

von Maydell, Richard

Conference Paper

Artificial Intelligence and its Effect on Competition and Factor Income Shares

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2023: Growth and the "sociale Frage"

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: von Maydell, Richard (2024) : Artificial Intelligence and its Effect on Competition and Factor Income Shares, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2023: Growth and the "sociale Frage", ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/290195>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Artificial Intelligence and its Effect on Competition and Factor Income Shares*

Richard von Maydell
CER-ETH – Center of Economic
Research at ETH Zurich
Leonhardstrasse 21
8092 Zurich, Switzerland
rbaron@ethz.ch

April 25, 2024

Abstract

We examine the impact of self-learning Artificial Intelligence (AI) on firm competition in a growth model with endogenous labor supply and heterogeneous agents. AI possesses the ability to improve autonomously through application, testing, and training. When firms incorporate AI into their production processes, they incur variable costs for acquiring the necessary software as well as fixed costs for installing AI infrastructure. The latter indeed drive productivity increases and economic growth but can also serve as an entry barrier for competing firms. Therefore, we examine how the rise of AI affects market concentration, firm competition, firm productivity and income inequality. We discuss potential policy interventions such as a profit tax or the modernization of competition and merger laws to prevent a significant increase in market concentration and income inequality within AI-intensive industries and to foster economic growth.

Keywords: Artificial Intelligence, Imperfect Competition, Markups,
Competition Policy, Factor Income Shares

JEL Codes: L11, L13, L52, O33, D21

*We thank seminar participants of the IEA World Congress 2023, VfS Annual Conference 2023, 6th Economics of Digitization Workshop, 16th RGS Doctoral Workshop, 11th Oxford Workshop on Global Priorities Research, Conference on Robots and Automation 2022, NIESR Workshop on Data Use and Productivity in the UK 2022, INFER International Workshop on Economic Growth and Macroeconomic Dynamics 2022, Rare Voices in Economics Conference 2022, ETH Astute Modelling Seminar 2021, Hans Gersbach, Margrit Buser, David Hémous, Evgenij Komarov, Klaus Prettnner, John Sturm, Ata Atabek and Arthur Schichl for valuable comments.

1 Introduction

The ascent of the "Tech-Giants"—e.g., Amazon, Alphabet, Meta or Apple—has led to a significant increase in market power in the industries for software, social media, or communication networks enclosing new policy challenges (Stigler Committee, 2019). There are more and more calls for increased market regulation of firms in the digital economy from the political and economic side. An essential cornerstone for the success of the Tech Giants has been the rise of Information and Communication Technologies (ICT)—and nowadays, the steadily increasing relevance of Artificial Intelligence (AI). We define AI as an intangible asset, as it is an "investment in organizational capabilities, creating or strengthening product platforms that position a firm to compete in certain markets" (Hulten, 2010). Other examples for intangible assets are e.g., software, databases, R&D, design, training, market research, branding or business process engineering (Haskel and Westlake, 2017). AI is distinct from "brute-force" (Makridakis, 2017) machines and has the crucial characteristic of being able to learn from application: AI is increasingly capable of autonomously improving without human aid and can improve itself by e.g., deep machine learning or reinforcement learning (Lu, 2020). As information can be broadly and cheaply assessed nowadays and evaluated using machine-learning algorithms, AI can autonomously develop over time (Brynjolfsson et al., 2017), which we call *self-learning* (Gersbach et al., 2022).

We suppose that a rise in AI algorithms will affect market concentration—not only in the software industry, where it facilitated the ascent of the Tech Giants, but also in industrial production. In the spirit of De Ridder (2019), we assume that AI may decrease variable costs but comes in hand with investments in AI infrastructure for its implementation in industrial production which we interpret as fixed costs. As a result, the main questions we address using our theoretical model are the following: When and why do firms prefer to implement AI, even if additional fixed and variable costs have to be borne? How do different groups benefit from AI implementation in industrial production? How does AI implementation affect market concentration, input factor allocation, markups and factor income shares? What are governmental interventions that guarantee an optimal economic integration of AI to broadly distribute the benefits of AI to all population groups?

Our main motivation is to theoretically examine how the industrial incorporation of AI—a self-learning intangible asset—may affect market concentration and factor income shares. The new feature of our model is that we endogenize the optimal AI infrastructure investments and analyze the resulting number of firms operating in industrial production in dependence of the development of AI and the preferences of different stakeholding groups. Therefore, the rise in market concentration is an endogenous result of our model. We focus on a neoclassical economy with heterogeneous agents and elastic labor supply and industrial firms that produce a single consumption good using a nested CES production

function that incorporates AI as a non-rival input factor. In our model, firms with a high AI productivity have incentives to invest in AI infrastructure—not only due to its positive effect on firm-specific productivity, but also to oust competing firms from the market, to obtain higher market shares and to be able to charge price markups. Depending on the productivity differences between firms in industrial production, we observe different levels of market concentration. In our approach, we focus on modelling the optimally-chosen AI infrastructure investments—promoting firms’ AI productivity but inhibiting market entry of competing firms—that maximize the income of different groups. Since we assume that ordinary workers benefit from their capital and labor income, whereas entrepreneurs additionally benefit from the sale of AI algorithms and firm profits, we examine how the level of AI affect the income of these agents. Our model shows that the evolution of the factor income shares and the (endogenous) labor supply depends on the elasticity of substitution between the input factors aggravating a conclusive statement on the effects of AI on the labor market. Yet, we generally observe an increase in market concentration, imperfect competition and rising markups if firm-owners can collectively decide on the optimal AI investments, leading to allocative inefficiency and socially sub-optimal output. Moreover, we observe a strong divergence of capital and consumption rates of entrepreneurs and workers in our economy with AI, especially in case of imperfect firm competition leading to decreasing capital and labor shares and an increasing profit share. This trajectory particularly supports wealth inequality promoted by AI, based on which we discuss potential policy interventions to counteract income divergence.

The remainder of the paper is organized as follows: In Section 2, we present related literature. We describe the essentials of our theoretical model in Section 3. Afterwards, the effects via which AI affects price markups and firm competition in industrial production, namely via investments in AI infrastructure and resulting firm exit and productivity effects, are described in Section 4. We define equilibrium conditions in Section 5 and describe the optimization considerations of different stakeholding groups and show how this affects AI growth and the long-run equilibrium in Section 6. For illustrative purposes, we provide a numerical exercise of the model in Section 7, assess comparative statics of the model and outline the effect of AI on several parameters of interest. We delineate inefficiencies in our economy and counteracting policy interventions in Section 8 and discuss our findings and portray the weaknesses of our model in Section 9. We conclude in Section 10. Additional graphs, tables and equations are relegated to the Appendix.

2 Relation to the Literature

There is a myriad of literature discussing the implications of automation and robots for economic outcomes, e.g., by Acemoglu and Restrepo (2018a), Acemoglu and Restrepo (2018b) and Hémous and Olsen (2014). Particular attention has been devoted to the

effect of automation on the labor force for assessing how new technologies may affect for example unemployment rates or the labor income share (Prettner, 2019; Korinek and Stiglitz, 2017; Grossman and Oberfield, 2021). Karabarbounis and Neiman (2014) describe that the effect of a new technology on the labor force strongly depends on the industry under consideration, as especially repetitive tasks that are easily routinizable such as clerical or simple assembly occupations may be at risk with a rise of automation. Trammell and Korinek (2020) state that automation allows capital to perform more tasks that were formerly performed by human labor and Aghion et al. (2017) argue that automation is labor-augmenting as it allows capital "to better complement to labor". On the opposite, Hémous and Olsen (2014) state that automation displaces especially low-skilled labor in the long-run and leads to a decline in the labor income share providing evidence for the connection between automation and the rise in income inequality which gets underlined by Acemoglu and Restrepo (2021).¹

Increasingly, however, the topic of data and AI is also being considered separately in economic literature (Aghion et al., 2017; Jones and Tonetti, 2020). In contrast to the definition of e.g., robots in economic literature, AI is not simply a technology that can be used as a substitute for labor (Trammell and Korinek, 2020) but rather a technology that improves endogenously via its application. Inspired by Hanson (2001) who revises learning by doing in light of new technological developments and the concept of machine intelligence, we focus on the self-learning feature of contemporary AI algorithms as defined by Gersbach et al. (2022). Moreover, AI is interpreted as a scalable intangible asset (Haskel and Westlake, 2017), as AI software can be used simultaneously by different firms and, in contrast to many tangible assets, AI can be re-employed repeatedly at a relatively little cost.

Being interpreted as a "general-purpose technology" (Eeckhout, 2021; Cockburn et al., 2019), AI has already a large impact not only on the Tech-industry, but also on industrial production. In the automotive industry, for example, software development is becoming part of the core business and AI plays a decisive role e.g., to support the endeavor for autonomous driving (Falcini et al., 2017). Patrício and Rieder (2018) reveal in a meta-analysis that the agricultural and food industry could benefit from AI, which could enhance production, quality control or food security as it e.g., facilitates the detection of plant diseases, grain classification or the determination of flowering stages. Cioffi et al. (2020) state that AI may be applied in supply chain management for demand forecasts, maintenance for predictive forecasts or improve the effective use and logistics of non-renewable resources. Ernst et al. (2019) refer to the possible applications of AI, in particular in the area of algorithmic trading, patients' pathology diagnosis or the automated review of legal contracts.

¹For more insights into the theoretical and empirical implications of task replacement and automation on human labor we reference to Hémous and Olsen (2014); Acemoglu and Restrepo (2018a,b, 2020).

Eeckhout (2021) states that although technological progress may displace workers, it benefits everyone due to higher real income for everyone—yet, only if markets are competitive. He determines different channels through which technology may foster a rise in market concentration, namely through the supply side (e.g., ownership of railway tracks) or through the demand side (e.g., network effects on social media platforms). In particular, he mentions a third factor affecting market power, namely learning, which becomes more important in times of machine-learning algorithms and a vast amount of data and cheap information. Goldin et al. (2020) analyze the link between the intensity of intangible assets—which also comprise AI—in a given industry and a declining degree of competition. De Ridder (2019) explains that the increasing use of intangible inputs leads to an increasing importance of fixed costs compared to variable costs. He assumes that expenditures on intangible inputs can reduce firms’ marginal production costs and highlights that the implementation of intangible assets change the production process of a firm. Despite large initial investments that have to be made for the infrastructure, development or maintenance of intangibles—defined as fixed or overhead costs—, variable costs decrease with scale, which makes the marginal costs converge to zero, indicating the increased importance of fixed costs compared to variable costs. Grossman and Oberfield (2021) note that the implementation of new technologies in the production process entails sizable fixed costs. In line with this argument, De Loecker et al. (2021) assume that firms have to pay fixed costs to enter an industry and to produce with a new technology or to pay nothing and to not enter the market. Yet, Haskel and Westlake (2017) state that intangible-intensive economies need new types of physical infrastructure and refer to the common theme in public debate that high fixed costs can inhibit economic growth. Schweitzer et al. (2022) point out that the training of AI algorithms after investing in software development and infrastructure is at near zero marginal costs. A well-known contribution to the industrial organization literature on the effect of fixed costs and market barriers on market competition is provided by Dixit (1979). He models a duopolistic economy, where an established firm faces potential competition from a prospective entrant and assesses the effect of fixed costs on discontinuities on firms’ reaction functions. Moreover, he analyzes how the fixed costs of an established firm have to be set to block the entry of a competing firm. Maskin and Tirole (1988) show that fixed costs can be that large that only one firm can generate positive profits on the market. Similarly, Osborne (1964) states that depending on the fixed costs, the entry of competitors can be ineffectively impeded, effectively impeded or, in an extreme case, even blockaded. The concentration of production in a few firms does not necessarily hint at weakened competition, but may also reveal that the most innovative and productive firms have an increased market share (Autor et al., 2020a; Bajgar et al., 2021). Nonetheless, Calvano et al. (2020) point out that the degree of possible collusion on the market declines with a rising number of competitors. Haskel and Westlake (2017) argue that due to the scala-

bility of intangibles, technologically-leading firms can outperform technological laggards, which facilitates the emergence of large and profitable firms which they empirically motivate by showing that the profitability of US domestic corporate businesses has recently become increasingly unequal. Bajgar et al. (2021) provide empirical evidence that market concentration has been increasing, particularly in markets with high levels of intangible assets, rising profits and markups and declining business dynamism, primarily in globalized and digital-intensive industries. Calligaris et al. (2018) note that the decline in business dynamism has been especially pronounced in information technology, telecommunication and services but highlight that technologically-laggard firms benefit less from the digital transformation as catching up to the technological frontier has become more difficult. In this spirit, Autor et al. (2020a) observe that large firms become more productive with more ICT and that intangible-intensive firms show higher markups and lower labor shares. Diez et al. (2021) provide empirical evidence that market concentration has especially risen in the digital economy, in particular in the ICT industry. Akcigit et al. (2021) state that "market concentration has risen, firms' price markups over costs have increased [...] and profitability has doubled" in the last decades. Ernst et al. (2019) point out that the digital nature of AI makes large "first-mover advantages" possible—leading to increased market concentration and economic inequality. Babina et al. (2021) provide empirical evidence that AI contributes to a rise in industry concentration as it especially benefits large firms. Li et al. (2017) emphasize that the AI-induced changes to competitiveness and to social and economic benefits need to be more thoroughly evaluated. On the opposite, Diez et al. (2021) conclude that technological changes coming in hand with rising fixed costs do not play a major role in explaining the rise in markups, but caution against the observation that a rise in market concentration and decline in competition may lead to imperfect factor allocation, welfare losses or inequality.

In addition to our analysis of an AI-induced rise in fixed costs and markups, we examine the effects of AI on factor income shares. Autor et al. (2020a) provide empirical evidence for a declining labor income share (LIS) due to a so-called rise of "superstar firms". One of the main findings of De Loecker et al. (2020) is a negative relation between the LIS and markups. Diez et al. (2019) state that rising markups have an inverse effect on firms' LIS and that national income paid to workers has been decreasing since the 1980s. Empirical studies reveal a decline in the LIS, both from a global perspective (Elsby et al., 2013) or specifically for the US, where a decline of around 7% in the labor share can be observed since 2000 (Karabarbounis and Neiman, 2014). With regard to industry-specific differences, Barkai (2020) points out that larger declines of the LIS are especially observable in industries with strong market concentration. There is a variety of reasons that are discussed in the literature for explaining the decline of the labor income share. Whereas Karabarbounis and Neiman (2014) hint at the decline in the relative price of investment goods, the reasoning of Raurich et al. (2012) is based on a larger

capital deepening and price markups. Grossman et al. (2017) explain the decreasing LIS with the human capital complementarity with physical capital. Trammell and Korinek (2020) conclude that the reason for a decline in the LIS is that wages rise, but less than output. In contrast to the literature showing that a decreasing labor share comes hand in hand with an increasing capital share, Barkai (2020) reveals that—independent of the industry—both labor and capital shares have been declining, which has been re-balanced by increasing profit shares.

3 Model

For the sake of simplicity, we choose a neoclassical economy where agents’ savings and consumption are endogenized and labor supply is elastic. We model an economy with an industrial sector in which firms produce a consumable good and an AI sector, where companies produce AI algorithms. Industrial firms with heterogeneous productivities use capital, labor and AI for production. In the model, AI algorithms can autonomously develop by being applied, tested and trained in industrial production. We focus on describing the supply side in detail. We illustrate the main features of our model in Figure 1 before subsequently explaining each feature in more detail.

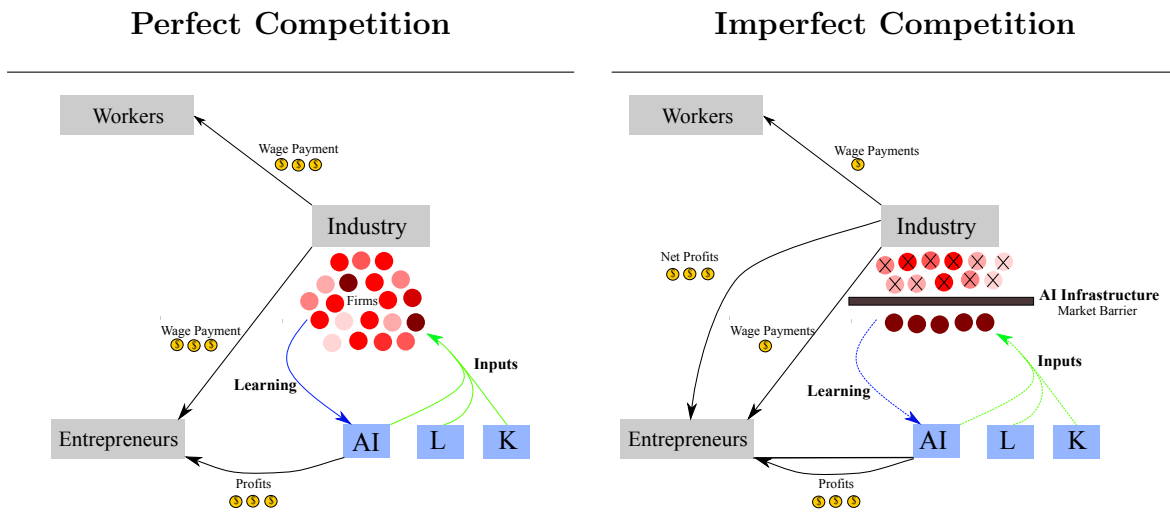


Figure 1: Diagram of the Model Concept.

In an initial state (left part of Figure 1), industrial firms with heterogeneous productivities (illustrated by dots with different intensities of red color) operate on a perfectly competitive market, entrepreneurs and ordinary workers obtain a competitive wage and industrial firms make zero profits. Yet, in particular, productive industrial firms have an incentive to build up AI infrastructure (e.g., the training of workers, installation of servers, acquisition of data) serving as market barriers that impedes market entry of industrial firms with a lower productivity, leading to imperfect competition (right part of

Figure 1). In such a setting, we show that markets are more concentrated and industrial firms can charge higher markups, leading to lower wages and positive profit payments for entrepreneurs.

3.1 Individuals

There are atomistically small agents of mass $L \in \mathbb{N}_+$ in the economy. The mass L of agents in the economy stays constant over time. In our growth model, we assume a multi period but finite time horizon, where $t \in \{1, \dots, T\}$ and $T \in \mathbb{N}_{\geq 1}$ is a discrete time index. All agents live for the entire time period. The mass L of workers in the economy is endowed with L units of labor that can be used for labor or leisure. We describe the labor leisure trade-off in more detail at a later stage. Moreover, at $t = 0$, a total capital endowment of K_0 is uniformly distributed amongst all agents in the economy. Capital can be accumulated over time and total capital in each period is given by $K_t \in \mathbb{R}_+$. Individuals inelastically rent out their entire capital supply to industrial production and obtain the interest rate r_t . Individuals differ with regard to their skill level η , with $\eta \in \{W, E\}$. First, agents with index $\eta = \{W\}$ are defined as ordinary workers, who work in industrial production. Second, agents with skill index $\eta = \{E\}$ have the same productivity as agents with $\eta = \{W\}$ but have entrepreneurial talent such that they own all firms in industrial production in all periods. In addition, agents with entrepreneurial skills own companies that develop AI algorithms.² Henceforth, we call agents with entrepreneurial skills entrepreneurs as they own all industrial firms and all AI-developing companies in the economy. Yet, all agents irrespective of their skill-level work in industrial production.³ Total labor supply of agents in the economy with endogenous labor supply is given by $L_t = L_t^W + L_t^E$, where L_t^E is the labor supply of agents with entrepreneurial skills and L_t^W is the labor supply of ordinary workers. We define the share of agents with entrepreneurial-skills as $l^E = \frac{L^E}{L} \in (0, 1)$ and of ordinary workers as $l^W = 1 - \frac{L^E}{L}$. The share of entrepreneurs and workers stays constant over time, but the time spent on leisure might change over time due to the labor leisure trade-off, such that we define the amount of supplied labor by group η at time t as L_t^η . The group of agents collectively decide on the time spent on leisure, given by $N_t^\eta = Ll^\eta - L_t^\eta$. Moreover, there are AI algorithms that can be used in industrial production. The concept of AI algorithms will be later described in more detail. A more detailed definition of the income stream of all agents in the economy is provided in Section 5.

²To avoid misunderstandings we use the term "firm" for corporations in industrial production and "company" for corporations in the AI-developing sector.

³For the sake of simplicity and in contrast to Lankisch et al. (2019) who assume different elasticities between automation and heterogeneously-skilled workers, all agents have the same elasticity of substitution with AI in our approach.

3.2 Industrial Production

There are $j \in \{1, \dots, N\}$ firms with $N \in \mathbb{N}_{\geq 1}$ in industrial production. All industrial firms produce the same consumption good. The nested CES production function of all j firms at time t in industrial production has the following form:

$$Y_{j,t} = \left(\alpha K_{j,t}^{\frac{\epsilon-1}{\epsilon}} + (1-\alpha) \left[\left(\gamma L_{j,t}^{\frac{\omega-1}{\omega}} + (1-\gamma)(\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right)^{\frac{\omega}{\omega-1}} \right]^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}. \quad (1)$$

The production function of each firm j is concave, strictly increasing and differentiable with respect to all input factors (Mas-Colell et al., 1995). Each firm j uses capital $K_{j,t} \in K_t$, labor $L_{j,t} \in L_t$ and AI $A_t \in \mathbb{R}_+$ as input factors to produce the consumption good. The parameters $\epsilon \in (0, \infty)$ and $\omega \in (0, \infty)$ define the elasticities of substitution between labor, capital and AI, and $\alpha \in (0, 1)$ and $\gamma \in (0, 1)$ determine the respective factor income shares. Moreover, we assume that labor L_t is not transferable across time, whereas capital K_t and AI A_t accumulate over time. Industrial firms can perform intra-period trade of labor and capital and can buy AI algorithms from AI-developing companies.

Depending on the state of the economy, only $m_t \in \{1, \dots, N\}$ firms are active in industrial production. We consider two types of firms: Active firms m_t have a strictly positive demand for the input factors for production, whereas $N - m_t$ passive firms do not produce and do not demand any input factors. Only a limited amount of the input factors labor and capital can be employed in production by each firm depending on the number of active firms $m_t \in \{1, \dots, N\}$ at time t .⁴ As labor and capital are rival and excludable goods, all m_t firms that are active at time t compete for the available input factors. We denote the labor and capital supply of firm j at time t depending on the number of active firms m as $L_{j,t,m}$ and $K_{j,t,m}$, respectively. Therefore, also the production $Y_{j,t,m}$ of a firm j at time t depends on the number of m active firms. We suppose that all industrial firms can acquire AI algorithms in a non-rival, but excludable fashion (Eeckhout, 2021; Ernst et al., 2019; Wagner, 2020): every firm that buys AI software A_t can use it. Recall that non-rivalry of AI implies that if one firm buys AI algorithms, it does not prevent other firms' use of AI (Acemoglu, 2009; Farboodi and Veldkamp, 2021). Yet, it is an excludable good as only firms that buy AI software can use it for production.

A main assumption in our model is that all industrial firms have a heterogeneous AI productivity. Our approach stands in line with Aghion et al. (2019), who assume that there is heterogeneity in firm-specific efficiency that persists over time—in our case, in firm-specific AI productivity, given by $\theta_{j,t}$. Thus, each firm j has a time-variant productivity $\theta_{j,t}$ drawn from a fixed probability distribution $\Phi(\theta_{j,t})$. We assume that the firm with rank $j = 1$ draws the highest productivity and the N -th firm draws the lowest from

⁴In the remainder of the paper we disregard the indices t if we use m_t in equations and only use m .

the distribution.⁵

Assumption 1

Each firm j has a firm-specific AI productivity $\theta_{j,t}$ drawn from a fixed productivity distribution $\Phi(\theta_{j,t})$. The higher the rank j of a firm, the higher its productivity $\theta_{j,t}$.

We assume that the productivity of each firm is discernible for all other firms. In this way, every firm knows its own AI productivity and that of all other firms. For the sake of simplicity, we suppose that the labor and capital productivity of all firms are identical. Productivity changes over time are disregarded.⁶ Thus, our model does not allow for i.e. firm-specific research and development or innovation, which could improve firms' AI productivity over time.

3.3 AI Algorithms

Already in 1965, Moore observed that the number of transistors in a densely integrated circuit doubles about every two years (Moore, 1998). This technological development gave rise to Information and Communication Technology (ICT) that has shaped the path of economic development in the last decades. Similarly, we assume that AI will grow exponentially in the upcoming years due to its self-learning characteristics and the immense rise in data availability.

Typical applications of AI (currently) are speech recognition, image recognition, natural language processing or shortest path derivation.

In our model, we conceptualize that the greater industrial production, based on the production of each active firm given by Eq. (1), the greater the application area for the testing and training of AI algorithms. Although a clear distinction between generic data, data, information and communication technologies and algorithms is often hard to make, we focus on the term Artificial Intelligence as we focus on the self-learning capabilities of AI algorithms. For example, in the agricultural industry, the greater total production of wheat grain, the better AI algorithms can be trained to detect grain diseases. In an example with autonomously-driving cars, the more cars are produced and sold, the more observations can be assessed and the better the learning of an AI-supported autonomous driving software. Therefore, we interpret total industrial production as the "training set" for AI algorithms and suppose that AI is an input factor that autonomously grows through its application in industrial production. In this sense, as stated by Trammell and Korinek (2020), "the production process itself contributes to the generation of productivity-increasing ideas". Thus, our model is in the fashion of Romer (1986, 1990),

⁵The initial productivities θ_{j,t_0} of all firms are equal at $t = 0$ and the subsequent dynamics of the productivity parameter are discussed at a later stage.

⁶Nonetheless, it could be an avenue for future research to examine how a stochastic evolution of the firm-specific AI productivity would affect the economy in more detail. For instance, De Loecker et al. (2021) model the productivities of firms entering the market as AR(1) processes.

stating that knowledge is a result of *learning by doing* (Arrow, 1962) in production and builds on a comparable assumption as Farboodi and Veldkamp (2021) stating that "data is a by-product of economic activity".

Motivated by Gersbach et al. (2022), we simplify our model and assume that AI has reached a stage at which it can develop autonomously due to its self-learning characteristics. This autonomous development implies that no labor or capital is required for the accumulation of AI—nonetheless, this extreme assumption has to be interpreted with caution and is only imposed for the sake of simplicity. Due to its autonomous development over time, AI is distinct from automation that serves as a substitute for labor or a technology that affects total factor productivity but does not learn by application. Inspired by an energy system analysis of Höök et al. (2011), we assume bounded growth, such that AI may grow interminably, but the rate of growth converges to zero in an infinite time horizon. In the spirit of their arguments, we suppose that the upper bound of AI may be virtually non-existent, but the steps in self-learning of AI software become increasingly smaller, thus slowing its growth process. For the sake of tractability and viability of our model, we assume that there is an upper bound B for the level of AI, caused by restricted hardware availability necessary for the implementation of AI in industrial production.⁷ We model the level of AI as follows:

$$A_{t+1} = \min\left[B, A_t \left(1 + b \left(\sum_{j=1}^m Y_{j,t,m}\right)\right)\right], \quad (2)$$

where $A_0 > 0$. We assume that $b(0) = 0$ and b is strictly positive for $x > 0$.

There is a unique company in the AI sector that develops AI algorithms. The AI company produces self-learning algorithms that autonomously accumulate over time without using any labor and capital as described in Eq. (2). Entrepreneurs collectively own the AI-developing company. The AI company sells AI algorithms to firms in industrial production. Laatikainen and Ojala (2014) argue that a "flexible and well-designed architecture" makes different pricing models for "software-as-a-service" possible, leading to firm-specific prices for the AI software. Therefore, we assume that entrepreneurs owning the AI company can apply competitive differentiation and can perfectly price-discriminate industrial firms that buy AI software, such that each industrial firm has to pay its marginal value for acquiring AI algorithms. All active firms buy the entire non-rival and excludable AI stock if they decide to produce using AI.⁸ It is a stark as-

⁷Even if we can assume that the hardware evolves together with the software, we suppose on the basis of the findings of Kumar (2015) that there are still physical scaling limits that restrict the growth of AI software. Moreover, also Farboodi and Veldkamp (2021) assume an upper bound of data productivity in their analysis of a data economy. Thus, AI cannot sustain growth in the long run, as it reaches an upper bound due to hardware restrictions.

⁸Entrepreneurs collectively decide on establishing this pricing mechanism. Otherwise, each entrepreneur would set a zero price for AI for the industrial firm s/he owns.

sumption that firms pay different prices for acquiring AI software. Yet, it can also be interpreted such that all industrial firms pay the same base price for the acquisition of AI software, but industrial firms with a higher AI productivity also pay for additional features, functions or services and thus face different variable AI costs. Another interpretation is that the rise in AI enables software-developing companies to better know customers' and firms' preferences and behavior such that the estimation of their price and income elasticities can be optimized allowing software producers to improve their pricing strategy and enabling them to charge personalized prices. The profit of the AI company is given by⁹

$$\Pi_{t,m}^{AI} = \sum_{j=1}^m p_{j,t,m} A_t - \Lambda(A_t)$$

We note that the profit of the AI company depends on the number of m_t active firms in industrial production that acquire AI algorithms for a firm-specific price. Moreover, the costs for the development of AI are defined as $\Lambda(A_t)$, but will not be discussed in more detail. We disregard the behavior and decisions of the AI-developing company, as we focus on the industrial firms that use AI for production. For further insights on how to model the behavior of AI-developing companies, we refer to Gersbach et al. (2022).

4 AI Incorporation and its Effect on Competition

In addition to being a non-rival input factor for production, we suppose that there are channels through which AI additionally affects firm competition in industrial production. Investments in AI infrastructure increase AI productivity but might serve as market barriers which affect firms' market entry and thus affect the number of active firms m_t .¹⁰ Furthermore, due to the self-learning feature of AI, its development is supported by a broad application area—which is defined as total industrial production, as given in Eq. 2.

4.1 AI Infrastructure Investments

Whether AI gets acquired by an industrial firm depends on the expected profit of producing with AI versus the profit without using AI. Referring to the corporate integration of AI, the European Commission (2021) points at the importance of (computational) infrastructure, in addition to software and an appropriate governance and coordination

⁹It needs to hold that the AI company does not pay any costs for AI development and gets endowed with an AI stock A_0 at $t = 0$ free of charge.

¹⁰Yet, in contrast to Hopenhayn (1992), we do not derive a stationary equilibrium of entry and exit rates of firms that are affected by productivity shocks.

framework. Recall that if industrial firms decide to employ AI in production, they have to buy AI software for the firm-specific price $p_{j,t,m}$ for each algorithm A_t . Moreover, we assume that industrial firms in our model need to invest in infrastructure for AI (fixed costs) in addition to the acquisition of AI software (variable costs). Infrastructure investments in AI can be e.g., the training of workers how to deal with AI, the installation of computational infrastructure such as computers or servers or the acquisition of data. Motivated by Noy and Zhang (2023) who experimentally show that output quality rises by 18% if workers have the possibility to use Chat-GPT—a new AI technology—we assume that AI increases firm productivity. We assume that investments in AI infrastructure increase the productivity of a firm, but they may serve as a market barrier for competing firms. Motivated by Markiewicz and Silvestrini (2021), we interpret investments in AI infrastructure as fixed costs that serve as barriers-to-entry for competing firms. In particular, exclusive ownership of big data (Rubinfeld and Gal, 2017) and data control (Stigler Committee, 2019) in digital ecosystems with AI-based services may constitute a barrier to entry. We claim that if a firm invests in AI infrastructure, all firms have to invest the same amount to be competitive, which will later be explained in more detail. In this way, investments in AI infrastructure can hinder market competitors from entering the market and we focus on modelling the optimally-chosen fixed costs inhibiting the market entry of competing firms that maximize the utility of different groups of agents.¹¹

In each period t , we assume that every firm j has to undertake a minimum infrastructure investment denoted as D_t to be able to incorporate AI in industrial production and to produce with Eq. (1).¹² Nonetheless, firms can make the decision to invest more than the minimum required level, denoted as $F_{j,t} = \overline{F}_{j,t} + D_t$, where $\overline{F}_{j,t}$ represents the amount of investment that goes beyond the minimum AI investment D_t . Each firm's productivity depends on its productivity in the preceding period. If a firm invests $F_{j,t-1}$ in AI infrastructure, this increases its AI productivity in the following period t . Moreover, we assume that the firm-specific learning of AI depends on the market share of a specific firm j , given by $\phi_{j,t-1}$. The larger the market share of a specific firm, the better tailored the AI algorithms for firm-specific production. In this way, the AI productivity of a firm depends both on the firm-specific investments in AI infrastructure, and on the firm-specific market share, which defines how well-tailored the AI algorithms are to a specific firm.¹³ Due to the abovementioned reasons, we define the following function to describe

¹¹Whereas e.g., Hopenhayn (1992) assumes a competitive market structure, Aghion et al. (2019) assume Bertrand competition and Markiewicz and Silvestrini (2021) assume Cournot competition for analyzing firm dynamics, we suppose that firms have market power and compete in quantities, but that the number of firms is variable and we endogenously determine the number of active firms.

¹²For the sake of simplicity, we set $D_t = 0$.

¹³There is no purposefully-directed innovation in our model, but productivity increases due to investments in AI infrastructure and due to the learning capabilities of AI software. The reason is that firms are myopic and do not anticipate the effect of infrastructure investments on the productivity in the following period.

the development of firm-specific AI productivity

$$\theta_{j,t} = \theta_{j,t-1} (F_{j,t-1})^\iota (1 + \phi_{j,t-1})^\eta \quad \text{with} \quad \eta < 1, \quad \iota < 1. \quad (3)$$

Both a higher market share and private investments have decreasing returns to scale, due to $\iota < 1$ and $\eta < 1$. Yet, we assume that $\iota > \eta$ such that the effect of investments in AI have a larger effect on firm-specific productivity than the market share. To sum up, the special feature of AI that we aim at conceptualizing in our model is that its productivity can be promoted by private investments, given by $F_{j,t-1}$, but also benefits from a higher market share, given by $\phi_{j,t-1}$, as algorithms are thus better personalized to firm-specific needs. Therefore, in addition to the self-learning of AI that is described in Eq. (2), we have a notion of firm-specific learning, described in Eq. (3).

4.2 Price Markups

If firms operate on a competitive factor and product market, this does not enable them to charge a markup on the price for the goods sold. Yet, as presented by Raurich et al. (2012) and Autor et al. (2020a), a firm operating on an imperfectly competitive product market can sell products at a higher effective price and charge a price markup. The literature claims that the markup trajectories have changed due to the rise of ICT and we aim at assessing the potential effect of AI on markups. Diez et al. (2019) state that the rise in market concentration after the turn of the millennium can be assigned to (especially digital) technology-driven changes in product markets and underpin the hypothesis of an increase in average markups in their empirical analysis. Markiewicz and Silvestrini (2021) show that markups are higher in high-ICT industries than in low-ICT industries using U.S. CompStat and European CompNet data. Calligaris et al. (2018) additionally state that markup differentials have risen between digitally high-intensive and low-intensive industries. Inspired by the abovementioned literature assessing the link between markups and intangible intensity, and based on the markup definition of De Loecker et al. (2020), we set up a function $\mu_{t,m}$ for price markups depending on the elasticity of demand and the number of active firms that determines total output. If less firms are active, firms can charge markups due to reduced competition, allowing them to finance their infrastructure investments. We define the consumers' demand function in the following way, where $q_t [Y_t(\nu)]$ is total consumers' demand at time t depending on price ν for the industrial good Y_t .

$$q_t [Y_t(\nu)] = \sum_{j=1}^m q_{t,j} = \sum_{j=1}^N Y_{j,t} \nu^{-\tau}$$

The elasticity of demand is given by $\tau \in (0, \infty)$, such that the lower the value of τ , the more elastic consumers demand to changes in prices of the consumption good. If firms have the possibility to charge a price markup $\mu_{t,j}$, the consumers' demand looks the following, where $\hat{\nu}$ is the price if firms operate under perfect competition such that $\nu = \hat{\nu} + \mu_{t,j}$ is the effective price for the consumption good. We can thus rewrite total consumers' demand depending on the effective price as

$$\sum_{j=1}^m q_{t,j} = \sum_{j=1}^m Y_{j,t} (\hat{\nu} + \mu_{t,j})^{-\tau}. \quad (4)$$

In this fashion, we additionally observe that firms' possibility to invest in AI infrastructure is hampered if the elasticity of demand is too high, as this enables firms to charge lower markups. Note that by construction, a situation in which $m_t < N$ firms are active implies imperfect competition.¹⁴ For this reason, firms have strategic incentives in choosing their AI infrastructure investments (fixed costs), defined as $F_{j,t}$, as it prohibits competing firms from market entry. Especially profitable firms with high AI productivity can invest in AI infrastructure to oust firms with low productivity from the market. In spite of the productivity differences between firms, we suppose that all firms charge the same markup—for the sake of simplicity. Therefore, we cannot distinguish between firms with high and low markups and neglect the possibility that changes in markups can arise due to reallocation effects (De Loecker et al., 2021), namely more output of higher markup firms. We only consider the case of homogeneous, symmetric and identical markups for all firms.¹⁵ The following condition needs to hold such that firms are able to finance their AI infrastructure investments and is thus a feasibility constraint for each firm

$$\mu_{j,t} \geq \frac{F_{j,t}}{Y_{j,t}}. \quad (5)$$

As a result, we observe a phenomenon where firms engage in cost-refinancing through markups to accommodate the expenses associated with AI infrastructure.¹⁶ The maximum fixed costs that a firm can impose are determined by the markups it can apply, which are influenced by factors such as demand, total output, and the production output of each individual firm. Building on the work of Jaimovich (2007), it is evident that markups tend to increase as firms invest more in AI infrastructure (fixed costs). Firms that have a high AI productivity and low marginal costs for AI are capable of offering the most competitive AI prices relative to their competitors. These firms, thanks to their

¹⁴A more detailed derivation of markups is provided in the Appendix.

¹⁵Our model could be extended by allowing the markups to depend on the output elasticity (Diez et al., 2021; Jaimovich, 2007), for modelling that only low markups can be charged on goods that are easily substitutable.

¹⁶Nonetheless, we disregard that markups could be set at a level such that they *overcompensate* the investments in AI infrastructure.

cost leadership, are able to capture larger market shares and generate higher profits, as emphasized by Markiewicz and Silvestrini (2021).

4.3 Productivity Effects and Market Barriers

AI addition to affecting firms' AI productivity, AI infrastructure investments can be interpreted as market barriers affecting the number of firms that can enter the market. Moreover, if a firm decides to invest in AI infrastructure, the emerging costs for the investments need to be financed. Firms anticipate that with higher $F_{j,t}$, less productive firms will not be able to afford the market entry, such that there will be less active firms. This allows active firms to apply stronger pricing power due to reduced firm competition. For the sake of simplicity, we assume that if one firm j invests $F_{j,t}$ in AI infrastructure, all other firms \tilde{j} , where $\tilde{j} \neq j$, have to invest the same amount in AI infrastructure $F_{j,t} = F_{\tilde{j},t}$ to be able to compete against firm j in equilibrium which will be explained in more detail in Section 5. We suppose that firms enter the market—implying that they have a strictly positive demand for labor, capital and AI—as long as their profit is expected to be larger than zero. Entrepreneurs are myopic as they only have static and intra-period considerations and are unable to anticipate the long-term effects of their present decisions.¹⁷ Therefore, if firm-owning agents make decisions at period t , they do not consider potential effects of their actions on outcomes in subsequent periods. Moreover, we assume that each firm j has to pay the fixed costs $F_{j,t}$ in each period t in which it produces, implying that infrastructure investments have to be renewed in each period.

Moreover, recall from Assumption 1 that the production of a firm j depends on its productivity, defined by its rank j and the productivity distribution $\Phi(\theta_{j,t})$. The profits of a firm j at time t depending on the number of active firms m_t are defined as

$$\Pi_{j,t,m} = (1 + \mu_{t,m}) Y_{j,t,m} - r_{t,m} K_{j,t,m} - w_{t,m} L_{j,t,m} - p_{j,t,m} A_t - F_{j,t}.$$

Firms pay competitive wages $w_{t,m}$, interest rates $r_{t,m}$ and AI prices $p_{j,t,m}$, which are equal to the marginal value of labor, capital and AI, respectively. The determination of the equilibrium factor prices is described in more detail in Section 5. The Euler theorem, which is explained in the Appendix encloses that the profits can be rewritten as

$$\Pi_{j,t,m} = \mu_{t,m} Y_{j,t,m} - F_{j,t}.$$

We now show how much any firm has to invest in AI infrastructure, such that firms with a lower AI productivity cannot afford to enter the market. Recall that firms are ordered

¹⁷This can also be interpreted such that firms can only operate for a single period t and thus only perform intra-period profit maximization and do not face a recursive multi-period problem.

by their AI productivity, where firm $j = 1$ has the greatest AI productivity and firm $j = N$ has the lowest AI productivity. The following equation defines the number of active firms m_t depending on the fixed costs $F_{j,t}$ chosen by any firm j at time t :

$$m_t(F_{j,t}) = \begin{cases} N & \text{if } F_{j,t} \leq \mu_{t,N}Y_{N,t,N}, \\ N - j & \text{if } F_{j,t} \leq \mu_{t,N-j}Y_{N-j,t,N-j} \quad \text{and} \\ & F_{j,t} > \mu_{t,N-j+1}Y_{N-j+1,t,N-j+1} \quad \forall j \in \{1, \dots, N-1\}, \\ 0 & \text{if } F_{j,t} > \mu_{t,1}Y_{1,t,1}. \end{cases} \quad (6)$$

The explanation of Eq. (6) is the following. If the fixed costs chosen by any firm j are lower than the net profit of the least productive firm $j = N$ operating under perfect competition, given by $\Pi_{N,t,N} = \mu_{t,N}Y_{N,t,N} = 0$, all firms are active on the market enclosing that $m_t = N$ as all firms can afford the fixed costs and still make positive profits. If the fixed costs are higher than the profit of the $(N - j)$ -th productive firm operating under imperfect competition, given by $\Pi_{N-j,t,N-j} = \mu_{t,N-j}Y_{N-j,t,N-j}$, there are $(N - j)$ active firms. If the fixed costs $F_{j,t}$ are larger than the profit of the second most productive firm, given by $\mu_{t,2}Y_{2,t,2}$, only one firm can afford to pay the fixed costs and will be the monopolist.¹⁸ In this fashion, we model an economy where investments in AI infrastructure (e.g., the acquisition of data) raise the barriers to market entry for competitors, pose a potential threat to competition and facilitate a rise of natural monopolies. By taking into account Eq. (6), we note that with infrastructure investments $F_{j,t}$ converging to 0, the number of active firms converges to the maximum number of active firms N , as $\lim_{F_{j,t} \rightarrow 0} m_t(F_{j,t}) = N$, and we therefore note that

$$\frac{m_t(F_{j,t})}{\partial F_{j,t}} \leq 0, \quad (7)$$

enclosing that the lower the fixed costs paid by any firm, the more firms can afford to enter the market. Furthermore, we note that the possibility to invest in AI infrastructure to oust firms from the market depends on the productivity distribution of firms in industrial production and therefore propose the following

Proposition 1

*The factor allocation in equilibrium between m_t active firms in industrial production depends on the productivity distribution $\Phi(\theta_{j,t})$.*¹⁹

For instance, if all firms have the same AI productivity, all firms operate under perfect competition and all firms make zero profits. Therefore no firm invests in AI infrastructure

¹⁸We assume that if a firm leaves the market, it can always re-enter the market by investing the necessary AI infrastructure costs in a future period.

¹⁹We show how the productivity distribution affects the number of firms and factor allocation in equilibrium to prove the Proposition in the Appendix.

and no firm is ousted from the market. On the opposite, if there is an unequal AI productivity distribution, high-productivity firms that can afford to pay higher fixed costs have an incentive to oust low-productivity firms from the market to be able to charge price markups on a market with imperfect competition. Note that by construction of our model, the average productivity of active firms in industrial production is higher the fewer firms are active. This conceptualizes the descriptive results of Effenberger et al. (2020) that indicate that firms in industries with higher market concentration tend to have a higher AI productivity. To sum up, we propose the following which is proved in the Appendix:

Proposition 2

By investing in AI infrastructure that serve as market-barriers, low-productivity firms are ousted from the market. The higher the infrastructure investments $F_{j,t} \geq 0$ of any firm j , the lower the number of firms that are active on the market and thus affects firm competition.

We conclude that depending on the fixed costs chosen by any firm, in particular low-productive firms can be hindered to enter the market. With less firms being active enclosing less competition, higher price markups can be charged, affecting each active firm’s profit, leading to a strategic behavior in deciding on AI infrastructure investments.

5 Equilibrium Definition

In Section 3, we showed our model setup using a simple neoclassical method with industrial production, self-learning AI algorithms and heterogeneous agents. Then, we highlighted the effects of incorporating AI on competition in Section 4 and underlined how AI affects market barriers, firm exit, productivity effects and price markups in the economy.

Now, we define equilibrium conditions such that firms select the optimal input factor allocation depending on K_t , A_t , L_t and Φ_t . We derive the equilibrium factor allocation of firms for and w_t , r_t , p_t and Y_t . We know that consumers demand $q[Y_t]$ is determined depending on total production and we derive the number of active firms, fixed costs and markups in equilibrium. Firms’ and agents’ optimization is jointly performed in a numerical example in Section 6.

5.1 Factor Market Equilibrium

Now, we examine the factor market equilibrium. In an initial step, we derive the optimal allocation of the rival and excludable input factors labor and capital—which are traded on a perfectly competitive input factor market—between firms with heterogeneous AI

productivities. Each firm j produces an industrial good $Y_{j,t}$ using a vector of primary inputs $(A_t, L_{j,t,m}, K_{j,t,m})$ and all firms produce using the production function, given by Eq. (1), maximize their expected profit and take the level of aggregate input factor demand and the level of marginal costs of all active firms as given. We define an input price vector, such that the prices clear the input factor markets to determine the factor market equilibrium. In this way, we derive wages, interest rates, firm-specific AI prices and the distribution of capital and labor amongst all firms depending on the number of m_t active firms. Recall that in contrast to AI, which is a non-rival good that can be simultaneously used by multiple firms, each unit of the labor and capital supply can only be used by a single firm. If a firm uses AI software for production, the same software can also be used by another firm. If a firm hires a worker, this worker cannot be employed by a different firm—the same holds for capital. Recall that total capital endowment at $t = 0$ is uniformly distributed amongst all agents in the economy, total labor supply is given by L and the initial level of AI in the economy is given by A_0 . Due to firms' concave production functions, first-order conditions are necessary and sufficient for the determination of optimal input factor demands of each firm. The optimal demand for each input factor $K_{j,t,m}$, $L_{j,t,m}$ and A_t of each firm j in period t , with m_t active firms, is given by:²⁰

$$r_{j,t,m} = \frac{\partial Y_{j,t,m}}{\partial K_{j,t,m}}, \quad w_{j,t,m} = \frac{\partial Y_{j,t,m}}{\partial L_{j,t,m}}, \quad p_{j,t,m} = \frac{\partial Y_{j,t,m}}{\partial A_t}. \quad (8)$$

Thus, the marginal costs of the input factors, given by $r_{j,t,m}$, $w_{j,t,m}$ and $p_{j,t,m}$, respectively, are equal to their marginal product.

²¹ Due to the concavity of the production function, an increasing use of a single input factor in production leads to an opposite effect on the factor price. The input factor equilibrium at time t with m_t active firms is defined as

$$(A_t^*, L_{j,t,m}^*, K_{j,t,m}^*) = ((A_t^*, L_{1,t,m}^*, K_{1,t,m}^*), \dots, (A_t^*, L_{N,t,m}^*, K_{N,t,m}^*)) \in \mathbb{R}_+,$$

where market clearing conditions for all j firms are given by

$$\begin{aligned} A_t^* &\in A_t(w_{j,t,m}, r_{j,t,m}, p_{j,t,m}), & L_{j,t,m}^* &\in L_{j,t,m}(w_{j,t,m}, r_{j,t,m}, p_{j,t,m}) \\ K_{j,t,m}^* &\in K_{j,t,m}(w_{j,t,m}, r_{j,t,m}, p_{j,t,m}), \forall j \in \{1, \dots, N\}, \forall t. \end{aligned}$$

The equilibrium production of each firm, depending on the number of m_t active firms, is

²⁰For the sake of simplicity, we fix the price of the industrial output to the numeraire.

²¹If the marginal benefit of a firm j with regard to capital is higher than for another firm \tilde{j} for all levels of capital, firm j will use the entire capital stock, whereas no capital is allocated to firm \tilde{j} in equilibrium. This implies a corner solution where firm \tilde{j} is not active and all capital is allocated to firm j . The same reasoning applies to the marginal benefit of labor and possible corner solutions. This implies that if firms productivity are very heterogeneous, it is not necessarily optimal that all firms are active.

defined as

$$Y_{j,t,m}^* = G(A_{j,t,m}^*, L_{j,t,m}^*, K_{j,t,m}^*) \quad \forall j \in \{1, \dots, N\}, \forall t.$$

The equilibrium wage and interest rate which are the same for all firms, are given by:

$$\begin{aligned} r_{t,m}^* &= \alpha \frac{K_{j,t,m}^*}{Y_{j,t,m}^*}^{\frac{-1}{\epsilon}} \\ w_{t,m}^* &= (1 - \alpha) \gamma L_{j,t,m}^* \frac{-1}{\omega} Y_{j,t,m}^* \frac{1}{\epsilon} \left[\left(\gamma L_{j,t,m}^* \frac{\omega-1}{\omega} + (1 - \gamma) (\theta_{j,t} A_t^*) \frac{\omega-1}{\omega} \right)^{\frac{\omega}{\omega-1}} \right]^{\frac{-1}{\epsilon}} \\ &\quad \left(\gamma L_{j,t,m}^* \frac{\omega-1}{\omega} + (1 - \gamma) (\theta_{j,t} A_t^*) \frac{\omega-1}{\omega} \right)^{\frac{1}{\omega-1}}. \end{aligned}$$

We note that the marginal values of labor and capital depend on the number of active firms m_t among which the input factors have to be distributed. As described in Section 3, all firms are charged their marginal value for AI for the acquisition of AI software such that the price for AI may differ between firms. Firms pay a firm-specific price $p_{j,t,m}$ for using the period-specific level of AI, A_t . As AI is a non-rival good, all firms can use the entire AI stock simultaneously. The price for AI in equilibrium is given by:

$$\begin{aligned} p_{j,t,m}^* &= (1 - \alpha) \theta_{j,t} (1 - \gamma) A_t^{\frac{-1}{\omega}} Y_{j,t,m}^* \frac{1}{\epsilon} \left[\left(\gamma L_{j,t,m}^* \frac{\omega-1}{\omega} + (1 - \gamma) (\theta_{j,t} A_t^*) \frac{\omega-1}{\omega} \right)^{\frac{\omega}{\omega-1}} \right]^{\frac{-1}{\epsilon}} \\ &\quad \left(\gamma L_{j,t,m}^* \frac{\omega-1}{\omega} + (1 - \gamma) (\theta_{j,t} A_t^*) \frac{\omega-1}{\omega} \right)^{\frac{1}{\omega-1}}. \end{aligned}$$

The higher the AI productivity $\theta_{j,t}$, the higher the marginal product of AI of a firm j and thus the higher the firm-specific price that is charged for acquiring AI algorithms. The input factor equilibrium determines $K_{j,t,m}^*$, $L_{j,t,m}^*$ and A_t^* for each firm, depending on its productivity and the number of competing firms. These equilibrium values can also be obtained by considering the approach of a revenue-maximizing social planner who maximizes the economy-wide revenue from production, which we describe in the Appendix.²² Therefore, we propose the following which we prove in the Appendix:

Proposition 3

With an increasing level of AI and a higher number of active firms, the effect on the allocation of the rival input factors—capital and labor—depends on the elasticities of substitution ω and ϵ between capital, labor and AI.

With regard to the effect of an increasing level of A_t on optimal factor allocation, recall from Eq. (1) that the marginal returns to labor, capital and AI depend on the level of the respective input factors and the elasticity parameters ω and ϵ , due to the functional

²²For more insights into the micro-foundation for the demand for industrial goods, see e.g., Atkeson and Burstein (2008); Jaimovich (2007).

form of the production function. For example, if labor and AI are substitutable, given by $\omega \geq 1$, firms demand less labor with an increasing level of AI (*ceteris paribus*). In the Appendix, we discuss conditions for e.g., the elasticity parameters ϵ and ω and how they affect capital, labor and AI income depending on the level of AI in more detail. Yet, recall that since AI is non-rival and can be simultaneously used by all active firms, and as each firm pays its marginal value for the acquisition of AI, the greatest AI income is obtained if all firms are active on a competitive market and all firms buy AI algorithms. Moreover, we know due to the properties of the production function, given by Eq.(1), that $\frac{\partial Y_{j,t,m}}{\partial A_t} \geq 0$ so that each individual firm produces more as the level of AI increases. Yet, the effect of the number of active firms on the firm-specific factor allocation remains unclear. The following example can illustrate this. On the one hand, more active firms lead to higher total production. On the other hand, with more active firms m_t , each firm can only obtain a lower share of the input factors labor and capital due to more competition for the rival input factors. Therefore, the effect of m_t on the input factor allocation and, thus firm-specific production $Y_{j,t,m}$ remains unclear. We discuss this in more detail in the Appendix.

5.2 Equilibrium for Consumers' Demand and Markups

The investments in AI infrastructure are financed by the charge of markups. This implies that firms need to charge markups such that they can (at least) refinance their expenditures on AI infrastructure in equilibrium where they make non-negative profits. Moreover, it holds that consumers' demand is satisfied such that

$$\sum_{j=1}^N q_{t,j} = \sum_{j=1}^N Y_{j,t} (\hat{\nu} + \mu_{t,j})^{-\tau} \geq \sum_{j=1}^N Y_{j,t} \left(\hat{\nu} + \frac{F_{j,t}}{Y_{j,t}} \right)^{-\tau}. \quad (9)$$

As a result, we can derive a lower bound for the markups charged by every firm, given by

$$\mu_{j,t} \geq \left(\frac{\sum_{j=1}^N q_{t,j}}{\sum_{j=1}^N Y_{j,t}} \right)^{-\frac{1}{\tau}} - \hat{\nu}. \quad (10)$$

This implies that

$$F_{j,t} \leq Y_{t,j} \left[\left(\frac{\sum_{j=1}^N q_{t,j}}{\sum_{j=1}^N Y_{j,t}} \right)^{-\frac{1}{\tau}} - \hat{\nu} \right]. \quad (11)$$

In our analysis, we focus on an equilibrium scenario where all firms within the market opt for the same fixed costs and apply identical markups. It is important to note that

in such a scenario, the greatest equilibrium output level is achieved when no markups are charged. This is the case in a market scenario with perfect competition, where firms operate in a manner that maximizes total output.

5.3 Individual Utility

Recall that the economy is populated by infinitely many agents with mass L . Entrepreneurs and ordinary workers do not only work, but also save and consume, maximizing the following life-time utility:

$$U_\eta = \sum_{t=0}^{\infty} \beta^t u(c_t^\eta, N_t^\eta). \quad (12)$$

The individual utility is given by an instantaneous concave utility function $u(c_t^\eta)$, which depends on the consumption c_t^η of an agent η in period t . Moreover, it depends on agents' leisure, given by $N_t^\eta = L^\eta - L_t^\eta$. Each individual discounts future consumption with the parameter β . Agents' utility is given by

$$u(c_t^\eta, L_t^\eta) = \log(c_t^\eta) + \kappa \log(N_t^\eta),$$

where parameter $\kappa > 0$ gives the weight that each agent attributes to leisure with regard to overall utility. All agents rent out their labor and capital supply to firms in all periods. Furthermore, we assume that capital depreciates at some exogenously given rate $\delta \in (0, 1)$ such that total capital of all agents has the following law of motion

$$K_{t+1}^\eta = (1 - \delta)K_t^\eta + s_t^\eta, \quad (13)$$

where s_t are the total savings. We consolidate our demand function, contingent on consumers' price elasticity, with the neoclassical framework, acknowledging that production can be either consumed or accumulated as capital for future production. This consolidation necessitates that the following condition remains valid at the aggregate level:

$$Y_t = c_t + K_{t+1} - (1 - \delta)K_t \quad (14)$$

In contrast to ordinary workers, entrepreneurs receive profits from the sale of AI algorithms such that they face different budget constraints.²³ Therefore, an agent's budget constraint in period t depends on the individuals' skill level η and the number of active firms in industrial production. Thus, consumption of agents with skills η and m_t active

²³Since individuals will optimally rent out all capital $K_{t,\eta} = K_{t,\eta}^S$, where $K_{t,\eta}$ is the capital an individual with skill level η has rented out to firms, we will only use $K_{t,\eta}$.

firms at period t is written as follows:

$$c_t^\eta + s_t^\eta = \sum_{j=1}^m [w_{t,m} L_{j,t,m}^\eta + r_{t,m} K_{j,t,m}^\eta] \quad \text{for } \eta \in \{W\},$$

$$c_t^\eta + s_t^\eta = \sum_{j=1}^m ([w_{t,m} L_{j,t,m}^\eta + r_{t,m} K_{j,t,m}^\eta] + p_{j,t,m} A_t + \mu_{t,m} Y_{j,t,m} - F_{j,t}),$$

for $\eta \in \{E\}$; with K_0, L_0, A_0 given.

Whereas ordinary workers with $\eta \in \{W\}$ only obtain labor and capital income, entrepreneurs with $\eta \in \{W\}$ additionally receive all profits from industrial production and from selling AI algorithms, but pay all fixed costs. We define $s_{t,\eta,m}$ as the savings made by agents with skills η at time t and m_t active firms. From now on, we assume that the equilibrium conditions are satisfied in the economy and assess firms' and consumers' utility maximization in the following section.²⁴

6 Optimization Problem

Our analysis involves an two-step optimization problem within a neoclassical model, where agents in the economy cannot anticipate firm decisions with regard to their infrastructure investments in AI, and take wages, prices and interest rates as given. Firms seek to maximize profits by selecting their AI infrastructure investments (fixed costs). First, firms make their decision with regard to their AI infrastructure investments. Afterwards, agents must recursively determine their capital, consumption, and labor supply such that we face a two-stage optimization problem.

Various stakeholding groups in the economy consider different factors when deciding on firms' optimal AI infrastructure investments, as these investments can act as market barriers, affecting the number of active firms, markups, profits, wages, and interest rates. We define the stakeholding group as the set of agents responsible for determining firms' optimal AI infrastructure investment strategy, which additionally affects firms' AI productivity and AI growth (see Eq. (2)). Three potential stakeholding groups are considered:

1. Entrepreneurs
2. Workers
3. Social Planner

Whereas entrepreneurs and workers select the AI infrastructure investments that maximize their group-specific total income, the social planner aims at maximizing total income of all agents. We address the optimization problem in two steps. First, we solve the static

²⁴For the sake of simplicity, we drop the notation using q and v for the demand that depends on the effective price that is affected by the markups and the price elasticity of demand but only use $Y_{t,j}$ for the realized equilibrium production from now on.

profit maximization problem of firms, determining the optimal values of AI infrastructure investment of each firm ($F_{j,t}$) depending on the stakeholding group. Investments in AI infrastructure have productivity effects (see Eq. 3), but can also serve as market barriers (see Eq. 6) affecting the number of active firms, markups, profits, total output and input factor demand. Furthermore, depending on the the AI infrastructure investments, the growth rate of AI and firm-specific AI productivity is affected. After determining the optimal AI infrastructure investments of different stakeholders and deriving firms' demand for labor, capital and AI, we subsequently model agents' decision-making with regard to c_t , K_t , and L_t .

In our analysis, we consider an economy where stakeholders do not have inter-temporal considerations when deciding on firms' AI infrastructure investments. Instead, stakeholders solely focus on maximizing their income within the current period, leading to a myopic optimization. Consequently, stakeholders do not take into account the indirect impact of AI infrastructure investments on e.g. the growth rate of AI.²⁵ To find potential equilibria and determine the preferred market equilibrium for each stakeholding group based on the level of AI, we employ constrained nonlinear optimization algorithms (Lagomarsino, 2020). Nonetheless, agents' decision-making with regard to capital, consumption and labor is optimized using a recursive inter-temporal utility maximization where agents take the period-specific prices, wages and interest rates as given. To sum up, whereas in the first-step, stakeholders only consider the intra-period effect of AI infrastructure investments on their income, agents perform a recursive utility maximization (in the usual neoclassical fashion) in the second step.

6.1 Initial State

In an initial state at $t = 0$, all N firms produce with AI productivity $\theta_{j,t} = 0 \quad \forall j \in N$. Therefore, AI has no effect on production and all firms produce with the same production function

$$Y_{j,0} = \left(\alpha K_{j,0}^{\frac{\epsilon-1}{\epsilon}} + (1 - \alpha) (\gamma L_{j,t})^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}. \quad (15)$$

It is straightforward to see that if the level of AI is zero, $A_t = 0$, or if AI has no positive effect on production, entailing that $\theta_{j,t} = 0$, all N firms have the same productivity and are all active, such that $m_t = N$. Every firm makes the optimal decision regarding its input factor demand in a competitive setting and obtains zero profits $\Pi_t(\theta_{j,t}) = 0$, such that no firm invest any amount on AI infrastructure. Therefore, all (symmetric) firms produce with the production function $Y_{j,0} = G\left(\frac{\sum_{j=1}^N K_{j,0}}{N}, \frac{\sum_{j=1}^N L_{j,t}}{N}\right)$ and make zero profits.

²⁵We assume that firms cannot predict future prices of input factors or the future level of AI.

6.2 Drawing of Firm-specific AI Productivity

At $t = 1$, each firm draws a firm-specific AI productivity from an AI productivity distribution $\theta_{j,t} \sim \Phi(\theta_{j,t})$ as explained in Section 3. Thus, all firms have a new AI productivity compared to $t = 0$. We suppose that each firm knows both, its own AI productivity and the AI productivity of all other firms. Recall that if a firm decides to incorporate AI in production, it has to build up minimum initial infrastructure capacities $D_l \geq 0$. Nonetheless, each firm can decide to invest more in AI infrastructure, namely $F_{j,t} = \overline{F_{j,t}} + D_l$ to increase its own AI productivity, but to also challenge the market entry of competing firms with a lower AI productivity, as described in Section 4.²⁶

6.3 First Step: Firms' Optimization

Assume that a firm j invests an amount $F_{j,t} \geq 0$ in AI infrastructure. All other firms \tilde{j} , with $\tilde{j} \neq j$ can fully observe the AI infrastructure investments of firm j . After observing $F_{j,t}$, a competing firm \tilde{j} can decide how much to invest in AI infrastructure. We have to distinguish between the following cases:

- 1: $F_{\tilde{j},t} = 0$. A firm \tilde{j} does not invest in AI infrastructure and thus cannot compete against firm j which is active and thus does not enter the market.
- 2: $F_{\tilde{j},t} \in (0, F_{j,t})$. A firm \tilde{j} invests to some degree into AI infrastructure, but no sufficient amount to be able to compete with firm j that invests $F_{j,t} > F_{\tilde{j},t}$. Although firm \tilde{j} invests in AI infrastructure, it cannot enter the market.
- 3: $F_{\tilde{j},t} = F_{j,t}$. A firm \tilde{j} invests the same amount into AI infrastructure as firm j . They compete for input factors and both have a strictly positive total industrial output.
- 4: $F_{\tilde{j},t} \in (F_{j,t}, \infty)$. A firm \tilde{j} invests more into AI infrastructure than firm j . Therefore, firm j cannot compete against firm \tilde{j} and is ousted from the market due to preceding considerations.

Consider the following example: If the technological frontier—the most productive firm with $j = 1$, where $\theta_{j,t} > \theta_{\tilde{j},t}$ spends a sufficient amount $F_{j,t}$ on AI infrastructure to be the monopolist, it will be the only producer. Competing firms $\tilde{j} \neq j$ (technological laggards) are always less productive than the technological frontier and can thus only afford to invest $F_{\tilde{j},t} < F_{j,t}$. As investing $F_{\tilde{j},t} \in (0, F_{j,t})$ only leads to negative profits for competing firms, they decide between not entering the market (Case 1) or competing

²⁶In contrast to the approach of Hopenhayn (1992), firms in our model first know their productivity and can then decide on how much they are willing to invest in fixed costs (in our case AI infrastructure). In line with the assumption of Antras and Helpman (2004), firms first observe their productivity level and then decide whether they want to start producing.

against the technological frontier by investing the same amount in AI infrastructure (Case 3).²⁷ Consequently, the investments in AI of firms \tilde{j} are given by $F_{\tilde{j},t} \in \{0, F_{j,t}\}$. Yet, as technological laggards would make negative profits if they invested $F_{j,t}$, they do not enter the market and the technological frontier is the unique active firm on the market. We note that with a decreasing amount paid for infrastructure investments by the technological frontier, more firms can afford to enter the market.²⁸

6.3.1 Stakeholder: Entrepreneurs

Recall that entrepreneurs obtain a share l^E of the total labor and capital income. Moreover, they receive the entire profits generated by selling AI algorithms as they collectively own the AI company. Furthermore, they benefit from the entire profits made by firms in industrial production, but have to bear all fixed costs for installing AI infrastructure. If entrepreneurs are the stakeholding group, they maximize their income in each period by collectively deciding on the profit-maximizing AI infrastructure investments of each firm. Entrepreneurs' income is given by

$$E_{t,m} := \sum_{j=1}^m \left(\underbrace{w_{t,m} L_{j,t,m}^E + r_{t,m} K_{j,t,m}^E}_{(I)} + \underbrace{p_{j,t,m} A_{j,t,m}}_{(II)} + \underbrace{\mu_{t,m} Y_{j,t,m} - F_{j,t}}_{(III)} \right) \quad (16)$$

Thus, if entrepreneurs are the stakeholding group, they chose the AI infrastructure investments that maximize their income in each period in line with

$$\begin{aligned} \max_{\{F_{j,t}\}} E_{t,m} \quad \text{s.t.} \quad & F_{j,t} \leq Y_{t,j} \left[\left(\frac{\sum_{j=1}^m q_{t,j}}{\sum_{j=1}^m Y_{j,t}} \right)^{-\frac{1}{\tau}} - \hat{\nu} \right], \\ & A_{j,t,m} = \hat{A}_t, \quad \sum_{j=1}^m L_{j,t,m}^E \leq L_t^E, \quad \sum_{j=1}^m K_{j,t,m}^E \leq K_t^E, \quad \forall t. \end{aligned}$$

There are different channels via which the number of active firms and the level of AI affect the income of entrepreneurs. Referring to (I), the highest labor and capital income can be obtained if firms operate on a competitive market where no markups are charged.²⁹ Yet, the effect of a higher level of AI on the capital and labor income depends on the elasticity of substitution, given by ϵ and ω and thus remains unclear, but is further discussed in the Appendix.³⁰ Moreover, entrepreneurs benefit from the profit by selling

²⁷Technological laggards can never afford to invest more than the technological frontier (Case 4).

²⁸As no other firm is able to pay more on AI infrastructure than the technological frontier, it suffices to only consider its decision. The decisions of all other firms follow analogously.

²⁹Linked to Proposition 3, we discuss how the number of firms affects total production and thus the capital and labor income due to the heterogeneous AI productivity across firms in the Appendix

³⁰Nonetheless, we see that the higher the share of entrepreneurs in the economy—the greater l^E —the more they benefit from wage and interest payments relative to net profits.

AI algorithms, given by (II), which depends on the level of AI, the number of firms to which the algorithms are sold and their respective AI productivity and total production. Recall that total production depends on the number of active firms, markups and the AI productivity of firms such that general statements on the dependence of (II) on the level of AI and the number of active firms are not possible. In addition, the effect of more active firms (a higher level of AI) on entrepreneurs' profits (III) remains unclear as less (more) fixed costs have to be paid, but also lower (higher) markups can be charged. For example, in a market with zero markups, entrepreneurs do not earn any profits (III), but receive the greatest total wage and capital income (I), and can sell their algorithms to many firms as total production is high. Conversely, in a monopoly situation with only one active firm $m_t = 1$, entrepreneurs gain monopoly revenue but incur high fixed costs to drive all competing firms from the market and a lower labor and capital income due to a reduced output. Furthermore, they are restricted to selling the AI algorithms exclusively to a sole industrial firm (II) with a lower output that has a lower output than it would have in a competitive setting. Thus, we cannot derive general conclusions with regard to the optimal AI infrastructure investments from the perspective of the entrepreneurs. The reason is that depending on the level of AI, the different channels (I) – (III) have changing relative importance for the income stream of the entrepreneurs.

6.3.2 Stakeholder: Workers

A share of l^W agents in the economy are ordinary workers who only receive capital and labor income. Workers' income is given by

$$W_{t,m} := \sum_{j=1}^m \left(\underbrace{w_{t,m}L_{j,t,m}^E + r_{t,m}K_{j,t,m}^E}_{(I)} \right). \quad (17)$$

Their income maximization problem when they are the stakeholding group is analogous to the one of entrepreneurs, as previously explained. As workers only benefit from labor and capital income (see (I)), they obtain the highest income if firms operate on a competitive market where no markups are charged. This is economically intuitive as wage and interest payments are the greatest on a perfectly competitive market in our model. Thus, workers and entrepreneurs have different considerations when deciding on the optimal AI infrastructure investments.

6.3.3 Stakeholder: Social Planner

A benevolent social planner aims at maximizing total income of all agents in the economy. The social planner chooses the AI infrastructure investments $F_{j,t}$ that maximize the

total income of all agents, given by $I_{t,m}$, jointly maximizing the income of workers and entrepreneurs. Therefore, the social planner maximizes the following income stream:

$$I_{t,m} := \sum_{j=1}^m [(1 + \mu_{t,m})Y_{j,t,m} - F_{j,t}] \quad (18)$$

Again, the income maximization is analogous to the one of workers and entrepreneurs. Note that it can be socially optimal for the total income of all agents that active firms operate on an imperfectly competitive market and charge markups. In such a case, fewer firms would be active, leading to lower capital and labor income, but higher profits would be distributed, with the latter outweighing the former effect. Then, although total income would be the greatest, there would be inequality between the incomes of workers and entrepreneurs if no redistribution mechanisms were embraced. Therefore, we will later show the income inequality depending on the stakeholding group.

6.4 Second Step: Individual Optimization

After the determination of the optimal AI infrastructure investments depending on the stakeholding group and obtaining the input factor allocation, the wages, interest rates, AI prices, markups, and profits are derived. Agents in the economy subsequently optimize with regard to their consumption, savings and endogenous labor supply which then determines the input factor availability in the next period. Due to income accounting in our economy, income either has to be consumed or saved, such that on aggregate, it holds that

$$\sum_{\eta} c_t^{\eta} + \sum_{\eta} s_t^{\eta} = \sum_{j=1}^m [(1 + \mu_{j,t,m})Y_{j,t,m} - F_{j,t}]. \quad (19)$$

Total consumption and saving need to be equal to the total demanded aggregate output in equilibrium, given by $\sum_{j=1}^m Y_{j,t,m}$ times one plus the possible markup, given by $(1 + \mu_{j,t,m})$ minus the total fixed costs spent on AI infrastructure, given by $\sum_{j=1}^m F_{j,t}$.

Agents assume the level of AI to be fixed and given exogenously. Thus, they do not anticipate its growth dynamics, but assume the value of AI to be fixed, which we call \hat{A}_t . The utility maximization problem of agents—if we regard a simplified framework with a single group of agents indexed by parameter η —reads as follows

$$\begin{aligned} & \max_{\{K_{t+1}^{\eta}, N_t^{\eta}\}} \sum_{t=0}^{\infty} \beta^t u(c_t^{\eta} + (1 - \delta)K_t^{\eta} - K_{t+1}^{\eta}, N_t^{\eta}) \\ & c_t^{\eta} \geq 0 \quad K_{t+1}^{\eta} \geq 0 \quad L_t^{\eta} \geq N_t^{\eta} \geq 0 \quad \hat{A}_t \geq 0 \quad K_0^{\eta} \text{ given} . \end{aligned}$$

In the Appendix, we show the utility optimization framework in a simplified framework with a single group of agents, including the first-order condition and the recursive formulation of the problem.³¹ Based on agents' decision on their capital, consumption and labor supply in the second step, the available input factors in the next period for the determination of the optimal AI infrastructure in the first step of the next period are specified. Thus, the two steps of our optimization problem are interlaced, where firms' decisions in the first step affect the equilibrium prices and agents' decisions in the second step, which again affect the available inputs for firms' optimization in the subsequent period.

To sum up, we note that different stakeholding have different preferences with regard to the optimal AI infrastructure investments due to different income streams. On the one hand, the profit interest of entrepreneurs can lead to high AI infrastructure investments enclosing imperfect competition, markups and reduced output. On the other hand, workers prefer a competitive environment without markups charged. Given the interlaced optimization problem and the intricate growth pattern of AI within the economy, it proves difficult to derive universally applicable insights through algebraic solutions in our model. As a viable alternative, we use a numerical model quantification to highlight the primary conclusions drawn from our framework. We additionally show the effect of different stakeholders' decision on the development of income inequality and factor income shares.

7 Numerical Exercise

As described in Section 6, we execute a two-step optimization procedure. First, we conduct firms' intra-period profit maximization and, subsequently, agents' recursive optimization problem of capital, labor and consumption. We illustrate the optimal AI infrastructure investments depending on three stakeholding groups: i) Workers, ii) Entrepreneurs and iii) the Social Planner. Depending on the stakeholding group, we compare the resulting market outcomes, namely the number of active firms, markups, labor supply and capital and consumption development. Moreover, we assess the growth rate of AI and the development of the factor income shares. We can partly motivate our parameter choice using findings of the literature. In our numerical exercise, we choose the following parameters, given in Table 1. For modelling each firms' AI productivity, we are motivated by trade literature (e.g., Helpman et al. (2004); Perla and Tonetti (2014)) and use a Pareto distribution.³²

³¹We set up the utility maximization problem for a framework with perfect competition and in a framework with imperfect competition.

³²The Pareto distribution is given by $\phi(j) = \frac{\zeta\lambda^\zeta}{(N-j+1)^{\zeta+1}}$, where we set the shape parameter ζ and the scale parameter λ . Yet, motivated by Nigai (2017), we use a lower value for ζ than in Melitz and Redding (2013).

Class	Parameter Choice	Literature	Literature Calibration
Production	$\alpha = 0.33$ $\omega = 1.25$ $\epsilon \in \{0.8, 1.25\}$ $\gamma = 0.9$	Mankiw et al. (1992); Kydland and Prescott (1982); King and Rebelo (1989) Lu (2020); Aghion et al. (2020) Klump and Saam (2008); Raurich et al. (2012) Lu (2020)	$\alpha = 0.33; 0.36; 0.33$ $\omega = 1.2222; 1.011$ $\epsilon \in [0.8, 1.2]; [0.63, 1.58]$ 0.76; 0.5
Individuals	$l_t^E = 0.15$ $\delta = 0.1$ $\tau = 2.5$	Lu (2020) King and Rebelo (1989) Cette and Lopez (2012)	$l_t^W = (0.0965, 0.3)$ $\delta = 0.1$ $\tau = 2.42$
Pareto Dist.	$\zeta = 1$ $\lambda = 1$	Melitz and Redding (2013) Melitz and Redding (2013)	$\zeta = 4.25$ $\lambda = 1$
AI	$x = 0.9$ $w = 4$ $B = 300$ $\iota = 0.1$ $\nu = 0.1$	N.A. N.A. N.A. N.A. N.A.	

Table 1: Parameters for the Numerical Exercise.

Parameters determining the growth dynamics of AI and the effect of AI infrastructure investments and better tailored algorithms due to a higher market shares cannot be motivated by economic literature. However, inspired by the analysis of Epoch, a team of researchers investigating and forecasting the development of AI, we refer to the predicted AI growth rate as given by <https://takeoffspeeds.com/> that we aim at rebuilding using an incomplete beta distribution for modelling the growth function of AI. Yet, we choose parameters determining the growth pattern of AI in line with the forecast of Epoch. Moreover, the concept of price elasticity has not yet been applied to (AI) software. Thus, we cannot motivate our chosen price elasticity of AI demand using related literature, but set it to $\tau = 2.5$. We base this value on findings on the price elasticity of ICT at its gradual introduction in the 1980s (Cette and Lopez, 2012).³³

Moreover, there is no estimate for κ in an AI-based economy which we set to 0.78.³⁴ Based on our knowledge and the statement of Lu (2020), there is no empirical data on the elasticity of substitution between human labor and AI and the factor share of AI, yet. Therefore, we infer the values for our numerical exercise based on the literature on automation, machines, robots and labor. For our default analysis, we set $\epsilon = 1.18$ and $\omega = 1.25$ as we assume substitutability between labor and AI, and capital and labor. In the Appendix, we consider complementarity between labor and capital and the effect on the factor income shares. We select the following starting values for capital, labor and AI in the economy: $K_0 = 50, L_0 = 50$ and $A_0 = 2$. Moreover, we observe the

³³Although we expect some gradual integration of AI into economic processes, we do not assume a time-variant price elasticity τ .

³⁴We use an incomplete beta distribution for modelling the growth function of AI, given by $b(\cdot)$, using $x = 0.5$ and $w = 4$. The incomplete beta function is defined as :

$$I_x(Y_t, w) = \frac{1}{B(Y_t, w)} \int_0^x h^{Y_t-1} (1-h)^{w-1} dh$$

where we use the beta function $B(Y_t, w)$

$$B(Y_t, w) = \int_0^1 h^{Y_t-1} (1-h)^{w-1} dh = \frac{\Gamma(Y_t)\Gamma(w)}{\Gamma(Y_t + w)}$$

which is based on the gamma function $\Gamma(Y_t) = \int_0^\infty h^{Y_t-1} e^{-h} dh$.

economy for $T = 35$ periods where AI grows endogenously via its self-learning features as described in Eq. (2) and assume that a maximum of $N = 50$ firms can operate in industrial production.

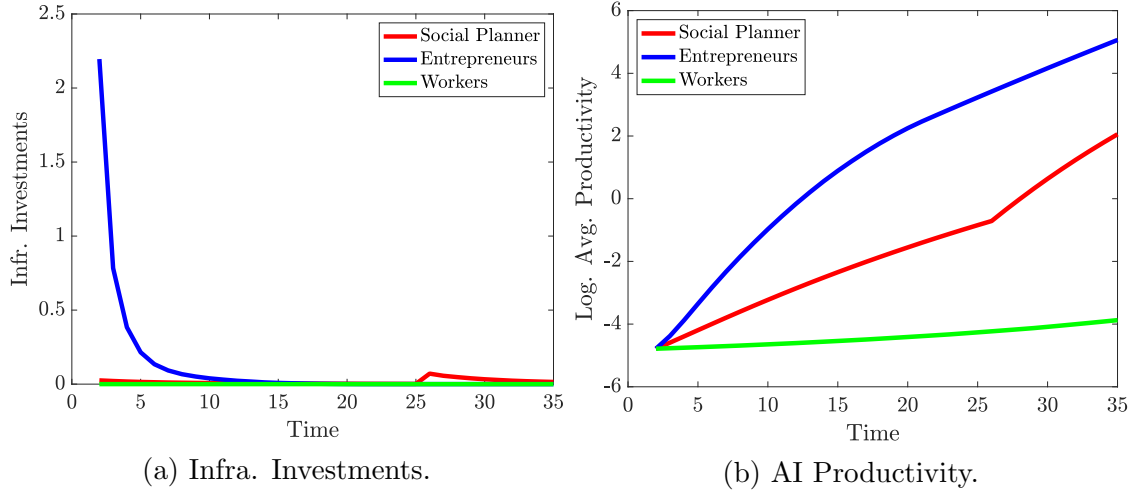


Figure 2: Development of AI Infrastructure Investments and AI Productivity Depending on the Stakeholder.

In Figure 2, we depict the optimal AI infrastructure investments depending on the stakeholder and the development of the AI productivity over time. We note that in particular if entrepreneurs are the stakeholding group, there are high investments in AI infrastructure to foster AI productivity growth and to maximize firms' profits.³⁵ The reason is that entrepreneurs aim at investing in AI infrastructure, not only to increase firms' productivity via Eq. (3), but to hamper competitors' market entry via Eq. (6) for obtaining larger profits. In contrast, workers solely aim at maximizing their capital and labor income, which is the greatest in a perfectly competitive environment without markups, such that there are no AI infrastructure investments going beyond the necessary investments D_l for the installation of AI in production.

Moreover, we note that there is a persistent increase in the average AI productivity—irrespective of the stakeholding group. This phenomenon stems from the advantageous position held by large firms, as AI algorithms are better tailored to their production, leading to significant increases in their AI productivity.³⁶ Increased market shares, in a reciprocal fashion, enable the continual fine-tuning of AI algorithms to better suit the unique needs of larger firms, thereby facilitating ongoing enhancements of their AI productivity. The social planner, taking into account the income of entrepreneurs and workers,

³⁵This trend persists even though agents in the economy do not consider the potential influence of increased infrastructure investments on firms' productivity due to their static and myopic profit considerations.

³⁶Two opposing effects influence the growth of AI: AI can be improved by being applied to more productive firms or by being applied to more firms. Determining which effect dominates is challenging and cannot be generalized.

chooses AI infrastructure investments that encompass a growth rate of AI productivity positioned between that of entrepreneurs and workers.

We note that the necessary AI infrastructure investments (fixed costs) that firms need to incur to oust competing firms from the market decrease over time. The reason is that large firms have a persistently-growing AI productivity making market entry of less productive firms increasingly difficult anyway. Firms with a low productivity cannot catch up to the productive firms as they invest less in AI infrastructure and algorithms are decreasingly tailored to their production, enhancing the increase in market concentration. We interpret this as an indicator that the self-learning feature of AI and the tailoring of algorithms to firm-specific needs can lead to monopolisation.

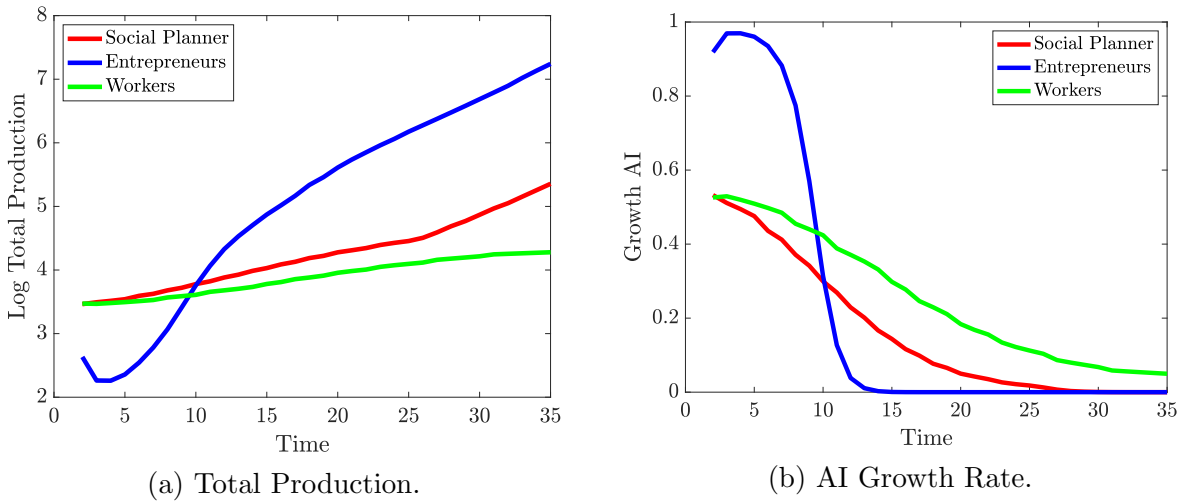


Figure 3: Development of Total Production and AI Growth Depending on the Stakeholder.

In Figure 3, we depict the development of the growth rate of AI and total production in our economy. Regardless of the decisive stakeholding group, we observe a persistent increase in total production due to the increasing level of AI and the higher AI productivity of firms. This finding highlights the importance of AI as a driver for growth. Yet, we observe that total production is initially higher if the social planner or workers decide on the optimal AI infrastructure investments. The reason is that entrepreneurs aim at charging markups to refinance their investments in AI infrastructure leading to reduced total production and an output gap. Yet, as entrepreneurs invest in AI infrastructure which enhances firms' productivity, we note that total production starts being the largest in case entrepreneurs are the stakeholding group after around 10 periods. Moreover, we note that the AI growth rate is especially high in the first periods due to entrepreneurs' high investments in AI infrastructure, as depicted in Figure 3b). In the long run, the growth rate declines over time for all stakeholding groups. After reaching its upper bound in the long run, the AI growth rate becomes zero due to the assumption of an upper bound for AI due to hardware restrictions, as explained in Eq. (2).

In addition, we illustrate the development of the market concentration with a rising level of AI and depending on the stakeholding group in Figure 4 using the Herfindahl-Hirschman Index (HHI).³⁷

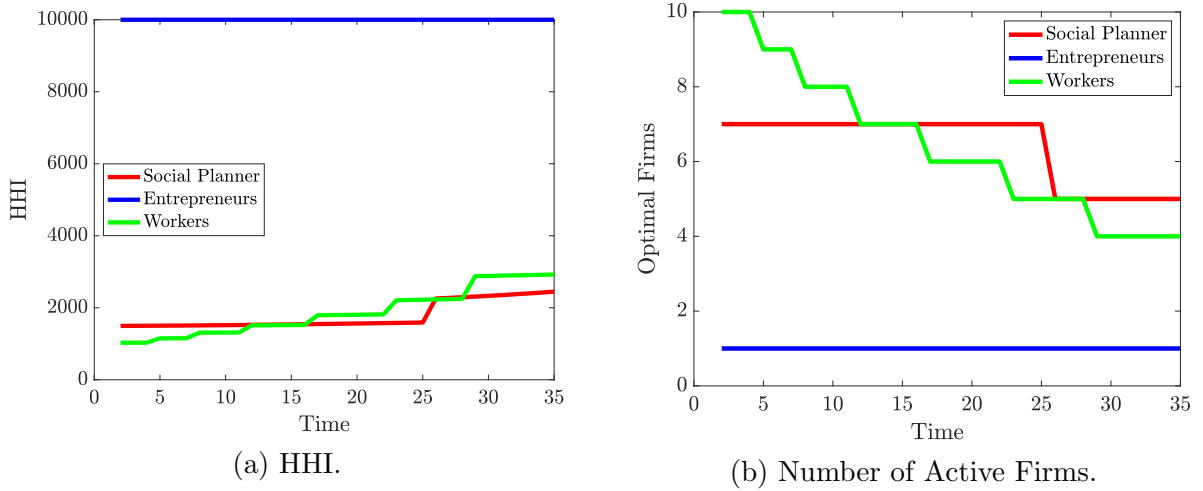


Figure 4: Development of Market Concentration Depending on the Stakeholder.

For the sake of simplicity, we set the parameters to assess a scenario, where entrepreneurs favor an imperfectly competitive market with a single operating firms such that they benefit from monopoly profits. However, in scenarios where workers or the social planner act as the stakeholding group, we still observe an increase in market concentration, driven by the rising productivity disparity between active firms.³⁸ The reason is that especially large and productive firms have an increased market share due to investments in AI infrastructure and the tailoring of AI algorithms that especially benefit large firms such that there is a decline in the number of active firms. We observe that if the social planner selects the optimal AI infrastructure investments, there is a decline in the number of active firms coming in hand with a rising market concentration. However, this increase in concentration is even more pronounced if workers are the stakeholding group such that a scenario eventuates where the HHI reaches a level of around 4000, with only 4 firms being active in the long-run.³⁹

³⁷We determine the HHI using the following equation, where x_i represents the output of a single firm i , and $H = \sum_{i=1}^N \left(\frac{x_i}{\sum_{j=1}^N x_j} \right)^2$. The value of HHI thus obtained is multiplied by 10000 and thus takes values between $\frac{10000}{N} \leq H \leq 10000$.

³⁸In the numerical example at hand, as the share of entrepreneurs in the economy is relatively small, the social planner solution mainly corresponds with the market outcome that maximizes workers' income.

³⁹We set the parameters in our numerical exercise such that the HHI at the first period is around 1500, which approximates the value for the HHI in the ICT sector in Germany in 2020 (Maydell and Menzel, 2023).

7.1 Markups and Output

In Figure 5b), we depict how the share of total output demanded is negatively correlated with the markups charged by active firms, as defined in Eq. 4.⁴⁰ Recall that markups can only be charged when there is a reduction in output, as defined by the demand equation (Eq. 4). This implies that higher markups lead to an increase in the output gap.

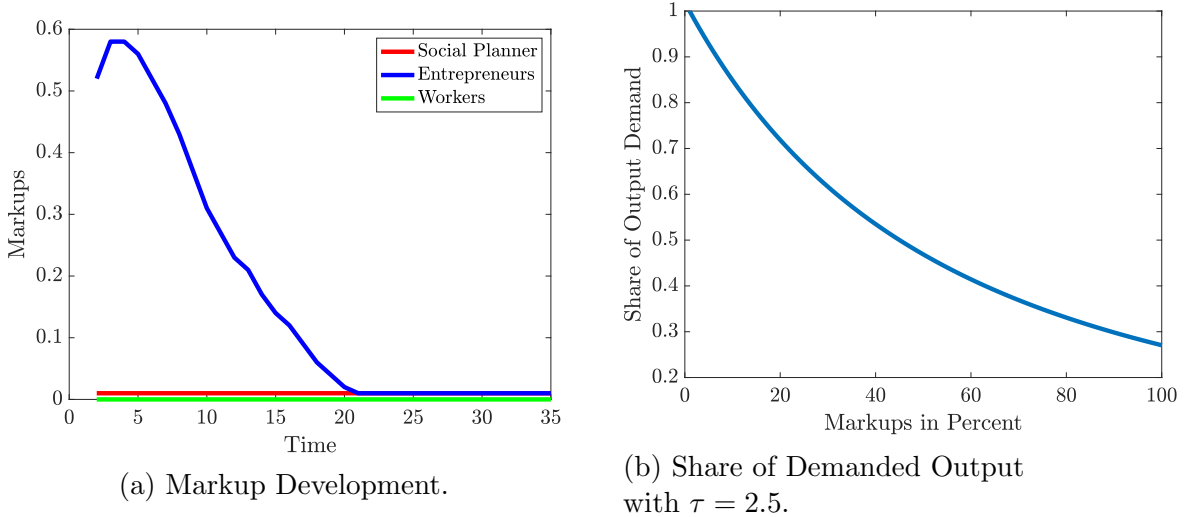


Figure 5: Markup Development and Share of Demanded Output.

Consequently, firms that charge markups on their products reduce their overall output, resulting in a decreased use of capital and labor in industrial production. We note that if workers are the stakeholding group, no markups are charged, such that the highest possible capital and labor income is paid out. Whereas the markups in the social planner’s solution remain at a low, but nonzero value, we observe decreasing markups if entrepreneurs decide on investing in AI infrastructure, as they charge markups for refinancing their costs. We note that if entrepreneurs are the stakeholding group, markups are charged especially in the first periods as firms with high AI productivity invest in infrastructure to oust competing firms from the market and to increase their profits and market shares. Subsequently, due to the self- and firm-specific learning of AI, investments in AI infrastructure increasingly decline coming in hand with lower markups. The reason is that the possibility to obtain larger AI profits by having a high total output and the possibility to sell many AI algorithms to industrial firms (Channel *(II)* in Eq. (17)) increasingly outweighs the profit surplus by reducing output and charging markups (Channel *(III)* in Eq. (17)) with a rising level of AI. In the long-run, even if entrepreneurs are the stakeholders, only low markups are charged as a single firm has reached such a high productivity that a scenario with low markups and a low output gap leads to the

⁴⁰We additionally emphasize that the magnitude of markups imposed by firms is inversely related with consumer price elasticity. When consumers are highly responsive to price changes and have a greater price sensitivity, given by τ , firms find themselves with limited pricing power.

highest income for entrepreneurs. Therefore, we note that the markups converge to a positive but non-zero value in the social planner’s and entrepreneur’s solution in the long-run.

In summary, our findings reveal that when entrepreneurs are the stakeholding group, they invest in AI infrastructure—in particular in the first periods after the introduction of AI—leading to larger productivity increases, in particular of firms with high AI productivity. This is accompanied by higher markups for the refinancing of the investments, albeit at the expense of lower total production. Increased AI infrastructure investments lead to enhanced AI productivity while additionally acting as market barriers for firms with a low AI productivity such that we observe high markups and high investments in AI infrastructure, particular at the beginning of the period under investigation if entrepreneurs are the stakeholders. We emphasize that due to the learning of AI algorithms that are increasingly tailored to the needs of large and productive firms that can afford to invest in AI infrastructure, it becomes increasingly difficult for small firms to compete against these large firms. Therefore, we note a growing market concentration due to the self-learning and firm-specific tailoring of AI algorithms that is especially pronounced if entrepreneurs are the stakeholding group, but can still be observed to a reduced extent if workers or the social planner are the stakeholder. Yet, we emphasize that AI is a driver for economic growth, productivity increases and total production, irrespective of the stakeholding group.

7.2 Income Divergence

Now, we aim at illustrating how the income of workers and entrepreneurs and in particular the divergence between the income of the two different groups is affected by the stakeholder deciding on the AI infrastructure investments. The income of workers—who solely benefit from renting out their labor and capital to firms—is directly linked to total production, which, in turn, is affected by the number of active firms and the markups charged. Following our theoretical considerations, it is optimal for workers to receive competitive wages and interest rates in a scenario where no markups are charged, irrespective of the level of AI. Recall, from Eq. (17), that entrepreneurs also generate income by renting out their capital and labor. Additionally, they obtain the entire profits from the AI-developing company and from industrial production. However, they are also responsible for covering the entire fixed costs associated with AI infrastructure investments. A benevolent social planner aims at choosing the AI infrastructure investments that maximize the total income of all agents in the economy which is a linear combination of the income of entrepreneurs and the income of workers.⁴¹

⁴¹Nonetheless, the social planner does not take into account potential income divergence between workers and entrepreneurs as only total income when determining the optimal infrastructure investments. The reason is that the social planner only regards the effect of income on the utility of agents when

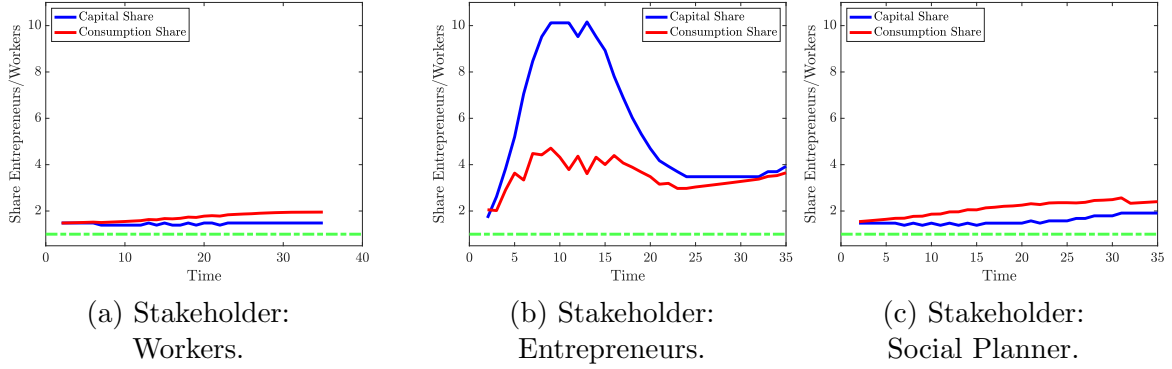


Figure 6: Income Inequality Depending on the Stakeholding Group.

We depict how the capital and consumption of entrepreneurs relative to that of workers develops in the economy over time in Figure 6. We differentiate between the scenarios if i) Workers, ii) Entrepreneurs or the iii) Social Planner are the stakeholding group deciding on the optimal AI infrastructure investments. In particular in the first 20 periods when there are high markups charged, we observe large differences between the capital and consumption between entrepreneurs and workers. Yet, also in the long run, income inequality is more pronounced in scenario ii) than in scenario i) or iii). Therefore, we contend that the rise of AI contributes to growing income inequality between workers and entrepreneurs, in particular in a scenario when entrepreneurs favor an imperfectly competitive market with high markups and reduced competition. Our findings suggest that the primary driver of this income gap is not solely the fact that only entrepreneurs benefit from selling AI algorithms. Rather, it is the preference of entrepreneurs for firms to operate on imperfectly competitive markets for charging markups and obtaining high profits, which amplifies income disparity.

7.3 Factor Income Shares

Now, we compare the evolution of the factor income shares depending on the stakeholding group. We examine the development of the factor income shares in Figure 7 for specific elasticity parameters, namely $\varepsilon = 1.18$, $\omega = 1.25$ and the additional parameters defined in Table 1. Note that the elasticities between labor and AI, given by ω , and between capital and labor and AI, given by ε are decisive for the development of the factor income shares. We argue that it is most reasonable that $\varepsilon > 1$ and $\omega > 1$ due to our expectation that AI will be able to substitute for labor.⁴²

selecting firms' optimal AI infrastructure investments in their intra-temporal profit maximization and cannot infer agents' resulting utility from their inter-temporal utility maximization.

⁴²Motivated by Aghion et al. (2020), stating that the elasticity of industry-level employment to industry-level automation is 1.011, we assume that the elasticity of substitution between labor and AI is given by $\omega > 1$. Furthermore, Karabarbounis and Neiman (2014) argue that the elasticity of substitution between labor and capital is $\varepsilon = 1.25$.

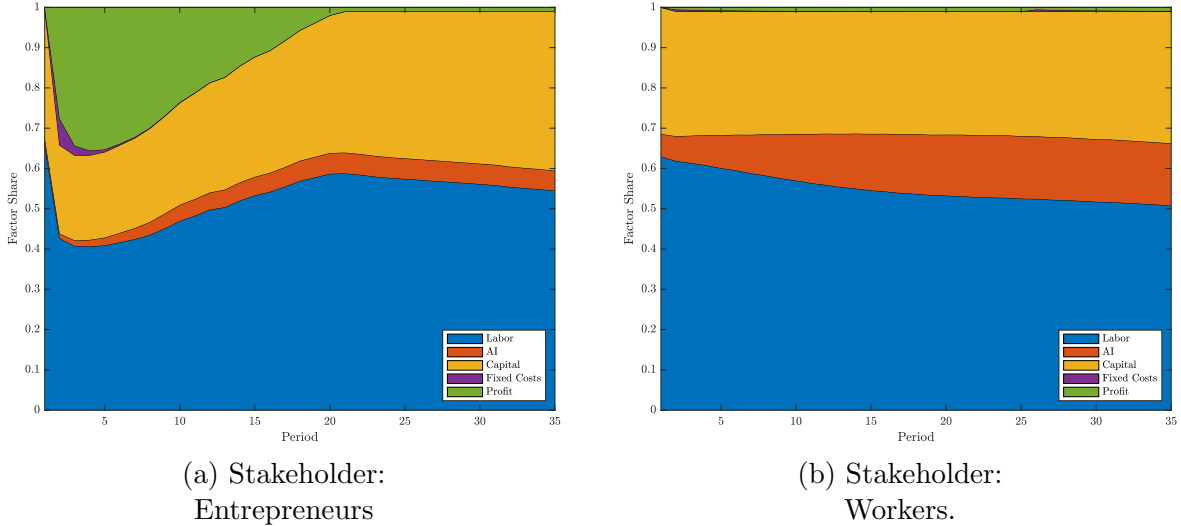


Figure 7: Development of the Factor Income Shares with $\epsilon = 1.18$ and $\omega = 1.25$.

As we suppose that $\omega > 1$, which implies that AI is a substitute for labor, we observe a decreasing labor share and an increasing AI share in both scenarios, irrespective whether workers or entrepreneurs are the stakeholding group.⁴³ However, we note that the development of the factor income shares strongly depends on the stakeholding group. Whereas the labor share decreases from 63% to 51% in an environment under perfect competition if workers are the stakeholder, the labor share would reach a level of only 55% 35 periods after the introduction of AI if entrepreneurs are the stakeholders, but reaches a temporary minimum of around 40%. After the introduction of AI on a market with imperfect competition, we note that the declining labor share is accompanied with an increasing profit share that only benefits entrepreneurs. The reason is that an increasing share of total income would be attributed to the profit payments due to markups on the output price. Yet, after 11 periods the profit share declines again as the markups decrease, leading to lower profits. However, there is a higher capital share on the market with imperfect competition than on the market with perfect competition after 35 periods. For the sake of completeness, we additionally illustrate the development of the factor income shares if we assumed complementarity between capital and labor, given by $\epsilon = 0.85$ in the Appendix.

In our model, the number of active firms and the total supplied output affect the demand for the input factors labor, capital and AI. Moreover, in our framework with a labor-leisure tradeoff, we observe that the labor supply depends on the level of AI, i.e., with a rising level of AI, all agents adapt their labor supply depending on the elasticity of substitution between the input factors and their income stream. Therefore, we illustrate the share of working agents depending on the stakeholding group in Figure 8, based on

⁴³If we assumed e.g. a larger substitutability between labor and AI, the decline of the labor share would be attenuated. The same reasoning holds for changes in the elasticity between labor and capital.

our baseline parameter values provided in Table 1.

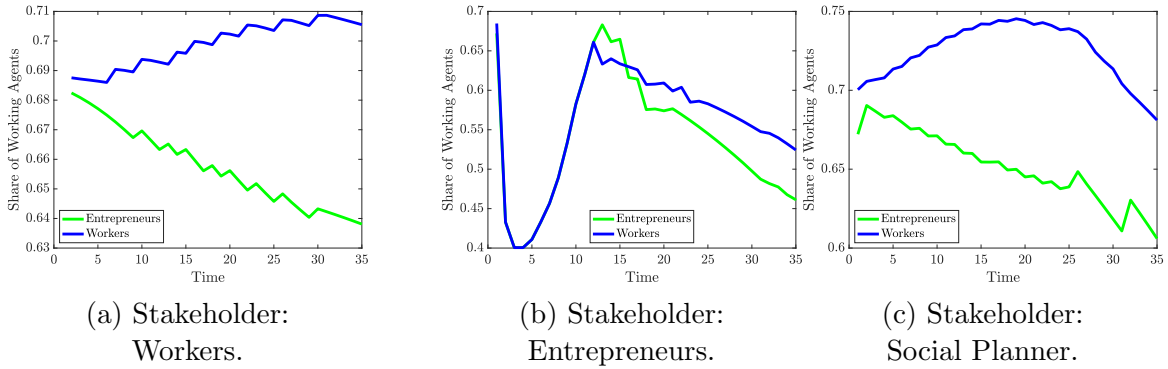


Figure 8: Share of Employed Agents Depending on Stakeholder.

We note that entrepreneurs reduce their labor supply over time irrespective of the stakeholding group. If workers are the stakeholding group, we note that workers increase their labor supply over time, whereas entrepreneurs decrease their labor supply as they increasingly benefit from AI and profit income that they obtain without *labor effort*. Irrespective of the stakeholder, it holds that entrepreneurs reduce their labor supply more than workers as they additionally benefit from firms' profits which makes them less dependant on the labor income. If entrepreneurs are the stakeholding group, we observe a lower level of employed workers in the long run. The observed phenomenon is not only caused by agents reducing their labor supply, as we might see if workers or the social planner are the stakeholder, but mainly due to a decreased labor demand. This is especially the case in the first 15 periods, where a lower total output is produced and markups are charged. Therefore, as these firms that charge markups and aim at achieving greater profits, active firms demand fewer input factors and reduce their labor demand such that not the labor supply but the labor demand determines the lower bound for the labor market equilibrium. However, we again observe that entrepreneurs reduce their labor supply to a larger extent than workers, also in the long-run. If the social planner is the stakeholding group, we still notice that entrepreneurs reduce their labor supply over time, but workers first increase and then decrease their labor supply.

As empirical research on the elasticities of substitution between labor, capital, and AI is limited, we are unable to reach a definitive conclusion regarding the impact of AI on factor income shares. However, regardless of the elasticities, we note that if entrepreneurs are the stakeholding group, there is a rise in imperfect competition coming in hand with the rise of AI and an increasing profit share particularly benefiting entrepreneurs. Moreover, entrepreneurs increasingly reduce their labor supply as they can benefit from *non-human work* and are the only group that benefits from AI profits. As a result, we observe a widening income gap between entrepreneurs and workers in particular on markets with AI-induced imperfect competition.

8 Policy Interventions

Comparing the equilibrium allocation in an economy where entrepreneurs are the stakeholders with the allocation preferred by a social planner, we observe that fewer firms are active, charge higher markups, (initially) produce less than the social optimum, and demand a smaller amount of labor than in the social optimum. Moreover, in particular on an imperfectly competitive market, due to the profit distribution solely to entrepreneurs, there is an increasing divergence between workers' and entrepreneurs' capital and consumption, enclosing increasing inequality. The group of entrepreneurs has a shared interest in colluding to intentionally set the investments in AI infrastructure at a higher level than desired by the social planner. This collaborative effort aims to create an environment of imperfect competition in industrial production, with the ultimate goal of maximizing entrepreneurs' income.

We aim at exploring potential mechanisms that help promoting the development of AI as a driver of growth but prevent the rise of an increasingly unequal society in an AI-based economy. Thus, we discuss the effect of a profit tax, new data sharing standards and a modernization of competition and merger legislation. The main political difficulty is to not stifle the investments of entrepreneurs in AI infrastructure that lead to high growth in AI productivity and total output accompanied by rises in market concentration and markups, but to also mitigate growing income inequality.

8.1 Profit Tax

Our model highlights that a social planner focuses solely on maximizing the total income of all individuals in the economy fails to consider the growing disparity in income distribution. By exclusively prioritizing total income, this approach overlooks the widening gap between the income of entrepreneurs and regular workers. Consequently, we underscore the importance of implementing re-distributive taxes to curb the exacerbation of income inequality resulting from the rise of AI. For instance, to counterbalance this trend, wage income could receive preferential treatment compared to income from capital and AI. This encloses that profits which are solely distributed to entrepreneurs, should be subject to higher tax rates than human-earned wage income. Similarly, Faulhaber (2019) highlights the necessity to modernize digital taxation to target multinational Tech companies and to establish an effective international tax system, e.g., using a minimum tax to guarantee international competition. Maintaining a competitive environment for broadly sharing technological rents (Ernst et al., 2019) could prevent firms from reaching market dominance, which may render a fair and socially-beneficial economic integration of AI difficult. One viable strategy for reducing income inequality involves the introduction of a profit tax for entrepreneurs. Through the imposition of a tax on entrepreneurs' profits, workers can also participate in the benefits arising from the integration of AI. By intro-

ducing a profit tax $\nu \in (0, 1)$ that is deducted by the entrepreneurs and re-distributed to the workers, the budget constraints of the agents would be rewritten in the following way:

$$c_t^W + s_t^W = \sum_{j=1}^m [w_{t,m}L_{j,t,m}^W + r_{t,m}K_{j,t,m}^W] + \nu\mu_{t,m}Y_{j,t,m},$$

$$c_t^E + s_t^E = \sum_{j=1}^m ([w_{t,m}L_{j,t,m}^E + r_{t,m}K_{j,t,m}^E] + p_{j,t,m}A_t + (1 - \nu)\mu_{t,m}Y_{j,t,m} - F_{j,t}).$$

In this way, the interest of entrepreneurs to invest in AI infrastructure such that firms operate under imperfect competition would decrease. The reason is that they would benefit less from profits as a share ν would be deducted. We note in our numerical exercise that a profit tax of $\nu = 0.498$ would be necessary to encourage entrepreneurs to invest in AI infrastructure in a way such that the economy coincides with the social planner's optimum.

8.2 Data Sharing and Intellectual Property Rights

Haskel and Westlake (2017) state that due to the transition to an intangible economy, new institutional foundations have to be defined to undermine the lobbying in intangible-intensive industries and to foster market competition. The first steps in this direction have been taken with the Digital Markets Act or AI Act in the European Union and comparable proposals in the United States, such as the American Choice and Innovation Online Act. Governments should establish policies that require data-rich companies to share certain types of data with competitors or third parties under specific conditions. These policies should strike a balance between promoting innovation and protecting privacy.

Governments should evaluate existing patent laws and consider reforms that explicitly address AI-generated inventions. This might involve revisiting the criteria for patentability and redefining concepts like inventorship, ownership, and disclosure requirements to better accommodate non-human labor. Collective ownership or public trusts for the generation of AI could ensure broader access and promote data sharing while still incentivizing innovation. Therefore, an avenue for future research that goes beyond this project might be to not only analyze market barriers due to investments in AI infrastructure, but to discuss the effect of the duration of patents, data sharing and knowledge spillovers on AI development and on income inequality.

8.3 Competition and Merger Law

Income inequality occurs primarily when firms are able to collude such that the AI infrastructure expenditures are chosen that maximise entrepreneurs' income. Brynjolfsson (2022) denotes the development that technological change through AI can disproportionately benefit or harm some groups, even if it is beneficial on average, as "Turing Trap" and speaks in favor of reaping the unprecedented benefits of AI by widely redistributing its economic profits. Thus, we emphasize that workers need to be restored as "stakeholders in collective bargaining and corporate decision-making" (Autor et al., 2020b) so that firms increasingly strive for the interests of the entire workforce instead of solely maximising the profits of the firm-owning entrepreneurs. There is a vicious circle that firms with large market shares can increasingly invest in lobbying to create legislation that enables them to even augment their market power (Eeckhout, 2021). With a rising level of AI, we expect that technological frontiers have increasing *technological* and *political* power. Yet, the extreme outcome in our numerical exercise that the technological frontier is powerful enough to oust all firms from the market has to be interpreted with caution. Nonetheless, potential collusion or cartel agreements for building up large market barriers of the most productive firms can be thwarted by a benevolent and neutral market-observing institution. In particular, there is a need for modernization of competition and anti-trust policies to ensure that companies entering digital markets are not impeded by high market barriers, in order to maintain competitiveness and firms' innovative potential. Gersbach (2017) proposes a tightening of competition law with regard to the tech-industry, ranging from ex-ante regulation of platforms facilitating contestability and data sharing requirements up to the break-up of Tech Giants. Authorities such as the Monopolkommission (2022) already emphasize the importance of limiting the market power that arises in the course of digitalization. We thus point towards governmental interventions for the promotion of competition in industries most susceptible to AI, a redistribution of profits to impede a rise in income inequality in digital economies and a modernization of competition and anti-trust policies.

9 Discussion

In our model, we only regard the effects that AI may substitute for labor leading to reduced employment. Yet, we disregard that with a rising level of AI, firms may have better possibilities to employ more input factors for production, as AI enables better coordination, combination and more efficient use of labor and capital. For example, Black and van Esch (2020) state that AI has affected and will further improve recruiting efficiency, which will give firms the possibility to "more effectively identify, attract, screen, assess, interview, and coordinate with job candidates", whereas Ernst et al. (2019) hint at

the improved matching process of workers with a rising level of AI. More general, Dogan and Birant (2021) provide a literature review on how machine learning applications could improve processes at the production line, human resource organisation or machine and material monitoring due to the increasing availability of manufacturing data. Li et al. (2017) argue that an integrated application of AI in the area of intelligent manufacturing may affect firms' production capacity.

In general, firm-specific investments in AI infrastructure can be interpreted as Research and Development (*R&D*) expenditures that may improve firms' AI productivity $\theta_{j,t}$ or may allow firms to produce goods with higher qualities (Aghion et al., 2019). On the one hand, if the technological frontier invested in *R&D* to expand its productivity advantages, it could easier oust competitors with inferior productivity from the market. In particular, this could be the case if competition policy is weakened (Haskel and Westlake, 2017), such that productive and advantaged firms can remain in powerful market positions. This is the channel that we particularly highlight in our model. Haskel and Westlake (2017) state that productivity differences between firms have risen, in particular in industries where firms employ intangible assets in production. On the other hand, if technological laggards can improve their AI productivity relative to the technological frontier, they may catch up to the technological frontier, leading to increased competition. For example, the European Fund for Strategic Investment aims at supporting start-ups and SMEs to strengthen their competitiveness via the AI/Blockchain Investment Support Programme (European Commission, 2021). In our model, especially large firms persistently increase their AI productivity. We do not allow for leap-frogging in the model and market entry of firms that can catch up to the technological frontier. Therefore, it might be worth investigating how the catching up of technological laggards in digital economies might be affected by the rise of AI.

Another shortcoming of our model is that imperfect competition due to less active firms encloses a higher average productivity of active firms. Nonetheless, Brynjolfsson et al. (2017) state that the average productivity has not improved noticeably in the last decades and our conceptual model has to be interpreted with caution. It is left aside by our model that reduced competitive pressure can also lead to negative incentives in terms of innovation activities, leading to lower productivity.

Furthermore, we assume homogeneous markups and do not consider potential composition or reallocation channels (Markiewicz and Silvestrini, 2021) that may enclose heterogeneous markups. Yet, it can be argued that less productive firms would need to set higher markups than highly productive firms to cover their expenditure for financing AI infrastructure. We disregard that larger firms might be able to choose higher markups, which is the case e.g., using a Klenow-Willis specification as in Edmond et al. (2018). For instance, heterogeneous markups can be obtained if we assumed a framework with price competition, where firms' markup is a decreasing function of the firm's market share

Atkeson and Burstein (2008).

A stable relationship between the labor and capital share in production is a main characteristic of CES models, which is one of Kaldor’s facts about economic growth (Kaldor, 1961). Nonetheless, there is empirical evidence that the assumption of a constant elasticity ϵ between capital K and labor L cannot be taken for granted, especially in advanced economies. For example, Piketty and Zucman (2015) provide historical evidence that the elasticity of substitution ϵ has been increasing over the last centuries. The validity of studies such as Karabarbounis and Neiman (2014) or Acemoglu and Restrepo (2020) analyzing LIS dynamics heavily depends on the elasticity of substitution between capital and labor. In our model, the elasticity of substitution between AI, capital and labor is pivotal for the long-term development of the factor income shares. An appropriate way to generalize our framework might be by incorporating a variable elasticity of substitution (VES) model with time-variant factor income shares to be able to address the question of how a time-variant elasticity of substitution between capital, labor and AI affects the factor income shares in an economy with a rising level of AI. Furthermore, more empirical research should be conducted to determine the elasticity between AI and labor.

In our model, the number of firms is endogenously defined by the investments in AI infrastructure in the first step of our optimization procedure. Moreover, the number of active firms affects the growth rate of AI. Thus, it might be necessary to model the number of firms as a state variable in an inter-temporal analysis when modelling the decisions of the different stakeholding groups. For the sake of computational tractability, we only solve an intra-temporal optimization problem in the first step as we assume that stakeholders in the economy are myopic. Yet, if entrepreneurs had inter-temporal considerations, they would anticipate the effect of a lower number of active firms and a lower total industrial output on the growth rate of AI. Indeed, an inter-temporal analysis with idiosyncratic exit, entry and productivity shocks could yield new insights on firm dynamics in AI-intensive economies but would make the analysis of endogenously determined optimal AI infrastructure investments more challenging. Moreover, it still remains hard to forecast how the future growth trajectory of AI will look like. Thus, further models and experimentation are needed to optimally model the growth pattern of AI in a straightforward and tractable fashion.

10 Conclusion

We model the effects of the incorporation of AI in industrial production on firm competition. The new feature of AI is that it learns by application—in contrast to existing automation and robotisation technologies. Therefore, we define AI as a type of self-learning technology that accumulates over time and can develop autonomously due to its self-learning characteristics. Industrial firms pay variable costs for the acquisition of AI

algorithms and invest in AI infrastructure to be able to integrate AI into production. While investments in AI infrastructure foster the continuous improvement of self-learning AI algorithms, especially trained and tailored for the production of firms with a large market share, they are also the foundation for increasing AI productivity, amplified market concentration, high markups and high profit shares. Depending on whether entrepreneurs, ordinary workers, or a benevolent social planner constitute the stakeholder that decide on the AI infrastructure investments of industrial firms, we examine the evolving economic growth trajectory with a rising level of AI.

We emphasize the findings of Autor et al. (2020b) that rising market dominance accelerates the decline in the labor share that is especially pronounced on imperfectly competitive markets if AI serves as a substitute for labor. Moreover, we show how the rise of AI leads to increasing income inequality between workers and entrepreneurs. Hence, we discuss potential instruments like profit taxes, modernized competition and merger laws, or new data sharing standards for preventing the rise of an increasingly unequal society in an AI-based economy.

References

- Acemoglu, D. (2009). *Introduction to Modern Economic Growth*. Princeton Univ. Press, Princeton, NJ.
- Acemoglu, D. and Restrepo, P. (2018a). Modeling Automation. *AEA Papers and Proceedings*, 108:48–53.
- Acemoglu, D. and Restrepo, P. (2018b). The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review*, 108(6):1488–1542.
- Acemoglu, D. and Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy*, 128(6):2188–2244.
- Acemoglu, D. and Restrepo, P. (2021). Tasks, Automation, and the Rise in US Wage Inequality. *NBER Working Paper*, No. 23928.
- Aghion, P., Antonin, C., Bunel, S., and Jaravel, X. (2020). What Are the Labor and Product Market Effects of Automation? *CEPR Discussion Paper*, No. 14443.
- Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., and Li, H. (2019). A Theory of Falling Growth and Rising Rents. *NBER Working Paper*, No. 26338.
- Aghion, P., Jones, B. F., and Jones, C. I. (2017). Artificial Intelligence and Economic Growth. *NBER Working Paper*, No. 23928.
- Akcigit, U., Chen, M. W., Diez, M. F. J., Duval, M. R. A., Engler, P., Fan, J., Maggi, C., Tavares, M. M. M., Schwarz, M. D. A., Shibata, M. I., et al. (2021). Rising Corporate Market Power: Emerging Policy Issues. *IMF Staff Discussion Note*, SDN/21/01.

- Antras, P. and Helpman, E. (2004). Global Sourcing. *Journal of Political Economy*, 112(3):552–580.
- Arrow, K. J. (1962). The Economic Implications of Learning by Doing. *Review of Economic Studies*, 29(3):155–173.
- Atkeson, A. and Burstein, A. (2008). Pricing-to-market, Trade Costs, and International Relative Prices. *American Economic Review*, 98(5):1998–2031.
- Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020a). The Fall of the Labor Share and the Rise of Superstar Firms. *Quarterly Journal of Economics*, 135(2):645–709.
- Autor, D., Mindell, D. A., and Reynolds, E. B. (2020b). *The Work of the Future: Building Better Jobs in an Age of Intelligent Machines*. MIT Press: Cambridge, MA.
- Babina, T., Fedyk, A., He, A., and Hodson, J. (2021). Artificial Intelligence, Firm Growth, and Product Innovation. *Unpublished Manuscript*.
- Bajgar, M., Criscuolo, C., and Timmis, J. (2021). Intangibles and Industry Concentration. *OECD Science, Technology and Industry Working Papers*, 2021(12).
- Barkai, S. (2020). Declining Labor and Capital Shares. *Journal of Finance*, 75(5):2421–2463.
- Black, J. S. and van Esch, P. (2020). AI-enabled Recruiting: What is it and how should a Manager use it? *Business Horizons*, 63(2):215–226.
- Brynjolfsson, E. (2022). The Turing Trap: The Promise & Peril of Human-like Artificial Intelligence. *Daedalus*, 151(2):272–287.
- Brynjolfsson, E., Rock, D., and Syverson, C. (2017). Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics. *NBER Working Paper*, No. 24001.
- Calligaris, S., Criscuolo, C., and Marcolin, L. (2018). Mark-ups in the Digital Era. *OECD Science, Technology and Industry Working Papers*, 2018(10).
- Calvano, E., Calzolari, G., Denicolo, V., and Pastorello, S. (2020). Artificial Intelligence, Algorithmic Pricing, and Collusion. *American Economic Review*, 110(10):3267–97.
- Cette, G. and Lopez, J. (2012). Ict demand behaviour: An international comparison. *Economics of Innovation and New Technology*, 21(4):397–410.
- Cioffi, R., Travaglioni, M., Piscitelli, G., Petrillo, A., and De Felice, F. (2020). Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions. *Sustainability*, 12(2):492.
- Cockburn, I. M., Henderson, R., and Stern, S. (2019). 4. The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis. In *The Economics of Artificial Intelligence*, pages 115–148. University of Chicago Press: Chicago, IL.
- De Loecker, J., Eeckhout, J., and Mongey, S. (2021). Quantifying Market Power and Business Dynamism in the Macroeconomy. *NBER Working Paper*, No. 28761.

- De Loecker, J., Eeckhout, J., and Unger, G. (2020). The Rise of Market Power and the Macroeconomic Implications. *Quarterly Journal of Economics*, 135(2):561–644.
- De Ridder, M. (2019). Market Power and Innovation in the Intangible Economy. *Cambridge Working Papers in Economics, Faculty of Economics, University of Cambridge*, No. 1931.
- Diez, F., Duval, R., Chen, W., Jones, C., and Villegas-Sanchez, C. (2019). The Rise of Corporate Market Power and its Macroeconomic Effects. *IMF World Economic Outlook - Chapter 2*.
- Diez, M. F. J., Fan, J., and Villegas-Sánchez, C. (2021). Global Declining Competition. *Journal of International Economics*, 132(103492).
- Dixit, A. (1979). A Model of Duopoly Suggesting a Theory of Entry Barriers. *Bell Journal of Economics*, pages 20–32.
- Dogan, A. and Birant, D. (2021). Machine Learning and Data Mining in Manufacturing. *Expert Systems with Applications*, 166:114060.
- Edmond, C., Midrigan, V., and Xu, D. Y. (2018). How Costly are Markups? *NBER Working Paper*, No. 24800.
- Eeckhout, J. (2021). *The Profit Paradox*. Princeton University Press, Princeton, NJ.
- Effenberger, A., Enkelmann, S., Menzel, C., Neumann, D., and Stolle, J. (2020). Marktkonzentration, Produktivität und Preisaufschläge: Deskriptive Evidenz auf Basis amtlicher Daten für Deutschland. *Bundesministerium für Wirtschaft und Energie Discussion Paper*, No. 9.
- Elsby, M. W., Hobijn, B., and Şahin, A. (2013). The Decline of the US Labor Share. *Brookings Papers on Economic Activity*, 2013(2):1–63.
- Ernst, E., Merola, R., and Samaan, D. (2019). Economics of Artificial Intelligence: Implications for the Future of Work. *IZA Journal of Labor Policy*, 9(1).
- European Commission (2021). Fostering a European Approach to Artificial Intelligence. *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions*, 205.
- Falcini, F., Lami, G., and Costanza, A. M. (2017). Deep Learning in Automotive Software. *IEEE Software*, 34(3):56–63.
- Farboodi, M. and Veldkamp, L. (2021). A growth model of the data economy. *NBER Working Paper*, No. 28427.
- Faulhaber, L. V. (2019). Taxing Tech: The Future of Digital Taxation. *Va. Tax Rev.*, 39:145.
- Gersbach, H. (2017). Co-voting Democracy. *Economics of Governance*, 18(4):337–349.
- Gersbach, H., Komarov, E., and von Maydell, R. (2022). Artificial Intelligence as Self-Learning Capital. *CEPR Discussion Paper*, No. 17221.

- Goldin, I., Koutroumpis, P., Lafond, F., and Winkler, J. (2020). Why Is Productivity Slowing Down? *Munich Personal RePEc Archive Working Paper*, No. 107644.
- Grossman, G. M., Helpman, E., Oberfield, E., and Sampson, T. (2017). The Productivity Slowdown and the Declining Labor Share: A Neoclassical Exploration. *NBER Working Paper*, No. 23853.
- Grossman, G. M. and Oberfield, E. (2021). The Elusive Explanation for the Declining Labor Share. *NBER Working Paper*, No. 29165.
- Hanson, R. (2001). Economic Growth Given Machine Intelligence. *Technical Report, University of California, Berkeley, CA*.
- Haskel, J. and Westlake, S. (2017). *Capitalism Without Capital*. Princeton University Press, Princeton, NJ.
- Helpman, E., Melitz, M. J., and Yeaple, S. R. (2004). Export versus FDI with Heterogeneous Firms. *American Economic Review*, 94(1):300–316.
- Hémous, D. and Olsen, M. (2014). The Rise of the Machines: Automation, Horizontal Innovation and Income Inequality. *CEPR Discussion Paper*, No. 10244.
- Höök, M., Li, J., Oba, N., and Snowden, S. (2011). Descriptive and Predictive Growth Curves in Energy System Analysis. *Natural Resources Research*, 20(2):103–116.
- Hopenhayn, H. A. (1992). Entry, Exit, and Firm Dynamics in Long Run Equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1127–1150.
- Hulten, C. R. (2010). Decoding Microsoft: Intangible Capital as a Source of Company Growth. *NBER Working Paper*, No.15799.
- Jaimovich, N. (2007). Firm Dynamics and Markup Variations: Implications for Sunspot Equilibria and Endogenous Economic Fluctuations. *Journal of Economic Theory*, 137(1):300–325.
- Jones, C. I. and Tonetti, C. (2020). Nonrivalry and the Economics of Data. *American Economic Review*, 110(9):2819–58.
- Kaldor, N. (1961). *Capital Accumulation and Economic Growth*. Palgrave Macmillan, London.
- Karabarbounis, L. and Neiman, B. (2014). The Global Decline of the Labor Share. *Quarterly Journal of Economics*, 129(1):61–103.
- King, R. G. and Rebelo, S. (1989). Transitional Dynamics and Economic Growth in the Neoclassical Model. *NBER Working Paper*, No. 3185.
- Klump, R. and Saam, M. (2008). Calibration of Normalised CES Production Functions in Dynamic Models. *Economics Letters*, 99(2):256–259.
- Korinek, A. and Stiglitz, J. E. (2017). Artificial Intelligence and its Implications for Income Distribution and Unemployment. *NBER Working Paper*, No. 24174.

- Kumar, S. (2015). Fundamental Limits to Moore’s Law. *arXiv preprint; arXiv:1511.05956*.
- Kydland, F. E. and Prescott, E. C. (1982). Time to Build and Aggregate Fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370.
- Laatikainen, G. and Ojala, A. (2014). SaaS Architecture and Pricing Models. *IEEE International Conference on Services Computing*, pages 597–604.
- Lagomarsino, E. (2020). Estimating Elasticities of Substitution with Nested CES Production Functions: Where do We Stand? *Energy Economics*, 88:104752.
- Lankisch, C., Prettnner, K., and Prskawetz, A. (2019). How can Robots affect Wage Inequality? *Economic Modelling*, 81:161–169.
- Li, B.-h., Hou, B.-c., Yu, W.-t., Lu, X.-b., and Yang, C.-w. (2017). Applications of Artificial Intelligence in Intelligent Manufacturing: A Review. *Frontiers of Information Technology & Electronic Engineering*, 18(1):86–96.
- Lu, C.-H. (2020). Artificial Intelligence and Human Jobs. *Macroeconomic Dynamics*, 25(5):1–40.
- Makridakis, S. (2017). The Forthcoming Artificial Intelligence (AI) Revolution: Its Impact on Society and Firms. *Futures*, 90:46–60.
- Mankiw, N. G., Romer, D., and Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics*, 107(2):407–437.
- Markiewicz, A. and Silvestrini, R. (2021). Heterogeneous Increase in Sectoral Market Power: the Role of ICT. *Working Paper*.
- Mas-Colell, A., Whinston, M. D., Green, J. R., et al. (1995). *Microeconomic Theory*, volume 1. Oxford University Press, New York (NY).
- Maskin, E. and Tirole, J. (1988). A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs. *Econometrica: Journal of the Econometric Society*, pages 549–569.
- Maydell, R. and Menzel, C. (2023). The Rise of Artificial Intelligence: Towards a Modernisation of Competition Policy. *mimeo*.
- Melitz, M. J. and Redding, S. J. (2013). Firm Heterogeneity and Aggregate Welfare. *CEPR Discussion Paper*, No. DP9405.
- Monopolkommission, W. (2022). XXIV. *Hauptgutachten der Monopolkommission gemäß § 44 Abs. 1 Satz 1 GWB, 5. Juli 2022*, 44.
- Moore, G. E. (1998). Cramming more Components onto Integrated Circuits, Electronics, Vol. 38, No. 8, April 1965. *Proceedings of the IEEE*, 86(1):82–85.
- Nigai, S. (2017). A Tale of Two Tails: Productivity Distribution and the Gains from Trade. *Journal of International Economics*, 104:44–62.

- Noy, S. and Zhang, W. (2023). Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence. *Science*, 381(6654):187–192.
- Osborne, D. K. (1964). The Role of Entry in Oligopoly Theory. *Journal of Political Economy*, 72(4):396–402.
- Patrício, D. I. and Rieder, R. (2018). Computer Vision and Artificial Intelligence in Precision Agriculture for Grain Crops: A Systematic Review. *Computers and Electronics in Agriculture*, 153:69–81.
- Perla, J. and Tonetti, C. (2014). Equilibrium Imitation and Growth. *Journal of Political Economy*, 122(1):52–76.
- Piketty, T. and Zucman, G. (2015). Wealth and Inheritance in the Long Run. In Atkinson, A. B. and Bourguignon, F., editors, *Handbook of Income Distribution*, Handbook of Income Distribution. Elsevier, Amsterdam.
- Prettner, K. (2019). A Note on the Implications of Automation for Economic Growth and the Labor Share. *Macroeconomic Dynamics*, 23(3):1294–1301.
- Raurich, X., Sala, H., and Sorolla, V. (2012). Factor Shares, the Price Markup, and the Elasticity of Substitution Between Capital and Labor. *Journal of Macroeconomics*, 34(1):181–198.
- Romer, P. M. (1986). Increasing Returns and Long-run Growth. *Journal of Political Economy*, 94(5):1002–1037.
- Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5, Part 2):71–102.
- Rubinfeld, D. L. and Gal, M. S. (2017). Access Barriers to Big Data. *Arizona Law Review*, 59:339.
- Schweitzer, H., Metzger, A., Blind, K., Richter, H., Niebel, C., and Gutmann, F. (2022). Data Access and Sharing in Germany and in the EU: Towards a Coherent Legal Framework for the Emerging Data Economy: A Legal, Economic and Competition Policy Angle-Final Report. *Report in behalf of the German Ministry for Economic Affairs and Climate Change*.
- Stigler Committee (2019). Stigler Committee Final Report on Digital Platforms. *The University of Chicago Booth School of Business*.
- Trammell, P. and Korinek, A. (2020). Economic Growth under Transformative AI. *GPI Working Paper*, 2020(8).
- Uzawa, H. (1961). Neutral Inventions and the Stability of Growth Equilibrium. *Review of Economic Studies*, 28(2):117–124.
- Wagner, D. N. (2020). Economic Patterns in a World with Artificial Intelligence. *Evolutionary and Institutional Economics Review*, 17(1):111–131.

Appendix

Markups in the Economy

For conceptualizing markups in a simple way, we consider a production function $Y_t = G(K_t, L, A_t)$ and take into account a single representative firm. Total income accrues as payments to production factors or is composed of profits minus fixed costs (Grossman and Oberfield, 2021). Thus, holding the goods' selling price fixed at the numeraire, Euler's theorem implies the following in a setting with perfect competition on the product market and zero profits⁴⁴

Proposition 4

Suppose that $G(K_t, L, A_t) \rightarrow \mathbb{R}$ is differentiable in $L \in \mathbb{R}_+$, $K_t \in \mathbb{R}_+$ and $A_t \in \mathbb{R}_+$, where the partial derivatives are given by w_t , r_t and p_t , respectively and fixed costs F_t have to be paid for production. Then,

$$Y_t - F_t = G(L, K_t, A_t) - F_t = w_t L + r_t K_t + p_t A_t - F_t.$$

On a product a market with perfect competition, we can thus define the labor, capital and AI share, respectively as

$$\kappa_{L,t} = \frac{w_t L}{Y_t}, \quad \kappa_{K,t} = \frac{r_t K_t}{Y_t}, \quad \kappa_{A,t} = \frac{p_t A_t}{Y_t}.$$

In case of imperfect competition on the product market, firms can sell products to a higher effective price and charge a price markup $\mu_t > 1$. Therefore, we write the net profit of a firm at time t as

$$\Pi_t = (1 + \mu_t)G(K_t, L, A_t) - r_t K_t - w_t L - p_t A_t - F_t.$$

Therefore, we define the labor share under imperfect competition in the following way:

$$\phi_{L,t}^{IC} = \frac{w_t L}{(1 + \mu_t)Y_t}.$$

We see that the higher the markup μ_t the lower the labor share in production. Using analogous definitions for AI and capital, we obtain $\phi_{K,t}^{IC} = \frac{r_t K_t}{(1 + \mu_t)Y_t}$ and $\phi_{A,t}^{IC} = \frac{p_t A_t}{(1 + \mu_t)Y_t}$. Moreover, we are in a position to show that the profit share is defined as

$$\phi_{P,t}^{IC} = \frac{\mu_t Y_t - F_t}{(1 + \mu_t)Y_t},$$

and the share of total income that has to be paid on fixed costs is given by

$$\phi_{F,t}^{IC} = \frac{F_t}{(1 + \mu_t)Y_t}.$$

By applying Proposition 4, we note that

$$\phi_{L,t}^{IC} + \phi_{K,t}^{IC} + \phi_{A,t}^{IC} + \phi_{P,t}^{IC} + \phi_{F,t}^{IC} = \frac{w_t L + r_t K_t + p_t A_t + \mu_t Y_t - F_t + F_t}{(1 + \mu_t)Y_t} = 1,$$

⁴⁴The proof for this this theorem is derived from Euler's homogeneous function theorem.

Proof Proposition 1

We show for two extreme cases, how the productivity distribution could look like to show how the number of active firms and the resulting factor allocation in equilibrium depends on $\Phi(\theta_{j,t})$:

1. All firms have the same AI productivity, such that $\Phi(\theta_{j,t}) = \Phi(\theta_{j+\nu}) \quad \forall \nu \in \mathbb{N}, \forall j \in N$
2. The technological frontier is the only firm that can produce with AI, such that $\Phi(\theta_j) > \Phi(\theta_{1+\nu}) = 0 \quad \forall \nu \in \mathbb{N}$

In the first extreme case, all firms have the same equilibrium production $Y_{j,t,m}^* = Y_{j,t,m}^* \quad \forall j$, as no firm has a comparative advantage compared to the others. Thus, no firm can afford to oust competing firms from the market and all firms are active, enclosing that $m_t = N$. In another extreme case, when $\Phi(\theta_1) - \Phi(\theta_{1+\nu}) > \delta$ such that there is a productivity difference between the technological frontier and all other firms, the technological frontier can afford to oust all firms out of the market. Thus, only one firm—namely the technological frontier—will be active. This stands in line with the findings of Ernst et al. (2019), stating that large productivity differences are indeed a barrier for technological progress leading to a rise in market concentration. Apart from the two extreme cases, where either all or only one firms are active, we conclude that the number of active firms $m \in [1, N]$ depends on the underlying productivity distribution.

Proof Proposition 2

Assume $F_{q,t} \leq F_{w,t}$, for arbitrary firms q and w . It holds that $q < w$ such that q is ranked lower in productivity than w .

Case 1: $F_{q,t} = \mu_{t,N} Y_{N,t,N}$

Given that $F_{q,t} = \mu_{t,N} Y_{N,t,N}$, it necessarily follows that $F_{w,t} > \mu_{t,N} Y_{N,t,N}$. Thus, due to profit maximizing considerations of the firms, only $N - w$ firms will be active on the market.

Case 2: $\mu_{t,N-q+1} Y_{N-q+1,t,N-q+1} \leq F_{q,t} \leq \mu_{t,N-q} Y_{N-q,t,N-q}$

In such a setting, there are a maximum of $N - q$ active firms in the market. Analogously, it holds that there are $N - w$ active firms if the fixed costs are $F_{w,t}$. As $F_{q,t} \leq F_{w,t}$ there are Nw active firms, which is fewer than Nq firms.

Proof Proposition 3

Increasing Number of Active Firms

There are different channels that affect firms' total output depending on the number of active firms in opposite directions

1. If less firms are active, there is less competition for the available rival input factors, leading to more available input factors for each active firm.

2. If less firms are active, only more productive firms are operating, leading to different productivity relations between active firms affecting the labor and capital allocation.

We thus note that

$$\frac{\partial L_{j,t,m}^*}{\partial m_t} \gtrless 0, \quad \frac{\partial K_{j,t,m}^*}{\partial m_t} \gtrless 0.$$

This encloses that

$$\frac{\partial r_{t,m}^*}{\partial m_t} = \frac{\partial}{\partial m_t} \alpha \frac{Y_{j,t,m}^*}{K_{j,t,m}^*} \frac{1}{\epsilon} \gtrless 0,$$

$$\begin{aligned} \frac{\partial w_{t,m}^*}{\partial m_t} &= \frac{\partial}{\partial m_t} (1 - \alpha) \gamma L_{j,t,m}^{*\frac{-1}{\omega}} Y_{j,t,m}^{*\frac{1}{\epsilon}} \left[\left(\gamma L_{j,t,m}^{*\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right)^{\frac{\omega}{\omega-1}} \right]^{\frac{-1}{\epsilon}} \\ &\quad \left(\gamma L_{j,t,m}^{*\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right)^{\frac{1}{\omega-1}} \gtrless 0, \end{aligned}$$

$$\begin{aligned} \frac{\partial p_{j,t,m}^*}{\partial m_t} &= \frac{\partial}{\partial m_t} (1 - \alpha) \theta_{j,t} (1 - \gamma) A_t^{\frac{-1}{\omega}} Y_{j,t,m}^{*\frac{1}{\epsilon}} \left[\left(\gamma L_{j,t,m}^{*\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right)^{\frac{\omega}{\omega-1}} \right]^{\frac{-1}{\epsilon}} \\ &\quad \left(\gamma L_{j,t,m}^{*\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right)^{\frac{1}{\omega-1}} \gtrless 0, \end{aligned}$$

Partial Derivatives of the Input Factors

We observe that due to the nested CES structure of our production function, the cross-derivatives with regard to AI are given by

$$\frac{\partial Y_{j,t,m}}{\partial K_{j,t,m} \partial A_t} = \frac{\partial r_{t,m}}{\partial A_t} = \frac{\partial}{\partial A_t} \alpha \left(\frac{Y_{j,t,m}}{K_{j,t,m}} \right)^{\frac{1}{\epsilon}}. \quad (20)$$

Due to the CES production function, we conclude that $\frac{\partial r_{t,m}^*}{\partial A_t} \geq 0$ irrespective of ϵ .

$$\begin{aligned} \frac{\partial Y_{j,t,m}}{\partial L_{j,t,m} \partial A_t} &= \frac{\partial w_{t,m}}{\partial A_t} = \frac{\partial}{\partial A_t} (1 - \alpha) \gamma L_{j,t,m}^{*\frac{-1}{\omega}} Y_{j,t,m}^{*\frac{1}{\epsilon}} \left[\left(\gamma L_{j,t,m}^{*\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right)^{\frac{\omega}{\omega-1}} \right]^{\frac{-1}{\epsilon}} \\ &\quad \left(\gamma L_{j,t,m}^{*\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right)^{\frac{1}{\omega-1}}. \end{aligned} \quad (21)$$

We note that $\frac{\partial w_{t,m}^*}{\partial A_t} \geq 0$ if $\omega > 1$. The opposite holds if $\omega \leq 1$. Due to the concavity of our nested CES production function, we additionally note that

$$\begin{aligned} \frac{\partial Y_{j,t,m}^2}{\partial A_t^2} &= \frac{\partial p_{j,t,m}}{\partial A_t} = \frac{\partial}{\partial A_t} (1 - \alpha) \theta_{j,t} (1 - \gamma) A_t^{\frac{-1}{\omega}} Y_{j,t,m}^{*\frac{1}{\epsilon}} \left[\left(\gamma L_{j,t,m}^{*\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right)^{\frac{\omega}{\omega-1}} \right]^{\frac{-1}{\epsilon}} \\ &\quad \left(\gamma L_{j,t,m}^{*\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right)^{\frac{1}{\omega-1}} \leq 0. \end{aligned} \quad (22)$$

We can conclude from Eq. 21 and Eq. 20 that the wage increases in the level of AI

if we assume an elasticity of substitution larger than 1 between labor and AI, given by $\omega \geq 1$. The opposite holds if we assume $\omega < 1$. In addition, we note that $\frac{\partial r_{t,m}}{\partial A_t} \geq 0$ if $\epsilon > 1$. The opposite holds if $\epsilon \leq 1$. Thus, the equilibrium allocation of the input factors is affected by the elasticities of substitution. Moreover, due to Proposition 1, the productivity distribution additionally affects the equilibrium factor allocation.

Revenue-maximizing Social Planner: Factor Market Equilibrium

The equilibrium factor allocation can also be obtained by assessing the approach of a revenue-maximizing social planner who maximizes the economy-wide revenue from production. A social planner maximizes the following:

$$\begin{aligned} & \max_{A_t, L_{j,t,m}, K_{j,t,m}} \sum_{j=1}^m G(A_{j,t,m}, L_{j,t,m}, K_{j,t,m}) \\ \text{s.t.} \quad & \sum_{j=1}^m L_{j,t,m} = L_t, \quad \sum_{j=1}^m K_{j,t,m} = K_t \quad \forall t. \end{aligned}$$

This approach resembles a revenue-maximizing problem, where the total production of all firms gets maximized. The equilibrium factor price for each production input must be equal to its aggregate marginal revenue (Mas-Colell et al., 1995). Therefore, the marginal rate of technical substitution (MRTS) of the input factors labor, capital and AI of all active firms have to be equal, given by

$$\begin{aligned} MRTS_{j,t,m}^L &= MRTS_{j,t,m}^L \\ MRTS_{j,t,m}^A &= MRTS_{j,t,m}^A \\ MRTS_{j,t,m}^{L,A} &= MRTS_{j,t,m}^{L,A} \quad \forall j, \forall t. \end{aligned}$$

Growth Path of the Economy

We know from Eq. (1) that the production function has the following form:

$$Y_{j,t} = \left(\alpha K_{j,t}^{\frac{\epsilon-1}{\epsilon}} + (1-\alpha) \left[\left(\gamma L_{j,t}^{\frac{\omega-1}{\omega}} + (1-\gamma)(\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right)^{\frac{\omega}{\omega-1}} \right]^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}.$$

Moreover, we assume that the available labor force remains constant, such that its growth rate is given by $g_L = 0$. Moreover, due to the assumption of bounded growth of AI, we see that $\lim_{t \rightarrow \infty} A_t = B$, such that the growth rate of AI in the long-run is given by $g_A = 0$. In addition, we know from Eq. (5) that the fixed costs have to comove with the level of production if the markups stay constant in the long-run and only depend on the number of active firms m_t if AI has reached its upper boundary. Thus, we note that the growth rate of fixed costs spent, given by g_F has to be proportional to the production of each firm g_Y , such that $g_F \propto g_Y$ in the long-run. As a result, as $g_A = g_L = g_\mu = 0$ in the long-run, we note that the capital growth rate, given by g_K is the decisive for determining the growth rate of production, as defined by Eq. (13). For any $K_t > 0$, we have $\frac{\partial K_t}{\partial t} > 0$ and therefore $\lim_{t \rightarrow \infty} K_t = \infty$. Due to Uzawa (1961)'s theorem, we can additionally conclude that $g_K = g_Y$. Furthermore, due to the constant savings rate and

the relation between consumption and total production, given by Eq. (19), we observe that $g_C = g_Y$.

Utility Maximization with Elastic Labor Supply: Perfect Competition

In this subsection, we set up the maximization problem for an economy with perfect competition, zero markups, a representative agent and no fixed costs. The first order conditions of the maximization problem described are given by:

$$\begin{aligned} u_c(c_t, L_t) &= \beta \left[f_K \left(K_{t+1}, L_{t+1}, \hat{A} \right) + (1 - \delta) \right] u_c(c_{t+1}, L_{t+1}), \\ u_l(c_t, L_t) &= -u_c(c_t, L_t) f_l \left(K_t, L_t, \hat{A} \right). \end{aligned}$$

This implies for the steady state that

$$\begin{aligned} f_k \left(k_{ss}, L_{ss}, \hat{A} \right) &= 1/\beta - 1 + \delta \\ u_l(c_{ss}, L_{ss}) &= -u_c(c_{ss}, l_{ss}) f_l \left(k_{ss}, L_{ss}, \hat{A} \right). \end{aligned}$$

Combining these results with our production function, defined in Eq. 1, we obtain that

$$\begin{aligned} \alpha \frac{K_t^{-\frac{1}{\epsilon}}}{Y_t} &= \frac{1}{\beta} - 1 + \delta, \\ \frac{\gamma}{1 - L_{ss}} &= \frac{1}{c_{ss}} (1 - \alpha) \gamma L_{ss}^{-\frac{1}{\omega}} Y_t^{\frac{1}{\epsilon}} \left[\left(\gamma L_{ss}^{\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} \hat{A})^{\frac{\omega-1}{\omega}} \right)^{\frac{\omega}{\omega-1}} \right]^{-\frac{1}{\epsilon}} \\ &\quad \left(\gamma L_{ss}^{\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} \hat{A})^{\frac{\omega-1}{\omega}} \right)^{\frac{1}{\omega-1}}. \end{aligned}$$

Consequently it has to hold for the minimal L_t^{min} to have $c_t >$ that

$$\left(\alpha K_t^{\frac{\epsilon-1}{\epsilon}} + (1 - \alpha) \left[\left(\gamma L_t^{\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right)^{\frac{\omega}{\omega-1}} \right]^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}} + (1 - \delta) K_t - K_{t+1} > 0.$$

This implies that

$$L_t^{min} > \frac{1}{\gamma} \left\{ \left[\frac{\left((K_{t+1} - (1 - \delta) K_t)^{\frac{\epsilon}{\epsilon-1}} - \alpha K_t^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}}{1 - \alpha} \right]^{\frac{\omega-1}{\omega}} - (1 - \gamma) (\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right\}^{\frac{\omega}{\omega-1}}.$$

The recursive optimization problem of the agents is given by the following equation Assuming a lower and upper bound for c_t and L_t enclosed the feasibility constraint that

$$\Gamma(K) = [0, f(K, 1, \hat{A}) + (1 - \delta)k].$$

Finally, we can set up a recursive formulation of the agents' problem that do not anticipate the growth of AI

$$V(K) = \max_{K' \in \Gamma(K)} F(K, K') + \beta V(K'),$$

where $F(K, K')$ is the value of the recursive maximization procedure

$$\begin{aligned} F(K, K') &= \max_{c, L} u(c, L) \\ \text{s.t. } c + K' &\leq f(K, L, \hat{A}) + (1 - \delta)K \text{ and } c \geq 0, 1 \geq L \geq 0. \end{aligned}$$

Utility Maximization with Elastic Labor Supply: Imperfect Competition

In this subsection, we set up the maximization problem for a representative agent in an environment with imperfect competition, non-zero markups and positive fixed costs. The first order conditions of the maximization problem described are given by:

$$\begin{aligned} u_c(c_t, L_t) &= \beta \left[(1 + \mu_t) f_K(K_{t+1}, L_{t+1}, \hat{A}) + (1 - \delta) \right] u_c(c_{t+1}, L_{t+1}), \\ u_l(c_t, L_t) &= -u_c(c_t, L_t) (1 + \mu_t) f_l(K_t^E, L_t^E, \hat{A}). \end{aligned}$$

This implies for the steady state that

$$\begin{aligned} (1 + \mu_t) f_K(k_{ss}, L_{ss}, \hat{A}) &= 1/\beta - 1 + \delta \\ u_l(c_{ss}, L_{ss}) &= -u_c(c_{ss}, l_{ss}) (1 + \mu_t) f_l(k_{ss}, L_{ss}, \hat{A}). \end{aligned}$$

Combining these results with our production function, defined in Eq. 1, we obtain that

$$\begin{aligned} (1 + \mu_t) \alpha \frac{K_t^{\frac{-1}{\epsilon}}}{Y_t} &= \frac{1}{\beta} - 1 + \delta, \\ \frac{\gamma}{L^{max} - L_{ss}} &= \frac{1}{c_{ss}} (1 + \mu_t) (1 - \alpha) \gamma L_{ss}^{\frac{-1}{\omega}} Y_t^{\frac{1}{\epsilon}} \left[\left(\gamma L_{ss}^{\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} \hat{A})^{\frac{\omega-1}{\omega}} \right)^{\frac{\omega}{\omega-1}} \right]^{\frac{-1}{\epsilon}} \\ &\quad \left(\gamma L_{ss}^{\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} \hat{A})^{\frac{\omega-1}{\omega}} \right)^{\frac{1}{\omega-1}}. \end{aligned}$$

Consequently it has to hold for the minimal L_t^{min} to have $c_t >$ that

$$\begin{aligned} (1 + \mu_t) &\left(\alpha K_t^{\frac{\epsilon-1}{\epsilon}} + (1 - \alpha) \left[\left(\gamma L_t^{\frac{\omega-1}{\omega}} + (1 - \gamma) (\theta_{j,t} A_t)^{\frac{\omega-1}{\omega}} \right)^{\frac{\omega}{\omega-1}} \right]^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}} \\ &- F_t + (1 - \delta) K_t - K_{t+1} > 0. \end{aligned}$$

This implies that

$$L_t^{min} > \frac{1}{\gamma} \left\{ \left[\frac{\left(\left(\frac{K_{t+1} - (1-\delta)K_t + F_t}{(1+\mu_t)} \right)^{\frac{\epsilon-1}{\epsilon}} - \alpha K_t^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}}{1-\alpha} \right]^{\frac{\omega-1}{\omega}} - (1-\gamma)(\theta_{j,t}A_t)^{\frac{\omega-1}{\omega}} \right\}^{\frac{\omega}{\omega-1}} .$$

Additional Figure on the Factor Income Shares

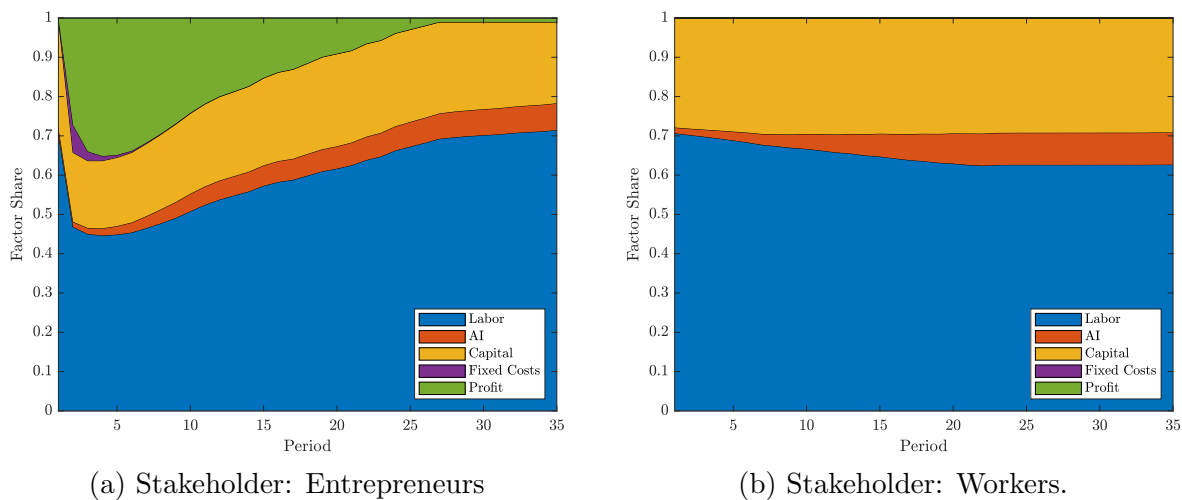


Figure 9: Development of the Factor Income Shares with $\epsilon = 0.85$ and $\omega = 1.25$.