of consumption (Acemoglu and Restrepo, 2022).

2.7 New Good Tasks

Adding new tasks will expand productivity, and this effect could be in principle larger than the cost savings due to automation and task complementarities. In addition, new tasks will tend to increase wages. In particular, using the same steps as before,

$$\frac{d \ln w}{dN} = \frac{1}{\sigma} \frac{d \ln Y}{dN} + \frac{1}{\sigma} \frac{\gamma_L(N)^{\sigma-1}}{\left(\int_I^N \gamma_L(z)^{\sigma-1} dz \right)} + \frac{\sigma - 1}{\sigma} \frac{B'(N)}{B(N)}. \tag{17}$$

This is always positive and could be large. Note also that the wage and productivity impact of new tasks can be potentially larger than cost savings in existing tasks, and this is particularly likely to be the case when new tasks improve the entire production process (as captured by the B term), or when they add new sources of cost improvements or complementary functions. Despite new tasks' central role in wage and productivity growth and in reducing labor income inequality (see Acemoglu and Restrepo, 2018, and Autor et al.,

---

[18]Around I, this amount is given by $k(I) = \frac{B^{\sigma-1} A_K^{\sigma-1} \gamma_K(I)^{\sigma-1} Y}{(R(K)\rho(I))^{\sigma}}$. If labor-intensive tasks were previously using capital as well, then the jump would be smaller but still positive.

2024), I will not focus on new good tasks generated by AI for the reasons discussed in detail in the Conclusion.

2.8   New Bad Tasks

AI may generate new tasks that increase revenue, but reduce consumer utility, such as through addiction or manipulation, or in the context of production, via security attacks by malicious actors. To capture this, let us suppose that welfare is given by $\mathcal{W} = C - E$, where the second term is an externality, for example from misinformation or manipulative activities, as I discuss below. In that case, the welfare effects of new tasks generated by AI will be

$$\frac{d\mathcal{W}}{dN} = \frac{dY}{dN} - \frac{dE}{dN}. \tag{18}$$

Although it is difficult to ascertain the magnitude of such negative welfare effects, I argue, based on some recent studies, that they may be nontrivial. Specifically, I use estimates on the relative magnitudes of the two terms in (18) and proxy for the magnitude of the first term by the revenues of tasks where AI can produce new bad tasks or socially harmful activities.

2.9   Wage and Inequality Implications of AI

A number of commentators and experts are cautiously optimistic that advances in generative AI could be beneficial for labor or at the very least not impact workers as adversely as previous waves of digital technologies, such as robotics and software systems, which were predominantly used for automation. There are three potential pathways via which such optimism may be realized.

1. AI can enable productivity increases in tasks currently produced by labor. This is the task-complementarities channel and can be captured either by an increase in A_L or an increase in $\gamma_L(z)$ for a subset of the tasks that are automated or by increases in λ_U and λ_H (which, recall, are the productivities of unskilled and highly-skilled workers). However, when $\sigma < 1$, these types of productivity improvements will reduce the labor share, and thus inequality between capital and labor will increase.

2. If AI generates large productivity gains, it may increase wages even though it reduces the labor share (Acemoglu and Restrepo, 2018, 2019b). This channel thus critically hinges on the magnitude of the productivity effects discussed above, but in any case, always increases inequality between capital and labor.

3. As already discussed, some early studies show that within narrow occupations, lower-performing or lower-expertise workers are the ones benefiting from generative AI. This raises the possibility that AI could be more complementary to lower-skill workers and may reduce labor income inequality. In my framework, this would be captured by an increase in λ_U relative to λ_H. However, even in this case, inequality between capital and labor will rise (provided that $\sigma < 1$).

4. If AI created new (good) tasks, these would reduce inequality between capital and labor, and if enough new tasks were targeting lower-skill workers, this could also reduce labor income inequality (Acemoglu and Restrepo, 2018).

I now argue theoretically that there are several reasons why 1 and 2 listed here are unlikely to be major sources of wage growth or significant limits on inequality. First consider a 1% increase in A_L (or an equivalent increase in $\gamma_L(z)$ for labor-intensive tasks). As explained above, this may not increase wages at all, or may lead to only small wage increases. In fact, from equation (11), abstracting from the productivity effect, the direct impact on the equilibrium wage will be a $(\sigma - 1)/\sigma\%$ change. In the plausible case where $\sigma < 1$, this is negative. The overall impact may still be an increase in wages because of the productivity effect, but as already noted, when σ is approximately equal to the share of capital in national income, the overall impact will be essentially zero. In conclusion, without the creation of a sufficient number of new tasks, inequality between capital and labor will increase and wage rises may be limited.

What about a reduction in inequality because lower-skill workers benefit more? In the model here, the earnings of high-skill workers relative to low-skill workers is always pinned down at λ_H/λ_U. So if new technologies reduce this ratio, they will reduce the gap between high-skill and low-skill workers. But even this conclusion needs to be qualified. Acemoglu and Restrepo (2022) show that in more general settings, with multiple skill groups, there will be ripple effects whereby impacted demographic groups can then compete for tasks previously performed by other groups. In such a situation, an overall increase in labor productivity of both high-skill and low-skill workers in some tasks can lead to their displacement from these tasks, and then the ripple effects can, in principle, affect low-skill workers even more adversely than high-skill workers.

While such adverse effects on low-skill workers are a general possibility in the framework of Acemoglu and Restrepo (2022), I am not aware of worked-out examples where an increase in the productivity of low-skill workers increases inequality. I now provide such an example.

2.10 How Greater Low-Skill Productivity Can Lead to Higher Inequality

Let me illustrate this possibility with a simple example, by relaxing the assumption that there are no comparative advantage differences between high-skill and low-skill workers. Suppose that the economy starts from an equilibrium in which tasks below some I are performed by capital, and tasks between I and N are performed by a combination of high-skill and low-skill workers. In particular, suppose that low-skill workers have a comparative advantage in tasks between I and I^*, and denote their (constant) productivity in these tasks by λ_U, while the productivity of high-skill workers in these tasks is also constant and equal to λ_H.

Suppose also that the relative productivity of high-skill workers in tasks between I^* and N is $\omega > \lambda_H/\lambda_U$—indicating that high-skill workers have a comparative advantage in higher-indexed tasks. However, because there is no strict comparative advantage, the equilibrium may involve both types of workers performing some of the same tasks.[19] Assume also that the elasticity of substitution between tasks is $\sigma < 1$.

Let us start with an equilibrium in which both high-skill and low-skill workers perform tasks $z \in (I, I^*]$, while only high-skill workers perform tasks $z \in (I^*, N]$. In this initial equilibrium, the relative wage of skilled workers will be pinned down by the relative productivities of the two types of workers in the tasks they are both performing—i.e., $z \in (I, I^*]$—and is given by λ_H/λ_U.

Suppose that labor productivity in $z \in (I, I^*]$ increases due to advances in AI, and this also is more helpful for lower-skilled workers, and so λ_H/λ_U declines. I now show that these advances could boost inequality. Because $\sigma < 1$, the prices of tasks $z \in (I, I^*]$ will decline more than the increase in physical productivity of labor, and there will be less labor assigned to these tasks. If this effect is significant, all high-skill workers may be allocated away from these tasks, and the amount of labor demanded in these tasks may fall short of the supply of low-skill workers. In this case, the post-AI allocation may involve only low-skill workers performing tasks $z \in (I, I^*]$, while both low-skill and high-skill workers perform tasks $z \in (I^*, N]$. Then, regardless of how much λ_H/λ_U declines, the relative wage of skilled workers will be determined by the tasks that both types of workers are performing, which are now those above I^*, and thus will be equal to $\omega > \lambda_H/\lambda_U$. Hence, inequality increases

---

[19]This structure makes the example similar to one that has a finite number of tasks—in this instance, three tasks. But I use the setting with a continuum of tasks for continuity with the rest of the paper. The "paradoxical" result I am highlighting here does not depend on this stark structure or on having just two types of labor.

following the rise in the productivity of low-skill workers.

Therefore, I have just proven that, in this general scenario, even a reduced productivity gap between low-skill and high-skill workers in some tasks *could* lead to greater inequality. Hence, the overall inequality implications of AI cannot be directly deduced from its effects on the performance of workers of different skills in a given set of tasks and requires a fuller empirical exploration. In the next section, I investigate AI's inequality effects by adapting the framework and estimates of Acemoglu and Restrepo (2022) to the current setting.

3   A Preliminary Quantitative Evaluation

In this section, I provide a preliminary quantitative evaluation of the possible effects of recent breakthroughs in AI over a horizon of 10 years. The centerpiece will be the use of Hulten's theorem and recent estimates of which tasks can be automated using AI and computer vision technologies and the cost savings thereof. Once I obtain these estimates from Hulten's theorem for TFP growth, I convert them to GDP growth estimates using another series of assumptions on how the capital stock of the economy will respond. I also discuss AI's effects through new bad tasks could be estimated. Finally, I combine these numbers with the more detailed framework in Acemoglu and Restrepo (2022) to obtain even more speculative estimates for wage and inequality implications. Before moving to the estimates, I first describe the sources I use, further discuss existing productivity estimates and motivate various parameter choices.

3.1   Data Sources and Parameter Choices

The centerpiece of the estimates in this paper is equation (14), and its refinement to (15). To implement this equation, two pieces of information are needed:

1. GDP share of tasks that are impacted by AI (inclusive of computer vision) within the next 10 years.

2. Average cost savings in these tasks due to AI, $\bar{\pi}$.

It is impossible to have accurate estimates of either of these two quantities, and hence my repeated caution that the numbers here—and for that matter, other estimates in the literature and the public debate—should be interpreted with great caution as suggestive numbers. Nevertheless, there are studies that have already shed light on these quantities. I now discuss what these are and how I use them.

GDP Share of Tasks Impacted by AI

The most careful estimates of which tasks are exposed to recent AI and computer vision advances come from Eloundou et al. (2023). These authors use two related methodologies for classifying which tasks are exposed to AI and computer vision. Both of those start from O*NET task and Detailed Work Activity (DWA) descriptions. The authors ask GPT-4 to classify all 19,265 tasks and 2,087 DWAs. They also develop a coarser index by manually classifying DWAs and then cross-validating their GPT-based measure with this "human" coding. Here, I focus on the GPT-based measure which allows greater granularity, as I explain next. In addition, Eloundou et al. (2023) distinguish between a direct exposure (α) measure, which is based on their assessment of what large language models (LLMs) can achieve now. They then develop a second, more aggressive measure (their so-called β measure), allowing "indirect" exposure to a hypothetical LLM+. This includes tasks that will be (possibly) exposed to new software and other advances building on current LLM and computer vision technologies.[20]

The index Eloundou et al. report in the paper is based on a binary coding, while they also construct an "automation" index, and in this case their β measure includes granular information about what fraction of the activities might be impacted by LLM+ (ranging across 0%, 25%, 50%, 75% and 100%). In what follows, I take their automation index to be able to use this granular information. This index still contains both automation and task complementarities, even if it emphasizes automation a little more than their other exposure indices (because they code the granular information on the basis of information about what activities can be automated). In particular, as the authors themselves note, the impact of generative AI may often involve automation of some subtasks, allowing human workers to focus on other activities. The granular information contained in this index is especially useful for my purposes, because it provides an assessment of which tasks/occupations are less likely to be impacted by AI. I set all tasks that the authors classify as having 50% or less of their activities impacted by AI and computer vision to zero, and I refer to the rest as "AI exposed tasks".

There are several problems that need to be tackled to turn these estimates into the quantities I need.

- Although Eloundou et al.'s automation index emphasizes automation, it still includes

---

[20]They also report a third, even more aggressive measure, which they refer to as ζ. This measure is more speculative and focuses on what can be ultimately performed by LLM+. In line with the authors' interpretation, this is unlikely to be the case in the near future, and I ignore this third measure.

elements of "augmentation" or task complementarity. This motivates my interpretation that combining their measure with equation (14) will capture cost savings from both automation and task complementarities.

- Eloundou et al.'s data need to be converted into GDP shares. To do this, I combine tasks into occupations, and then I aggregate across occupations using their wage bills, computed from the US Bureau of Labor Statistics (BLS) National Occupational Employment and Wage Estimates pooled across the years 2019-2022. This procedure yields a wage bill-weighted share of exposed occupations equal to 20%. I interpret this number to be the same as the GDP share of tasks exposed to AI.

- Eloundou et al.'s approach is to determine tasks that can be ultimately performed by generative AI and computer vision technologies (such as the technology already incorporated in Dall-E). Two things are missing from the information they provide. The first is how much of the task impact is likely to be realized within the next 10 years. The second is whether, in all of these cases, it is profitable to use AI (e.g., whether automating using AI is cost-effective). Svanberg et al. (2024) make an attempt to provide answers to these questions in the case of computer vision technologies, which are a subset of the technologies Eloundou et al. consider. I take Svanberg et al.'s estimates and extrapolate them to all of the tasks Eloundou et al. consider.[21] Namely, Svanberg et al.'s base estimates imply that among computer vision-exposed tasks, 23% can be feasibly (and profitably) automated within 10 years.[22] Applying this number to Eloundou et al.'s estimates, I arrive at the GDP share of tasks impacted by AI within the next 10 years as $0.23 \times 0.200 = 4.6\%$ of all tasks (or occupations).

Cost Savings from AI

I base my estimate of $\bar{\pi}$ on the experimental studies that have already provided "proof-of-concept" estimates of productivity improvements or labor cost reductions due to AI. Three

---

[21]This is not an innocuous step. Svanberg et al. focus on automation using computer vision, while Eloundou et al.'s exposed technologies include task complementarities as noted above and go beyond computer vision tasks. One could imagine that the share of tasks that can be profitably used with AI within 10 years is different between these two categories. Unfortunately, I do not have another source for such an estimate for non-computer vision AI.

[22]Svanberg et al. also make further extrapolations assuming cost declines in computer vision technologies. They base these on estimates from Besiroglu and Hobbhahn (2022), who project a 22% yearly decrease in computation costs attributed to the expansion of compute power from GPUs. However, a doubling of GPU capacity will not necessarily translate into a 50% decline in costs in general due to diminishing returns and bottlenecks created by other inputs, and because of the limitation of the current architecture. I therefore do not include these in the baseline but return to them in the robustness discussion.

studies, which I have already mentioned, are particularly notable here, and I now describe each one of them.

- Peng et al. (2023) design an experiment where freelance computer programmers are given access to and encouraged to use GitHub Copilot, which (at the time of the experiment) was powered by OpenAI Codex (GPT-3). Participants were given the task of implementing a HTTP server on JavaScript, a popular language for which resources and subroutines are readily available, and GPT-3 was already trained on these resources. They compare the experimental treatment group to a control group that is not given access to GitHub Copilot. They find that the treatment group performed the assigned tasks on average 55.8% faster than the control group. As an aside, like the other studies I mention here, Peng et al. find that these improvements come from otherwise less well-performing subjects.

- Noy and Zhang (2023) design an online experiment where individuals in a range of white-collar occupations are recruited and presented with simple writing tasks (in particular, tasks like writing press releases, short reports and analyses that are designed to take 20 to 30 minutes and resemble real life tasks in the participants' occupations). The treatment group is given access to and encouraged to use ChatGPT-3.5, while the control group is not. They verify that there is very low usage of ChatGPT in the control group. They estimate that ChatGPT enables, on average, 40% faster completion of the task at hand. They also estimate an 18% improvement in quality scores, as judged by peers and ChatGPT-based scoring of the output. In their case, too, the gains come mostly from subjects that performed less well before the experiment.

- Brynjolfsson et al. (2023) is the only study that I am aware of that looks at the use of generative AI tools in a real business setting with a careful experimental design. The business they focus on is a customer service provider, which uses a custom generative AI tool to help customer service associates. The rollout took the form of a treatment group getting access to this tool, while the control group did not. Brynjolfsson, Li and Raymond evaluate the impact of the rollout on cost savings, by focusing on how quickly tasks (open customer tickets) are resolved. They also look at self-reported customer satisfaction. They find a significant improvement in the speed with which tasks are completed by customer service associates—an effect of about 14% on average. However, they additionally estimate a slight and statistically insignificant decline in

quality, as judged by the users themselves. Like the other studies, Brynjolfsson, Li and Raymond find that the results are predominantly among the lower-performing, less expert employees. In fact, they estimate that the top quintile of associates experience no improvement at all.

I interpret all three studies as providing labor cost savings from AI, broadly construed—in particular, meaning that all three studies include both automation and task complementarity elements. For instance, for the GitHub Copilot users in Peng et al., the authors' interpretation is that some of the subtasks previously performed by programmers, such as the writing of common routines, are now done by the Copilot. Along the lines of the discussion in Section 2.2, suppose that programming in JavaScript involves N subtasks, which need to be completed for a successful program. These include initial planning (which approach to adopt, how to organize the program, etc.), composition of subroutines, putting the subroutines together, debugging the subroutines, debugging the master program and then assessing whether the program achieves the planned aims. When all of these subtasks are performed, then the task at hand—the writing of a specific computer program—is completed. We can think of generative AI as taking over a subset of the composition of subroutine tasks. Then in line with the framework in Acemoglu and Restrepo (2018, 2019b), the use of this new technology will create displacement and productivity effects. Overall performance will increase because of productivity effects, provided that this technology is better than/faster than humans at composing some subroutines (but humans are still needed for the other subtasks, including planning, debugging and checking). Demand for human labor in this task may increase or decrease depending on the magnitude of the displacement and productivity effects and the elasticity of substitution between this task and other tasks (as well as the demand elasticity for the product that is ultimately being produced for the market). In addition, when simple subroutines are taken over by the Copilot, this may enable human workers to specialize in other subtasks, potentially generating task complementarities and further productivity gains.[23]

What I have just described for programming also applies to the other two studies. Chat-GPT is doing some of the drafting, which human subjects can then take and incorporate into

_____

[23]However, in this experimental setting, we do not see the displacement effects, because each of the treatment subjects are given the task and there is no possibility of reducing the number of workers performing this task. Hence, from these experiments, we can only learn about the productivity gains, inclusive of any task complementarities—and not about the displacement effects. This interpretation clarifies why I am comfortable bundling automation and task complementarities together for the purposes of estimating productivity gains from AI.

their writing with some verification and modification. The same applies for the customer service associates, who are allowed to copy and paste text suggested by the generative AI tool in the setting studied by Brynjolfsson, Li and Raymond.

I also assess that the tasks in these three studies are broadly comparable to the exposed tasks considered in Eloundou et al. (2023), even though I argue later that they are more likely to be in the easy-to-learn category. Hence, these studies are most informative about cost savings (productivity improvements) for the exposed tasks in Eloundou et al.

As a baseline, I ignore the quality effects (which, as noted, are not uniform between the three studies) and focus on the average increase in speed and interpret this as average cost savings. I return to the inequality implications below. Finally, as a baseline, I use the average of the estimates from Noy and Zhang (2023) and Brynjolfsson et al. (2023), and turn to the average of the three studies as a robustness check. The reason for this is that Peng et al.'s setting is less likely to be relevant to other tasks and occupations, since the task in question is a very well-defined one for which Github Copilot was extensively trained, and this has no direct equivalent in the other tasks we are focusing on. Under these assumptions, the average labor cost savings are 27% (= 0.27).

Recall that these numbers refer to declines in labor costs in occupations, where many tasks involve use both capital and labor, and what is relevant in equation (14) is overall cost savings. To convert labor cost savings into overall cost savings, I use the industry-level estimates from Eloundou et al. and combine them with industry labor shares from the Bureau of Economic Analysis (BEA), as described in the Appendix.[24] This gives an average (AI exposure-adjusted) labor share of 0.535, and thus the average (overall) cost savings from AI are about $0.27 \times 0.535 = 0.144$.

If we add the numbers from Peng et al. to the mix, the average labor cost savings become 0.36, and thus the average overall cost savings come to 0.193.

---

[24]Briefly, I follow Eloundou et al. and map exposed tasks to occupations, and then map occupations to 14 aggregated NIPA (National Income and Product Accounts) industries using the fractions of workers in each occupation employed across industries. I then compute the labor share of exposed tasks using industry labor shares weighted by the value-added shares of these industries in gross national income. In these calculations, the self-employed are assumed to have the same hourly wage rate as employees in the same industry. I use the 14 NIPA industries to be able to have information on hours worked. These data are for the years 2018-2022 from the US BEA NIPA tables, and I provide a list of these 14 industries in the Appendix.

For reference, the average (value-added-weighted) labor share in 2018-2022, without the exposure adjustment, is 0.576. This suggests that exposed tasks are, on average, in industries with slightly lower labor shares than the national average. This labor share estimate is similar to the one reported in Elsby et al. (2013), 0.583, for private non-farm business sector for the years 2010-2012 using the same approach as here. Elsby et al. (2013) also report similar numbers under alternative methodologies for incorporating self-employment income.

3.2   Aggregate Productivity Gains: A First Pass

A first-pass estimate of productivity (TFP) gains at the aggregate level can be obtained simply by combining the numbers derived in the previous subsection with equation (14). This implies:

$$\text{TFP gains over 10 years} \quad = \quad \text{GDP share impacted by AI over the next 10 years}$$
$$\times \; \text{average cost savings of impact tasks.}$$

Therefore, on the basis of the numbers from the previous subsection, I can approximate:

$$\text{TFP gains over the next 10 years} \quad = \quad 0.046 \times 0.144$$
$$\simeq \quad 0.0066.$$

In other words, according to this basic estimation strategy, TFP gains over the next 10 years from AI are about 0.66%—meaning that relative to the baseline without the current suite of AI and computer vision advances, TFP will be higher by 0.66 percentage points in 10 years, or annual TFP growth will be higher by around 0.064%. This is a nontrivial, but modest effect, and certainly much less than both the revolutionary changes some are predicting and the less hyperbolic but still substantial improvements forecast by Goldman Sachs and the McKinsey Global Institute, which I discussed in the Introduction.

If we were to consider the higher productivity numbers from Peng et al., 0.144 would be replaced by 0.193, and the 10-year TFP gains would be 0.89%, instead of 0.66%.

The only modification that would make a sizable difference to these numbers is to increase the fraction of tasks that will be impacted over the next 10 years. One way of doing this is to inflate the numbers from Svanberg et al.. This could be because either the fraction of tasks that can be feasibly automated will be different for generative AI than for computer vision, or because within 10 years this fraction will increase significantly. For example, in Svanberg et al.'s scenarios where costs for computer vision decline very rapidly, such as 10% a year, the fraction of tasks that are feasibly automated may be as high as 30%. This would raise the GDP share impacted by AI to approximately 6%, and correspondingly increase the TFP gains over the next 10 years to about 0.9%. Note that even this number is quite modest, and moreover, 10% per annum cost declines is quite aggressive (since, as already noted above, even if GPU costs were to decrease by 10% or even 20%, this would not lead to a 10% decline in costs of performing computer vision tasks, given the presence of other inputs, such as programming and data, as well as the inherent limitations of the current generative AI architecture).

Finally, I note three important considerations missing from these computations.

1. These adoption numbers ignore the fact that there is still very little investment in AI in the US corporate sector. Acemoglu et al. (2022) estimate that less than 1.5% of US businesses had any investment in AI in 2019, and this is particularly true beyond the very large companies in manufacturing, information services and business services. Since many of the tasks considered in Eloundou et al. (2023) are performed in small and medium-sized enterprises, this is unlikely to change quickly. If generative AI tools become monopolized in the hands of a few companies, this might further slow down their adoption by small and medium-sized firms. These considerations suggest that even the 0.046% number for the share of GDP impacted by AI may be a big overestimate, and the true numbers could be much smaller.

2. Any major technology creates adjustment costs when adopted at large scale, because other organizational aspects need to evolve as well and this is typically quite costly and slow. In the context of digital technologies, Greenwood and Yorukoglu (1997) and Brynjolfsson et al. (2021), among others, have argued that productivity gains will take a J-shaped pattern, and the former paper predicts that the flat part of the J-curve lasts no less than 20 years for digital technologies. If so, the 14.4% overall cost reductions may be a significant overestimate for the next 10 years.

3. As already discussed above, some of the tasks in the list of Eloundou et al. (2023) are hard-to-learn, where productivity gains may be significantly less than those based on the experimental studies that have focused on the easy-to-learn tasks.

In the next subsection, I make a preliminary attempt at incorporating the third possibility, but I will ignore the first two. Nevertheless, these considerations make me conclude that even the 0.66% increase in TFP within the next 10 years due to AI is likely to be an upper bound on this technology's medium-run effects.

3.3   Aggregate Productivity Gains: Incorporating Hard Tasks

In this subsection, I refine the estimates from the previous subsection by switching to equation (15), which can be rewritten as:

$$
\begin{aligned}
\text{TFP gains} \quad = \quad & \text{GDP share of impacted easy tasks } \times \text{ average cost savings from easy tasks} \\
& + \text{GDP share of impacted hard tasks } \times \text{ average cost savings from hard tasks.}
\end{aligned}
$$

I take the 27% labor cost savings, from Noy and Zhang (2023) and Brynjolfsson et al. (2023), to apply to easy-to-learn tasks. The discussion in the previous section suggests that most hard tasks have not been impacted or automated yet, and hence it is impossible to know what their cost savings will be. Here I take those productivity gains to be 7%. My reasoning is as follows. I consider the tasks involved in the Peng et al. (2023) study to be very easy-to-learn for generative AI for reasons explained above. Those studied in Noy and Zhang (2023) are also on the easy side, and led to cost savings of about 40%, which is about two thirds of the cost savings in Peng et al., while the customer service tasks in Brynjolfsson et al. (2023) are already moving towards somewhat more complex tasks, and these had cost savings of only 14%. I imagine that many of the hard-to-learn tasks are more challenging for AI models than the simpler end of the customer service tasks to which Brynjolfsson et al.'s numbers refer. This motivates my choice of half of the cost savings of their study, 7% (which is also about a quarter of the baseline 27% cost reduction estimate I used in the previous subsection).

The cost-saving numbers for both easy and hard tasks are again multiplied with 0.535 to convert them into overall cost savings—0.144 and 0.037 for easy and hard tasks, respectively. To obtain the shares of easy and hard tasks, I start from Eloundou et al.'s data and methodology, and then develop a procedure, implemented using GPT-4 like they do, for sorting these into easy and hard tasks. The key characteristic for easy tasks is the presence of a well-observed outcome and a straightforward rule that links actions/recommendations to characteristics of the problem at hand.

To implement this procedure, I start from the 4089 exposed tasks determined from the above procedure. Each one of these tasks has a statement on O*NET that includes at least one verb. Each task also belongs to the higher level of aggregation of Detailed Work Activites (DWAs) and the (even coarser) 332 Intermediate Work Activities (IWAs). The procedure proceeds in four steps:

1. **Classification of verbs**: Each task statement includes at least one verb which is located at the beginning and can be easily identified. These primary verbs describe much of what is "hard" or "easy" that needs to be learned about the task. We classified tasks manually into easy and hard categories.[25] The full list of tasks is provided in the Appendix. As an illustration, verbs for easy tasks include, among others: compute, resolve, count, draft, grade, transcribe, classify, standardize, write, and record. Many

---

[25]This step and other manual coding steps in this subsection were carried out by Can Yeşildere.

of these verbs are associated with simple actions that follow a clear set of steps and also implicitly have a well-defined metric for success (such as accounting or grading). In contrast, verbs for hard tasks include: participate, advise, instruct, diagnose, educate, hire, represent, testify, and care. The latter set of verbs describes more open-ended activities for which there is less likely to be a clear metric of success. Yet other verbs, such as analyze, maintain or inspect, do not fall into either of these categories, and are coded as "uncertain".

2. **Classification of IWAs:** IWAs provide additional context for verbs, especially for actions that verbs alone lack. We manually classify the 332 IWAs into easy and hard tasks. Some of the easy IWAs are: Evaluate project feasibility; Maintain sales or financial records; Explain regulations, policies, or procedures; Issue documentation; and Teach safety procedures or standards to others. Some of the hard ones include: Monitor health conditions of humans or animals; Evaluate scholarly work; Evaluate the quality or accuracy of data; Maintain safety or security; and Train animals. These examples highlight the same principle mentioned in the discussion of verbs—easy-to-learn activities are those for which there is a clear metric of success and simple rules that can achieve this successful outcome, while these are absent in hard-to-learn activities.[26]

3. **Latent Dirichlet Allocation (LDA) topic modeling**: While tasks may share IWAs and verbs, each task is also worded uniquely to capture the subject and the more detailed description of activities. While "writing an audit report" and "writing a letter of recommendation" are both associated with the same verb and the same IWA, the content and skills required for these tasks are different. To better disambiguate the contexts of these tasks, we use the LDA to allocate tasks into clusters. LDA procedure is unsupervised and clusters tasks using the co-occurence matrices of words extracted from each task statement. It assigns a probability to each task belonging to a topic cluster (Blei et al., 2003). We feed all $19,281$ tasks into the algorithm and use LDA to identify 100 clusters and the probability that each task belongs to one of those 100 clusters.

4. **Final assignment:** The final step of our procedure is to derive a probability that

---

[26]Of course, there is considerable ambiguity in some cases. For example, "Maintain safety and security" also includes IT security activities, such as making sure that a new password can be issued to an authorized person, and this would be an easy-to-learn task. The third step is aimed at dealing with such ambiguities.

each exposed task is easy or hard. We first manually classify a random sample of 500 exposed tasks as easy or hard. We then use the classification of verbs, classification of IWAs, and the LDA-derived probabilities to train a gradient-boosted tree (Friedman, 2001) to match the manual classifications of the training sample of tasks. We finally obtain a probabilistic assignment of each one of the 4089 exposed tasks into easy or hard task category from this algorithm.[27]

The end result is that 72.6% of wage bill-weighted exposed tasks in Eloundou et al. are easy. This means that easy exposed tasks comprise 3.3% of GDP, while the remaining exposed tasks (making up 1.3% of GDP) are hard. Incorporating this information together with the assumptions on cost savings yields the following tighter upper bound on TFP gains in the next 10 years:

$$
\begin{aligned}
\text{TFP gains over the next 10 years} \quad &= \quad 0.033 \times 0.144 + (0.046 - 0.033) \times 0.037 \\
&\simeq \quad 0.0053.
\end{aligned}
$$

Unsurprisingly, this is smaller than the estimate in the previous subsection. Since automating hard-to-learn tasks—and even more so, introducing task complementarities for them—will be more challenging, I view this estimate to be more reasonable. Either way, the TFP gains within the next 10 years appear quite modest.

3.4   From TFP to GDP

Given the TFP effects of AI over the 10 year horizon, the total GDP effects of AI over the same horizon can be computed from equation (13). Namely, this equation can be written as:

GDP gains over the next 10 years =   TFP gains over the next 10 years

+ capital share $\times$ proportional increase in capital stock.

As a simple benchmark, I start by assuming that the capital stock will grow proportionately with TFP, which implies:

GDP gains over the next 10 years = TFP gains over the next 10 years $/(1-$ capital share$)$.

---

[27]As examples, the algorithm assigns a probability of 0.98 that "Maintaining equipment service records" is an easy task, while only a probability of 0.098 that "Interviewing credit card applicants by telephone or in person" is an easy task. Because writing and drafting are classified as verbs associated with easy tasks, activities that involve writing are classified generally as easy. Some of this may understate the difficulty of some writing-related tasks. For example, "Writing reports or academic papers to communicate findings of climate-related studies" is classified to be an easy task with probability 0.67, which is likely to be an underestimate of how hard such a task will continue to be for AI in the foreseeable future.

Using the capital share for the entire private business sector, 0.43, this implies that GDP gains will be equal to the TFP gains multiplied by 1.75 ($\simeq 1/(1 - 0.43)$). Hence taking the baseline estimate of an increase in TFP of 0.66%, I obtain a first estimate for GDP growth due to AI of 1.16% over 10 years, or taking the presence of hard tasks into account, a lower estimate of 0.93%.

I update the GDP effects of AI advances in Section 3.6 when I model the between-task and between-industry substitution patterns and the resulting investment response.

3.5 Consequences of New Bad Tasks

The calculations so far leave out the effects of new tasks introduced thanks to AI (and equivalently, the system-wide adjustments that AI may enable in some businesses, as emphasized by Bresnahan (2019) and Agrawal et al. (2023). It is even more challenging to put numbers on the effects of new tasks. If AI helps create new tasks that increase productivity and especially contributes to the reinstatement of workers of different skill levels into the production process, its consequences can be much more positive.

Here my purpose is simply to point out that in the case of AI, because some of the new tasks may be of the "bad" type, there may be an overstatement of the welfare gains from AI when we look at GDP numbers. As an extremely preliminary attempt to argue that this could be important, I will draw on two sources of data.

The first is the recent study by Bursztyn et al. (2023), which provides suggestive estimates for the extent of this problem in the case of social media. Bursztyn et al. (2023) run an experiment to evaluate the welfare effects of social media for college students, the demographic group most engaged with social media. They ask TikTok and Instagram users how much they would need to be paid to not use the platform for a month. Users of these two platforms are willing to pay up to $59 and $47 a month, respectively, to continue to use social media—or on average $53 per user-month. However, they find that users are also willing to pay $28 and $10 every month to get everyone from their social network off TikTok and Instagram, respectively—or an average $19 per user-month. If non-users are also included in the analysis, the willingness to pay to stop everybody using social media increases to $47 for TikTok and $13 for Instagram. This suggests that while companies can profitably market AI-based social media, the net effect on welfare may be negative. Quantitatively, I ignore the non-users (since the population in the study is college students, who are more likely to be impacted by social media activity even when they do not use it than the adult population in

the United States). I also start with a benchmark in which AI-powered platforms can capture the full (average) $53 willingness to pay per user (for example, because they are effectively price discriminating by varying the intensity with which they are collecting and monetizing users' data). Then taking the average between the two platforms, I conclude that for every $53 of revenue, there is a net negative effect on users of $19. Put differently, the total effect from this class of new bad tasks, the equivalent of (18) in theory, is $-19/53 \simeq -0.36$ in proportionate terms. Note also that if I assumed instead that social media companies can capture less than the full $53 per user-month, then the denominator in this expression would be smaller, and thus the proportionate damage per dollar of revenue would be higher.

I combine the number -0.36 with estimates of (i) revenues from social media and digital ads, and (ii) spending on malicious IT attacks and IT security against these attacks to compute apparent and real gains from new bad tasks. Specifically, the total revenues in 2022 of Meta (Facebook plus Instagram), Alphabet (Google and Youtube), Snapchat, TikTok and Twitter ("X") comes to 460 billion, or about 1.64% of US GDP,[28] while a lower bound on IT security is 78 billion or 0.28% of US GDP.[29] Assuming that spending by malicious actors on IT attacks is at least one third of this and combining these three estimates, I arrive to 2% of US GDP. This number could be a significant underestimate of other manipulative activities enabled by new AI technologies. On the other hand, it could also overstate the problem, since only a fraction of digital ad revenues will come from such manipulative activities (and this may be particularly true for digital ad revenues from Google search). These considerations suggest that the numbers here should be taken as merely suggestive.

Taking the 2% of US GDP as revenues from new bad tasks and using the numbers from Bursztyn et al. (2023) suggests that the total negative effects of these manipulative activities is $0.02 \times 0.36 = 0.072\%$ of GDP (in consumption equivalent units). This number points to a sizable negative impact on welfare, and could be larger if manipulative uses of generative AI becomes more widespread. In contrast, if one were to simply count the revenue coming from these new tasks, one might conclude that they would increase GDP by 2%. This discussion thus suggests caution in interpreting all increases in GDP coming from the use of generative AI as a positive effect on consumer welfare.

---

[28]The numbers for Meta ($130 billion), Alphabet ($307 billion) and Snapchat ($4.6 billion) are from these companies' 2023 10K filings. Fortune reported on December 12, 2023 that X's annual revenue was $2.5 billion in 2023, while Financial Times reported on March 15, 2024 that revenues from TikTok's US business had reached $16 billion per annum (see Fortune and Reuters).

[29]See Statista.

3.6 Wage and Inequality Implications

Finally, I evaluate the wage and inequality consequences of generative AI advances. To do this, the present framework needs to be extended to include multiple demographic groups that have different comparative advantages across different tasks, as in Acemoglu and Restrepo (2022). To save space, I do not introduce this generalization and refer the reader to that paper. Instead, I start with the following equation from their paper, which is a generalization of (11) above to a setting with multiple sectors and multiple demographic groups (but, for simplicity, without the task complementary and labor-augmenting changes):

$$d \ln w_g = \Theta_g \cdot \left(\frac{1}{\sigma} d \ln y + \frac{1}{\sigma} d \ln \boldsymbol{\zeta} - \frac{1}{\sigma} d \ln \boldsymbol{\Gamma}^{\mathrm{auto}} \right).$$

Here, g refers to demographic group g, and following their paper, I will focus on 500 demographic groups, defined by education, age group, gender, ethnicity and native vs. foreign-born status. In addition, $d \ln y$ is the change in GDP resulting from the technology change (capturing the productivity effect), σ is the elasticity of substitution between tasks, and $\boldsymbol{\zeta}$ is the vector of induced industry shifts (e.g., because automation of tasks in one industry affects prices and causes a reallocation of spending across sectors). Most importantly, $d \ln \boldsymbol{\Gamma}^{\mathrm{auto}}$ is the vector of demographic group-level displacement caused by the technology shock—namely, it is a column vector of 500 entries, and Θ_g is the gth-row vector of the propagation matrix, which summarizes how the displacement of other demographic groups affects demographic group g (this is the reason why it is pre-multiplying the vectors of industry shifts and displacements for all groups).

The propagation matrix represents the full "ripple effects"—the impact of the displacement of one demographic group on others, as they leave the tasks they were previously performing and compete with other groups to be employed in other tasks. Such reallocations are the key channel via which direct productivity gains for a group may end up harming it (as my example in the previous section illustrated). They are also the mechanism via which the displacement of a demographic group may end up being more damaging to another demographic group.

The ripple effects estimated in Acemoglu and Restrepo (2022) may be context-specific—meaning that the magnitudes of these effects could be quite different for the tasks impacted by AI and automation technologies that my previous work focuses on. Nevertheless, since it is impossible to estimate these ripple effects for the future impact of generative AI technologies, I will use these existing estimates.

Figure 1: Distribution of AI exposure across the wage distribution
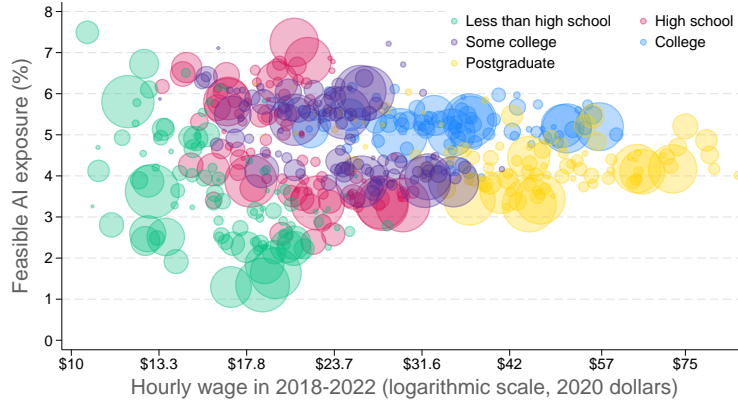
Notes: This figure depicts the AI exposure measure (both easy and hard tasks, combined) across 500 demographic groups. The horizontal axis gives the average hourly wage of each demographic group between 2018 and 2022, computed from the American Community Survey five-year sample. Marker sizes are proportional to the average 2018-2022 employment level of each group and different colors indicate the education level of the group.

An additional issue is that I have so far interpreted AI exposure to include both automation and task complementarities. In this subsection, I ignore the task complementarities and presume that all of the AI exposure quantified so far will take the form of automation.[30]

The methodology in Acemoglu and Restrepo (2022) starts from the $d \ln \mathbf{\Gamma}^{\text{auto}}$ vector, which is at the demographics group level. To construct an equivalent of this measure, I take the set of exposed occupations, and then use the wage bill shares of different demographic groups in these occupations to map the AI-generated displacement to the demographic group level. For example, if for demographic group g 5% of the wage bill share in 2019-2022 were in fully-exposed occupations, then $d \ln \Gamma_g^{\text{auto}}$ would be 0.05. I also assign these occupations to industries using wage bill shares in order to compute industry-level impacts on costs and prices.[31] I compute the induced sectoral reallocations in the same way as in Acemoglu and Restrepo (2022), using their parameterization of inter-sector elasticity of substitution and

---

[30]If there were task complementarities, this would increase the productivity of exposed demographic groups in the remaining subtasks, but would also reduce the automation-driven cost savings by a corresponding amount. If task complementarities and automation affected different demographic groups within occupations symmetrically, this would have minor effects on the conclusions: although the wage effects of productivity gains from task complementarities are a little different than the cost savings generated by automation (Acemoglu and Restrepo, 2022), the results would remain broadly similar. If, on the other hand, some groups benefited more from task complementarities (for example, because they are overrepresented among middle-expertise workers that can benefit most from generative AI tools), the distributional consequences could be different. Since I do not have a way of distinguishing productivity effects from task complementarities and automation, I am unable to explore this issue further.

[31]This exercise uses the 49 BEA industries as in Acemoglu and Restrepo (2022).

the estimate of $\sigma = 0.50$ as in Humlum (2021) and take the elasticity of substitution between sectors in consumption to be $\eta = 0.2$ as in Buera et al. (2022), which was also imposed in Acemoglu and Restrepo (2022).

Figure 1 is the equivalent of Figure 5b in Acemoglu and Restrepo (2022) and presents the distribution of AI exposure across demographic groups sorted by their hourly wage in 2018-22. It shows that AI exposure is much more equally distributed across demographic groups than pre-AI automation (which was based on robotics, dedicated advanced machinery and software systems).

Table 1 presents the main results. The seven rows are for the five education groups (aggregating demographic groups according to their education level), for the average of the workforce and for GDP. The first column gives the AI exposure for each one of the demographic groups, using our baseline exposure measure that does not distinguish easy and hard tasks. The second column presents the direct impact of AI exposure on each one of these groups, while the third column contains the full wage impact, taking into account induced substitution between industries and the ripple effects. The next four columns are similar but now refer to the AI exposure measure that separates easy and hard tasks—column 4 gives exposure to easy tasks and column 5 is for exposure to hard tasks. Finally, column 8 shows the exposure measure from Acemoglu and Restrepo (2022) for comparison (but recall that this measure refers to a 36-year period, rather than the 10-year timescale here). The comparison of columns 1 and 4-5 to column 8 confirms that AI exposure is much more equally distributed across demographic groups than pre-AI automation exposure. Workers with less than high school have the lowest exposure, followed by postgraduates, while workers with college degrees and those with associate degrees or some college have the highest exposure.

Consequently, predicted wage impacts do not appear to have a big impact on inequality between education groups. Focusing on the estimates in column 7, which incorporate easy and hard tasks, there is a slightly higher wage growth for workers with less than a high school degree—about 1.3% within 10 years—but the gap between postgraduate versus high school and college graduate workers also widens somewhat. In fact, the between-group standard deviation of log wages (weighted by employment) increases slightly, from 0.35 to 0.36. The results are quite similar in column 3 for the baseline AI exposure measure.

Baseline exposure | Exposure adjusted for easy and hard task

	(1) Baseline AI exposure	(2) Direct effect of AI exposure	(3) Total wage effect of AI exposure	(4) Exposure to easy tasks	(5) Exposure to hard tasks	(6) Direct effect of exposure to easy and hard tasks	(7) Total wage effect of exposure to easy and hard tasks	(8) Direct task displacement 1980-2016
Workers with less than high school degree	0.0318	-0.0323	0.0157	0.0235	0.0083	-0.0356	0.0132	0.2690
Workers with high school degree	0.0464	-0.0617	0.0106	0.0337	0.0128	-0.0649	0.0079	0.2706
Workers with some college	0.0494	-0.0676	0.0139	0.0357	0.0138	-0.0709	0.0112	0.1886
Workers with Bachelor's degree	0.0523	-0.0733	0.0060	0.0382	0.0140	-0.0765	0.0040	0.0684
Workers with postgraduate degree	0.0405	-0.0498	0.0098	0.0299	0.0106	-0.0530	0.0081	0.0343
Average worker	0.0462	-0.0612	0.0101	0.0336	0.0126	-0.0644	0.0077	0.2107
GDP			0.0156				0.0140	

Table 1: Exposure and wage effects by education groups

Note: This table summarizes the effects of AI exposure on the real wages of different education groups and all workers. The first three columns are for the AI exposure measure that does not distinguish easy and hard tasks, while columns 4-7 are for the measure that introduces this distinction. Columns 1, 4 and 5 present levels of AI exposure for these different cuts. Columns 2 and 6 provide the direct effects of AI exposure, while columns 3 and 7 include the full equilibrium wage impacts, incorporating the ripple effects. Column 8, for comparison, includes the direct task displacement measure for 1980-2016 from Acemoglu and Restrepo (2022). The rows are for different education groups as well as for the average worker and for the overall GDP impact. The propagation matrix estimates are for 1980-2016, as reported in column 2 of Table 8 in Acemoglu and Restrepo (2022).

Figure 2: Decomposition of productivity effects, industry shifts, direct displacement effects and ripple effects
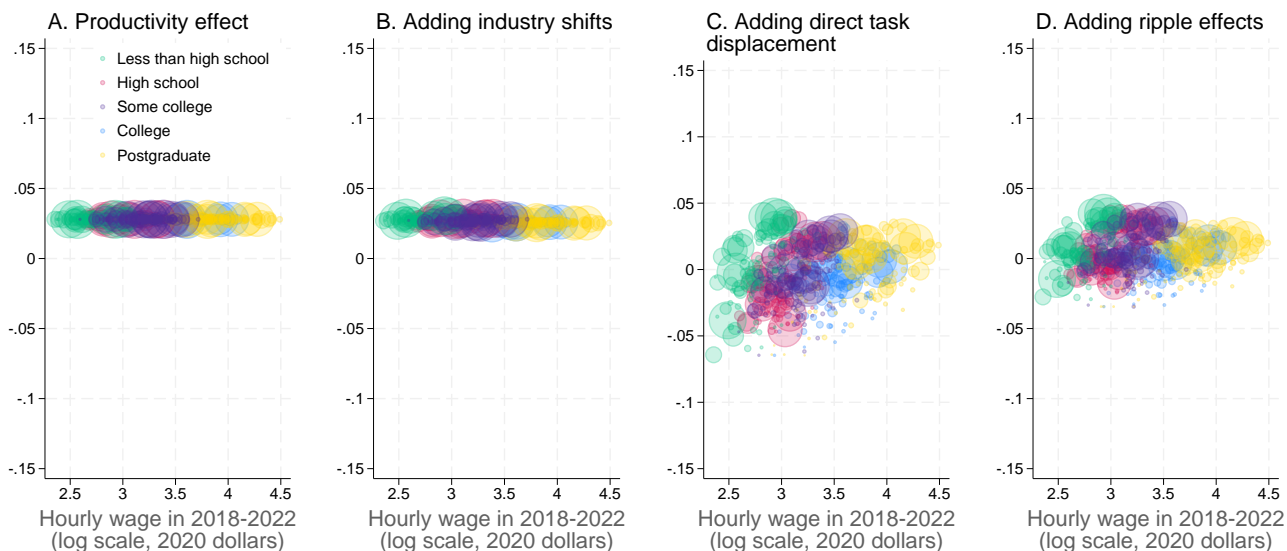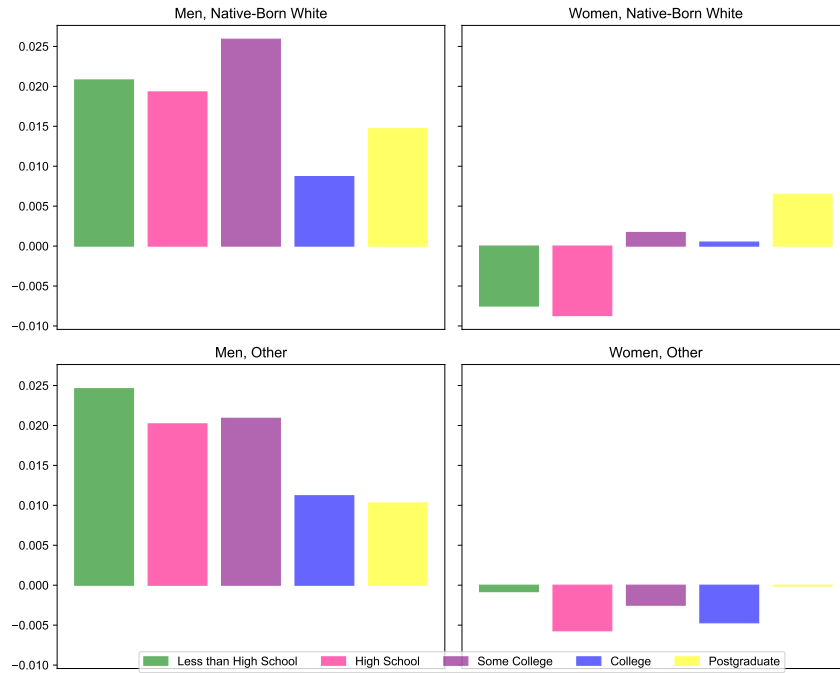

Notes: This figure is based on the estimates of the propagation matrix from Acemoglu and Restrepo (2022) and combines this with the measure of exposure to easy and hard AI tasks in this paper. The first panel includes just the productivity effect. The second panel adds the industry shifts induced by AI exposure. The third panel incorporates the direct displacement effect, while the final panel adds the ripple effects. The horizontal axis gives the average hourly wage for the relevant demographic group between 2018 and 2022, computed from the five-year American Community Survey sample. Marker sizes are proportional to the average 2018-2022 employment level of each group and different colors indicate the education level of the group. See text for details.

Finally, row 7 presents the estimate of the impact on GDP, taking into account the equilibrium increase in the capital stock implied by the model. This is under the assumption that all of the capital stock adjustment will take place within 10 years (while in practice it may take longer) and that the required rate of return on capital investments does not change (whereas with a sizable investment, we may expect an increase). This leads to an upper bound GDP impact of 1.4% in column 7, when I distinguish between easy and hard tasks (or 1.56% in column 3 when this distinction is not drawn). GDP therefore increases substantially more than average wages, and as a result, the capital share of national income increases by about 0.31 percentage points. This confirms that inequality between capital and labor is likely to rise as a result of the rollout of AI.[32] Figure 2 performs the same

---

[32]Recall, however, that it is the TFP gains that are the relevant numbers for consumer welfare, since the additional investment comes out of consumption and may involve additional heavy energy use, as remarked in footnote 6.

Figure 3: Total wage effect of exposure to AI, by gender

exercise as Figure 7 in Acemoglu and Restrepo (2022), focusing on the exposure measure that distinguishes between easy and hard tasks. As in that paper, productivity effects are (by construction) uniform across groups, and there is also not much inequality generated by the cross-industry shifts shown in the second panel. The third panel confirms that AI's direct effects are more equally distributed across demographic groups and throughout the wage distribution. In contrast to the findings in Acemoglu and Restrepo (2022), ripple effects do not change the inequality patterns by much, largely because the direct effects are already fairly equally distributed. Figure 2 also indicates that there is a lot of variability in the experience of low-education groups—and this is the reason why the between-group standard deviation of log wages increases, as noted above. Figure **??** explores this issue further by depicting the real wage changes of finer groupings, distinguished by gender, by education and by white and native-born status versus the rest. It reveals that low-education women, especially white, native-born low-education women, are likely to experience declines in real

wages as a result of AI.

Overall, this exercise suggests that the inequality consequences of AI will not be as adverse as pre-AI automation, because AI exposure is more equally distributed across demographic groups. Nevertheless, there is no evidence that AI will reduce inequality, as some are forecasting. Rather, my analysis suggests that it may have a small positive effect on overall (between-group) inequality and reduce the real earnings of low-education women. It will also further widen the gap between capital and labor income.

4   Concluding Remarks and Discussion

Following its release on November 30, 2022, ChatGPT became the fastest spreading tech platform in history, reaching approximately 100 million monthly users within just two months. Its impressive features, and the greater capabilities of the newer version ChatGPT-4, released in March 2023, soon captured imaginations, both among the general public and economic commentators. Forecasts of large productivity gains have now become commonplace.

While there is no question that generative AI models, including ChatGPT, have impressive achievements and have great potential for beneficial economic effects, the extent of their macroeconomic consequences remains an open question.

There are four potential types of macro effects that AI technologies can have in the medium run:

1. They can quickly revolutionize every aspect of the economy, and lead to massive improvements in productivity, even taking us close to "singularity". While this is a possibility that cannot be completely ruled out, there is so far no evidence of such revolutionary effects (Nordhaus, 2021), and these are not addressed in the current paper.

2. They can have more modest but still notable effects on the macroeconomy by improving productivity and reducing costs in a range of tasks. Some of the forecasts have focused on these types of improvements and still produced relatively large numbers, such as a $1.5 - 3.4$ percentage point per annum increase in economic growth within a 10 year horizon.

3. They can impact wages and inequality because of their automation effects, or conversely, lead to large wage increases, especially for lower-pay workers, as forecast by The Economist (2023).

4. They can have macroeconomic effects by producing deepfakes, misinformation, manipulation and other "bads".

In this paper, I use the task-based framework of Acemoglu and Restrepo (2018, 2019b, 2022) to evaluate the second and the third effects, and I also take some steps to formalize how the fourth set of effects might work out in a task-based macro framework.

I base my approach on existing experimental studies that estimate productivity gains and time savings from the use of generative AI tools in a number of settings. By building on these studies, I am explicitly taking on board the idea that generative AI will lead to productivity improvements. Nevertheless, combining these numbers with estimates of exposed tasks from Eloundou et al. (2023) and Svanberg et al. (2024) leads to much more modest productivity effects than most commentators and economists have claimed so far. These numbers become even smaller once we take into account that many of the tasks for which we have evidence of cost savings are relatively easy for AI, while in several other tasks the integration of AI will face more formidable difficulties—mostly because these are likely to involve more complex interactions between action and context and because they lack clear metrics for success that are observable, and hence necessitate AI models to learn from the (average) behavior of humans previously performing the same tasks.

Taking these considerations into account, I estimate that TFP effects from AI advances within the next 10 years will be modest—an upper bound that does not take into account the distinction between hard and easy tasks would be about a 0.66% increase in total within 10 years, or about a 0.064% increase in annual TFP growth. When the presence of hard tasks among those that will be exposed to AI is recognized, this upper bound drops to about 0.53%. GDP effects will be somewhat larger than this because automation and task complementarities will also lead to greater investment. But my calculations suggest that the GDP boost within the next 10 years should also be modest, in the range of 0.93% − 1.16% over 10 years in total, provided that the investment increase resulting from AI is modest, and in the range of 1.4% − 1.56% in total, if there is a large investment boom.

If AI is used to create new tasks and products, these will also add to GDP and can boost productivity growth. Nevertheless, when we incorporate the possibility that new tasks generated by AI may be manipulative, the impact on welfare can be even smaller. Based on numbers from Bursztyn et al. (2023), which pertain to the negative effects of AI-powered social media, I provide an illustrative calculation for social media, digital ads and IT defense-attack spending. These could add to GDP by as much as 2%, but if we apply

the numbers from Bursztyn et al. (2023), their impact on welfare may be -0.72%. This discussion suggests that it is important to consider the potential negative implications of AI-generated new tasks and products on welfare.

Finally, I borrow heavily from the estimates of Acemoglu and Restrepo (2022) on the economy-wide productivity, wage and inequality effects of pre-AI automation technologies to provide some guidance on what the impact of new AI advances will be. Because AI-exposed tasks are more equally distributed within the population than tasks exposed to pre-AI automation, I do not find substantial negative wage effects for any education group. Nevertheless, the estimates do not point to significant reductions in inequality either, and in fact, my findings suggest that low-education women may experience small wage declines, overall between-group inequality may increase slightly, and the gap between capital and labor income is likely to widen further.

These results should not be interpreted as arguing that there are no major benefits from AI. First, an increase of about $0.53 - 0.66\%$ in TFP within 10 years is modest but still far from trivial. Second and more importantly, there may be other ways in which AI can be used to generate much more notable benefits. I have suggested in previous work (Acemoglu, 2021, Acemoglu and Restrepo, 2020b) that if AI is used for generating new tasks for workers, it can have more beneficial productivity, wage and inequality consequences, and it can even increase wages. This may be doubly true for generative AI, which could be used for providing better information to workers and boosting their expertise, as argued in Acemoglu et al. (2023) and explained briefly here.

Many production workers today, including electricians, repair workers, plumbers, nurses, educators, clerical workers, and increasingly many blue-collar workers in factories, are engaged in problem-solving tasks. These tasks require real-time, context-dependent and reliable information. For instance, an electrician dealing with the malfunctioning of advanced equipment or a short-circuit on the electricity grid will be hampered from solving these problems because he or she does not have sufficient expertise and the appropriate information for troubleshooting. Reliable information that can be provided quickly by generative AI tools can lead to significant improvements in productivity. Similarly, generative AI in classrooms can lead to a major reorganization of how teaching takes place, with greater levels of personalization, as these tools help teachers identify specific aspects of the curriculum with which subgroups of students are having problems and propose new context-dependent teaching strategies. Reliability of AI models will be key for successfully creating such new

tasks and delivering improvements in the quality of education.

Productivity improvements from new tasks and new products, which have been important for previous transformative technologies, such as electricity and the Internet, are not incorporated into my estimates. This is for three reasons. First and most parochially, this is much harder to measure and is not included in the types of exposure considered in Eloundou et al. (2023) and Svanberg et al. (2024). Second, and more importantly, I believe it is right not to include these in the likely macroeconomic effects, because these are not the areas receiving attention from the industry at the moment, as also argued in Acemoglu (2021), Acemoglu and Restrepo (2020b) and Acemoglu and Johnson (2023). Rather, areas of priority for the tech industry appear to be around automation and monetization of personal data, such as through search or social media digital ads. This makes it less likely that many new good tasks will be created quickly.Third, and relatedly, more beneficial outcomes may require new institutions, policies and regulations, as also suggested in Acemoglu and Johnson (2023) and Acemoglu et al. (2023).

My assessment is that there are indeed much bigger gains to be had from generative AI, which is a promising technology, but these gains will remain elusive unless there is a fundamental reorientation of the industry, including perhaps a major change in the architecture of the most common generative AI models, such as the LLMs, in order to focus on reliable information that can increase the marginal productivity of different kinds of workers, rather than prioritizing the development of general human-like conversational tools. The general-purpose nature of the current approach to generative AI could be ill-suited for providing such reliable information. To put it simply, it remains an open question whether we need foundation models (or the current kind of LLMs) that can engage in human-like conversations and write Shakespearean sonnets if what we want is reliable information useful for educators, healthcare professionals, electricians, plumbers and other craft workers.

References

Acemoglu, Daron, "AI's Future Doesn't Have to Be Dystopian.," *Boston Review*, 2021. https://www.bostonreview.net/forum/ais-future-doesnt-have-to-be-dystopian/.

＿ **and David Autor**, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in "Handbook of Labor Economics," Vol. 4B, Elsevier, 2011, pp. 1043–1171. edited by David Card and Orley Ashenfelter. Amsterdam: North-Holland.

＿ **and Pascual Restrepo**, "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, 2018, *108* (6), 1488–1542.

＿ **and ＿** , "Artificial Intelligence, Automation, and Work," *The Economics of Artifical Intelligence: An Agenda*, 2019, pp. 197–236. Edited by Ajay Agrawal, Joshua Gans and Avi Goldfarb. Chicago, USA: University of Chicago Press.

＿ **and ＿** , "Automation and New Tasks: How Technology Displaces and Reinstates Labor," *Journal of Economic Perspectives*, 2019, *33* (2), 3–30.

＿ **and ＿** , "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy*, 2020, *128* (6), 2188–2244.

＿ **and ＿** , "The Wrong Kind of AI? Artificial Intelligence and the Future of Labour Demand," *Cambridge Journal of Regions, Economy and Society*, 2020, *13* (1), 25–35.

＿ **and ＿** , "Tasks, Automation, and the Rise in US Wage Inequality," *Econometrica*, 2022, *90* (5), 1973–2016.

＿ **and ＿** , "Automation and Rent Dissipation: Implications for Inequality, Productivity, and Welfare," 2023. MIT Working Paper.

＿ **and Simon Johnson**, *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity*, London, UK: Hachette, 2023.

＿ **, David Autor, and Simon Johnson**, "Can We Have Pro-Worker AI?," Technical Report, MIT Shaping the Future of Work Initiative 2023.

＿ **, Gary W Anderson, David N Beede, Cathy Buffington, Eric E Childress, Emin Dinlersoz, Lucia S Foster, Nathan Goldschlag, John C Haltiwanger,**

Zachary Kroff et al., "Automation and the Workforce: A Firm-Level View from the 2019 Annual Business Survey," Technical Report, National Bureau of Economic Research Working Paper 30659, 2022.

_ , **Martin K. Jensen, and Pascual Restrepo**, "Technology and Wages in the Long Run," 2024. Work in Progress.

Agrawal, Ajay, Joshua S Gans, and Avi Goldfarb, "Artificial Intelligence Adoption and System-Wide Change," *Journal of Economics & Management Strategy*, 2023, pp. 1–11.

Autor, David, "Applying AI to Rebuild Middle Class Jobs," Technical Report, National Bureau of Economic Research Working Paper 32140, 2024.

_ , **Caroline Chin, Anna Salomons, and Bryan Seegmiller**, "New Frontiers: The Origins and Content of New Work, 1940–2018," *The Quarterly Journal of Economics*, 2024. Forthcoming.

_ , **Frank Levy, and Richard J Murnane**, "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 2003, *118* (4), 1279–1333.

Besiroglu, Tamay and Marius Hobbhahn, "Trends in GPU Price-Performance," Technical Report, EpochAI 2022. https://epochai.org/blog/trends-in-gpu-price-performance.

Blei, David M, Andrew Y Ng, and Michael I Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 2003, *3*, 993–1022.

Bresnahan, Timothy, "Artificial Intelligence Technologies and Aggregate Growth Prospects," *Prospects for Economic Growth in the United States*, 2019, pp. 132–172.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson, "The Productivity J-curve: How Intangibles Complement General Purpose Technologies," *American Economic Journal: Macroeconomics*, 2021, *13* (1), 333–372.

_ , **Danielle Li, and Lindsey R Raymond**, "Generative AI at Work," Technical Report, National Bureau of Economic Research Working Paper 31161, 2023.

Buera, Francisco J, Joseph P Kaboski, Richard Rogerson, and Juan I Vizcaino, "Skill-Biased Structural Change," *The Review of Economic Studies*, 2022, *89* (2), 592–625.

Bursztyn, Leonardo, Benjamin R Handel, Rafael Jimenez, and Christopher Roth, "When Product Markets Become Collective Traps: The Case of Social Media," Technical Report, National Bureau of Economic Research Working Paper 31771, 2023.

Chui, Michael, Eric Hazan, Roger Roberts, Alex Singla, Kate Smaje, Alex Sukharevsky, Lareina Yee, and Rodney Zemmel, "The Economic Potential of Generative AI: The Next Productivity Frontier," *McKinsey & Company*, 2023. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-AI-the-next-productivity-frontier#business-value.

Davidson, Tom, "Could Advanced AI Drive Explosive Economic Growth?," *Open Philanthropy*, June 2021. https://www.openphilanthropy.org/research/could-advanced-ai-drive-explosive-economic-growth/.

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock, "GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models," Technical Report, arXiv 2023.

Elsby, Michael WL, Bart Hobijn, and Ayşegül Şahin, "The Decline of the US Labor Share," *Brookings Papers on Economic Activity*, 2013, *2013* (2), 1–63.

Friedman, Jerome H, "Greedy Function Approximation: a Gradient Boosting Machine," *Annals of Statistics*, 2001, pp. 1189–1232.

Goldman Sachs, "Generative AI could raise global GDP by 7 percent," 2023. https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html.

Greenwood, Jeremy and Mehmet Yorukoglu, "1974," *Carnegie-Rochester Conference Series on Public Policy*, 1997, *46*, 49–95.

Hulten, Charles R, "Growth Accounting with Intermediate Inputs," *The Review of Economic Studies*, 1978, *45* (3), 511–518.

Humlum, Anders, "Robot Adoption and Labor Market Dynamics," 2021. University of Chicago Working Paper.

Korinek, Anton and Donghyun Suh, "Scenarios for the Transition to AGI," Technical Report, National Bureau of Economic Research Working Paper 32255, 2024.

Kurzweil, Ray, *The Singularity Is Near: When Humans Transcend Biology*, New York City, USA: Viking Press, 2005.

Nordhaus, William D, "Are We Approaching an Economic Singularity? Information Technology and the Future of Economic Growth," *American Economic Journal: Macroeconomics*, 2021, *13* (1), 299–332.

Noy, Shakked and Whitney Zhang, "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence," *Science*, 2023, *381* (6654), 187–192.

Peng, Sida, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer, "The Impact of AI on Developer Productivity: Evidence from Github Copilot," Technical Report, arXiv 2023.

Susskind, Daniel, *A World without Work: Technology, Automation and How We Should Respond*, London, UK: Penguin, 2020.

Svanberg, Maja, Wensu Li, Martin Fleming, Brian Goehring, and Neil Thompson, "Beyond AI Exposure: Which Tasks are Cost-Effective to Automate with Computer Vision?," 2024. MIT Working Paper.

The Economist, "Blue-Collar Bonanza: Why Conventional Wisdom on Inequality is Wrong," December 2023. https://www.economist.com/weeklyedition/2023-12-02.

Online Appendix

A-1 Exposed Share of Tasks

Eloundou et al. (2023) present task-level exposures estimates based on O*NET, which lists tasks by occupation at the level of the eight-digit Standard Occupational Classification (SOC). We link thess data to the BLS Occupational Employment and Wage Statistics (OEWS) for 2019 to 2022 to calculate the wage bill-weighted share of exposed tasks in the economy. BLS OEWS data are reported at the six-digit SOC level, and we combine these two sources of data as follows:

1. **Six-digit direct matching:** We compute occupational exposures scores at the eight-digit SOC level based on Eloundou et al. (2023). We then aggregate these eight-digit numbers to the six-digit SOC level (with equal weights). This direct match covers 90.8% of the total wage bill in the economy.

2. **Aggregation to five-digit:** In this step, we aggregate the six-digit SOC occupations to five-digits. For five-digit occupations that include both six-digit occupations with exposure estimates and wage bill information in BLS OEWS, we compute the wage bill share-weighted average of the six-digit occupations. This weighting scheme ensures that the information at the six-digit level is preserved when we aggregate up to five-digits. BLS OEWS does not report wage bill information for some small occupations, and as a result, for some five-digit occupations we have no direct match to six-digit occupations with exposure and wage bill data. In such cases, we compute the five-digit exposure by taking the unweighted average of all corresponding six-digit occupations with exposure data. These steps together cover 99.3% of the total wage bill in the economy. The remaining five-digit occupations, which make up 0.7% of the wage bill and have no exposure information in Eloundou et al. (2023), are assumed to have zero exposure. This aggregated occupational exposure score is denoted by Γ_0.

A-2 Industry Labor Share and Exposure-Adjusted Labor Share

Industry-level exposure to AI, Γ_i, is computed following the strategy used by Eloundou et al. (2023), leveraging industry-level occupation distributions and occupational exposure data.

Specifically:

$$\Gamma_i = \frac{\sum_{o \in O} n_{oi} \Gamma_0}{\sum_{o \in O} n_{oi}},$$

where Γ_o is occupation o's exposure to AI and n_{oi} are the number of people employed in occupation o in industry i, drawn from the BLS OEWS.

The labor share in non-farm private business income is computed in the usual manner as:

$$s_L = \frac{\sum_{i \in I} s_Y^i s_L^i}{\sum_{i \in I} s_Y^i},$$

where s_Y^i is the share of industry i's value added in non-farm private sector GDP, and s_L^i is industry i's labor share of value added. I follow the BLS and Elsby et al. (2013), and impute self-employment income in line with average hourly wage of employees in the same industry according to the number of hours worked of the self-employed. Since the number of hours worked is available only at the level of 14 aggregate NIPA industries, provided in BEA, I use these 14 industries for this exercise.[33] The labor share for 2018-2022 is estimated as 0.576. This number is similar to the estimates provided in Elsby et al. (2013) for the years 2010-2012, when they use the same methodology for imputing self-employment income. They also report a slightly higher labor share of 0.593, leveraging an alternative methodology based on the idea that the implied rate of return to capital used with employees has to be the same as the rate of return to capital used in self employment.

Exposure-adjusted labor share income, $\hat{s}_L$, is then calculated as:

$$\hat{s}_L = \frac{\sum_{i \in I} \Gamma_i s_Y^i s_L^i}{\sum_{i \in I} \Gamma_i s_Y^i}.$$

The only difference from the labor share computer previously, s_L, is the adjustment by the level of industry exposure, Γ_i. This leads to a slightly lower number than s_L, 0.536.

---

[33]Specifically, we averaged data for the years 2018-2022 from Section 6 of NIPA Tables, available through the BEA, here. Data on wages and salaries come from Table 6.2, on value added by in are from Table 6.1, the number of self-employed from Table 6.7, hours worked by employees from Table 6.9 and corporate taxes from Table 6.18. In addition, we follow the BLS and use the Current Population Survey (CPS) to impute weekly hours worked by the self-employed in each of these 14 industries.

The 14 NIPA industries are: (1) agriculture, forestry, fishing, and hunting, (2) mining , (3) utilities, (4) construction, (5) manufacturing, (6) wholesale trade, (7) retail trade, (8) transportation and warehousing, (9) information, (10) finance, insurance, real estate, rental, and leasing, (11) professional and business services, (12) educational services, health care, and social assistance, (13) arts, entertainment, recreation, accommodation, and food services, and (14) other services, except government.

A-3 Easy and Hard Tasks

The procedure used for classifying "easy-to-learn" (or "easy") and "hard-to-learn" (or "hard") tasks is described at a high level in the text. Here I provide further details.

We start with Eloundou et al's exposure measure and focus on three features of tasks (i) primary verbs, (ii) IWAs, and (iii) topic clusters coded by LDA.

1. **Primary verbs:** Each task on O*NET has a task description that starts with at least one verb. These "primary" verbs are informative about what the task involves and how easily a task can be automated using AI. In particular, verbs that indicate objective measurement possibilities for the outcome (e.g., *calculate* or *quantify*) or involve predictable rules that allow for easy mapping to outcomes (*estimate*) are more likely to be associated with easy tasks.

   We manually classified verbs as easy, hard or uncertain, where the last category contains verbs that are harder to classify or involve both harder and easier elements (for which the context will be important as I describe next). The full list of verbs and their classifications are provided in the replication package, and the most common hard and easy verbs are listed in Table A-1 in this Appendix.

2. **Intermediate Work Activities:** Each task on O*NET is situated within a hierarchy and is associated with at least one Detailed Work Activity (DWA), which is a coarser description of the task. Each DWA is in turn associated with at least one Intermediate Work Activity (IWA), which is at a higher level aggregation. There are 19264 tasks and 332 IWAs in O*NET. While verbs are informative on whether the observed act is associated with a predictable outcome, IWAs provide context about the observability of the outcome. For example, *explaining* can be hard-to-learn when it is in the context of teaching key concepts to elementary school students, but easy-to-learn when it involves explaining the rules of a parlor game. We manually classified each IWA as easy or hard taking this context into account. The replication package also includes the exact list of IWA classifications.

3. **Topic clusters:** We use a separate unsupervised learning algorithm, the Latent Dirichlet Allocation (LDA) algorithm (Blei et al., 2003), to determine probabilistic topic clusters. Topic clusters provide an additional layer of context for the algorithm. We apply the LDA algorithm to all 19264 tasks to generate 100 topic clusters. LDA lever-

ages word co-occurence within task statements to determine probabilistic assignment of each task to topic clusters.

We use the output of these three steps with our supervised learning model, which is a gradient boosted classifier (GBC) (Friedman, 2001). GBCs maximize goodness-of-fit over a training set by iteratively generating partitions. We produce the training set by manually classifying 757 of the 4089 exposed tasks. The GBC produces the best fit for these 757 tasks, and we obtain the output of the classifier for the remaining tasks. The inputs into the GBC are: the verb classification, the IWA classification, and probabilities to belonging to each one of the 100 topic clusters from the LDA. In addition, we also feed the GBC the empirical frequencies of each verb being associated with an easy or uncertain task (hard task being the omitted category), and each IWA being associated with an easy task in our training data. This information is useful as it provides an empirical measure of the degree of certainty with which each verb and IWA is classified (analogous to the prior in a Bayesian calculation). When a task has more than one primary verb, we take the average across these.

The goodness-of-fit of the GBC is quite good. It matches 90% of the manual classification on our training data and has an AUC (area under the curve) score of 0.99. As output, we generate the probability of each task being easy-to-learn or hard-to-learn. When calculating the adjusted TFP scores, we weight each task by the degree of exposure (according to Eloundou et al) multiplied by the probability scores, to obtain the share of easy-to-learn tasks in the economy. Further details on the GBC and the training sample are contained in the replication package.

Finally, it is important to note that because the GBC algorithm is trained and tested exclusively on exposed tasks, its output is most meaningful for this sample. For example, for verbs that are absent or rare in this sample, it would not provide a meaningful classification. Therefore, its output should not be directly extrapolated to non-exposed tasks. Moreover, because Eloundou et al's sample is already for exposed tasks, we would expect a higher fraction of these to be easy-to-learn than for the entire set of tasks. Indeed, as reported in the text, about three quarters of the wage bill-weighted exposed tasks are classified as easy-to-learn within this sample.

A-4   Wage and Inequality Calculations

To estimate the general equilibrium impact of AI on wage outcomes, I follow the methodology used in Acemoglu and Restrepo (2022).

Hard-to-Learn Verbs

- direct (325)
- supervise (275)
- clean (270)
- select (223)
- participate (197)
- repair (174)
- advise (161)
- train (156)
- instruct (110)
- study (105)
- verify (98)
- consult (97)
- attend (94)
- teach (93)
- oversee (57)
- negotiate (56)
- diagnose (54)
- present (50)
- disassemble (40)
- counsel (39)
- prescribe (36)
- educate (31)
- hire (31)
- visit (28)
- represent (21)
- testify (20)
- patrol (19)

- pour (19)
- scrape (18)
- feed (16)
- care (11)
- accompany (11)
- travel (10)
- mentor (10)
- talk (8)
- nail (8)
- hammer (7)
- hang (7)
- advocate (7)
- feel (7)
- sprinkle (6)
- heat (6)
- mediate (5)
- decorate (5)
- play (5)
- disinfect (5)
- audition (5)
- groom (4)
- furnish (4)
- inject (4)
- cook (4)
- tutor (4)
- wipe (4)

Easy-to-Learn Verbs

- write (219)
- record (191)
- measure (176)
- schedule (81)
- document (78)
- calculate (74)
- drive (67)
- compute (62)
- estimate (62)
- inform (61)
- refer (56)
- process (50)
- respond (43)
- sort (42)
- draw (35)
- update (34)
- notify (33)
- count (22)
- edit (14)
- draft (13)
- revise (13)
- transcribe (11)
- classify (10)
- pay (10)
- send (9)
- solicit (9)
- announce (8)

- translate (7)
- quote (6)
- standardize (6)
- model (6)
- code (6)
- scribe (6)
- proofread (5)
- optimize (5)
- convert (5)
- comply (5)
- adhere (5)
- summarize (5)
- upload (4)
- alert (3)
- detect (3)
- score (3)
- format (2)
- digitize (2)
- diagram (2)
- download (2)
- rewrite (2)
- visualize (1)
- debug (1)
- encode (1)
- reformat (1)
- telephone (1)

Notes: This table displays a subset of verbs which were manually classified as hard-to-learn or easy-to-learn. The frequency with which each verb appears as primary in the task list among all tasks is displayed in parentheses. Verbs are sorted by frequency.

Table A-1: Manual Classification of Verbs as Easy versus Hard-for-AI

We focus on 500 demographic groups defined by age (five age groups), sex, race (white, Black, Asian, Hispanic and other), education (less than high school, with high school degree, some college, Bachelor's degree, and postgraduate degree) and immigration status (native or foreign-born). We use the 2018-2022 five-year American Community Survey data to determine hourly wages for workers and windsorize to deal with outliers as in Acemoglu and Restrepo (2022).

Estimates of the propagation matrix, Θ, are directly from Acemoglu and Restrepo (2022) and are for exposure to robots, software and dedicated machinery between 1980 and 2016. I take their estimates of task displacement based on industry labor share declines, reported in column 1 of Table A-X in the Online Appendix of Acemoglu and Restrepo (2022). The propagation matrix calculates the total spillovers between industry, occupation and age-education groups for observed changes between 1980-2016. The only difference from Acemoglu and Restrepo (2022) is that we feed in exposure to AI and its TFP implications, rather than those related to automation between 1980-2016 due to other technologies, as in their paper.

The quantitative implications are then produced using a fixed-point iteration algorithm to solve the following set of equations, as in Proposition 4 of Acemoglu and Restrepo (2022):

$$d \ln w_g = \Theta_g \cdot \left(\frac{1}{\sigma} d \ln y + \frac{1}{\sigma} d \ln \zeta - \frac{1}{\sigma} d \ln \Gamma^{\text{AI}} \right)$$

$$d \ln \zeta_g = \sum_{i \in I} \omega_{gi} \cdot \left(\frac{\partial \ln s_i^Y(p)}{\partial \ln p} \cdot \ln p + (\sigma - 1) \cdot d \ln p_i \right), \forall g$$

$$d \ln p_i = \sum_{g \in G} s_{gi}^L \cdot (d \ln w_g - d \ln \Gamma_{gi}^{\text{AI}} \cdot \pi_{gi}), \forall i$$

$$d \ln y = \frac{1}{1 - s^K} \cdot \left(d \ln \text{tfp} + s^K \cdot d \ln s^K \right)$$

$$d \ln s^K = -\frac{1}{s^K} \sum_{g \in G} s_g^L \cdot (d \ln w_g - d \ln y).$$

Here w_g is the wage of group g, Θ_g refers to the row in the propagation matrix associated with group g, y is economy-wide GDP, ζ is the vector of industry shares, Γ_g is AI exposure for g, ω_{gi} is the share of wages earned by group g in industry i relative to this group's total wage income, s_i^Y is the income share of industry i, p_i is the price index for industry i, s_{gi}^L is the share of wages earned by members of group g as a fraction of total earnings in industry i, π_{gi} is cost savings in tasks performed by group g in industry i, s^K is the capital share of income. Finally, Acemoglu and Restrepo (2022) provide an equation for $d \ln tfp$ as a

function of cost savings and industry shares, but rather than using this equation explicitly, I take the change in TFP from my previous computations. All variables (except y) when denoted without a subscript refer to $G \times 1$ vectors.

The variables s_i^Y, s_{gi}^L and ω_{gi} are computed directly from the American Community Survey data, supplemented with industry-level labor share data from the BEA at the level of 49 industries as in Acemoglu and Restrepo (2022). Γ^{AI} is derived from our aforementioned occupation-level exposure numbers. Specifically, we aggregate from the occupation level to the industry-group level by first matching occupations with occupation-level AI exposure, then aggregating this to the industry $\times$ demographic group level. For the baseline exposure results, we set $\pi_{gi} = \bar{\pi} = 0.27$, which is the average of Noy and Zhang (2023) and Brynjolfsson et al. (2023), as explained in the text. Finally, $\frac{\partial \ln s_i^Y(p)}{\partial \ln p}$ is computed as in Acemoglu and Restrepo (2022) by assuming that sectoral outputs are combined in GDP with the constant elasticity of substitution of $\eta = 0.2$ (as explained in the text).

For easy versus hard tasks, I repeat the same procedure, except that I use separate exposure measures for exposure to AI in easy tasks and exposure to AI in hard task for each demographic groups.