Artificial Intelligence and the Rents of Finance Workers*

Jean-Edouard Colliard† and Junli Zhao‡

June 27, 2025

Abstract

This paper studies how artificial intelligence (AI) affects the finance labor market when humans and AI perform different tasks in investment projects, and workers earn agency rents that grow with project size. We identify two key effects of AI improvement: A free-riding effect raises worker rents by increasing the probability of successful investment when the worker shirks; A capital reallocation effect shifts investment toward workers with higher or lower rents, depending on which tasks AI improves. Contrary to standard predictions, AI can raise both worker rents and labor demand. We derive implications for capital allocation, labor demand, compensation, and welfare.

Keywords: Artificial intelligence, labor market, automation, rents in finance.

JEL classification: D21, G20, O33.

*We are grateful to seminar participants at INSEAD for helpful comments and feedback. Part of this research was conducted while Colliard was visiting INSEAD and Sciences Po. Their hospitality is gratefully acknowledged.

†HEC Paris, Department of Finance, 1 rue de la Libération, 78351 Jouy-en-Josas, France. Email: colliard@hec.fr

‡Bayes Business School, City University of London, 106 Bunhill Row, London, EC1Y 8TZ, United Kingdom. Email: junli.zhao@city.ac.uk

Introduction

Since the public release of ChatGPT in November 2022, the number of tasks that artificial intelligence (AI) can complete as well as high-skilled humans has grown rapidly. Finance features spectacular examples of AI capabilities, including forecasting earnings on par with human analysts (Cao *et al.*, 2024), predicting stock price movements (Lopez-Lira and Tang, 2024), and even making corporate finance decisions (Campello *et al.*, 2025). These are only the latest developments in a longer trend towards automation in finance, which previously relied on large datasets and machine learning.[1]

An extent literature in economics has relied on a "task-based approach" to study the impact of automation on the labor market (e.g., Acemoglu and Restrepo (2018b)). Under this approach, production requires the completion of different tasks, allocated to human "workers" or to "machines". In a competitive economy, workers' wages and jobs are hurt by improvements in the machine if they work on tasks that are substitutes to the machine. Conversely, wages and jobs increase for complementary occupations.[2] Under this approach, the future looks bleak for finance workers: a recent report by Accenture ranks banking as the sector of the economy as the sector with the highest share of working hours with "higher potential for automation" (54% of hours), followed by insurance second (48%) and capital markets fourth (40%).[3]

However, a competitive model may not be sufficient to explain the outcomes on the finance labor market. Indeed, Philippon and Reshef (2012) document a premium for finance jobs after 1990 that cannot be explained by standard factors. This unexplained component is likely an agency rent, due to the widespread use of incentive pay in finance. Célérier and Vallée (2019) confirm this hypothesis, and also point out the role of high complementarity

[1]See, e.g., Abis and Veldkamp (2023) for the case of asset management, and Bonelli (2025) for venture capital. For an industry view, see for instance "We have 'surrendered more to the machines', says quant fund titan Cliff Asness", *Financial Times*, 4 June 2025.

[2]Some substitution may have already occurred during the "big data" wave of the 2010s. Abis and Veldkamp (2023) show that over 2010-2018 the production function of the asset management industry has changed in a way that requires fewer workers to analyze a given stock of data, which reduced the labor share of income by 5%.

[3]Bick *et al.* (2024) report that the use of AI at work was already widespread in the U.S. financial sector in August 2024, with 38.1% of respondents in "Business / Finance" occupations reporting using AI at least once a week.

between talent and project size as an explanation of the finance wage premium.

In this paper, we ask how the finance labor market is impacted by the development of "machines" that are substitutes to workers, taking into account that workers are paid for performance, and that their rents increase with project size.

To answer this question, we consider a model in which investors optimally allocate capital to heterogeneous "financiers" (e.g., banks or active mutual funds) and to a risk-free asset. Each financier has access to an investment project, with a marginal cost increasing in project size. The success of the project depends on the completion of a continuum of tasks, which can be performed by either a worker or a machine. The worker and the machine are perfectly substitutable on a given task, and tasks are perfectly substitutable to each other. The machine is free to use, while the worker is subject to moral hazard. The worker is more productive than the machine at least on some "high-level" tasks. We interpret AI as a machine that is also productive on many high-level tasks. The financier has to decide which tasks to allocate to the machine and to the worker, and what contract she offers to the worker. The market for capital is competitive, so that financiers need to maximize the expected return they offer to investors.

We obtain that machines impact the labor market very differently from what would happen in a competitive model. Two important effects are at play. First, as long as the financier decides to use the worker, machine improvement has a "free-riding effect": the worker is paid conditionally on success, and a better machine increases the probability of success even when the worker shirks. This leads to higher rents for the worker. Second, machine improvement leads to a reallocation of capital across financiers, a "capital reallocation effect". This effect can be positive or negative, depending on which tasks are affected by machine improvement.

Overall, we find that under many configurations these two effects lead to larger total rents for workers in the financial sector (sometimes, shared across fewer workers). In fact, there are only two cases when machine improvement can be negative for workers: (i) if the machine improves so much that it is no longer necessary to incentivize workers; (ii) if machine improvement disproportionately affects tasks already allocated to machines, which, according to our interpretation, is not the case of generative AI for instance. Thus, our conclusion is that current developments in AI in finance may in large part be appropriated by workers, contrary to what a standard competitive model would predict. More broadly, our results suggest that empirical research on the consequences of automation on the labor market should look into the moral hazard component of occupations in addition to the tasks involved.[4]

Our model delivers several additional insights and predictions.

First, if a financier uses a worker in equilibrium, then too many tasks are allocated to the worker relative to a benchmark without moral hazard. Indeed, the worker earns a rent based on the success of the project, and is not rewarded for the marginal disutility of work. Hence, allocating a marginal task to the worker is free, from the financier's perspective. This result implies that the machine is under-used at the intensive margin.[5] Hence, it is possible that the potential for automating finance tasks will not be fully realized, because the efficiency gains would be captured by workers.[6]

Second, if capital is scarce in the economy, so that investors allocate all of it to the financiers, the result above implies that total labor demand is higher in equilibrium than in the absence of moral hazard. Hence, finance workers both earn more rents and are more numerous, due to the presence of moral hazard. If capital is abundant instead, investors allocate some of their capital to the risk-free asset. Without moral hazard the financiers would offer higher returns, attract more capital, and this would create a need for more workers.

Third, we can study different forms of machine improvement, which will have different consequences via the reallocation effect. More specifically, our model assumes two types of financiers: high types financiers have access to a more productive worker, and low type financiers to a less productive worker. Low type financiers optimally use the machine more than high types. An improvement in the machine will thus tend to benefit the low types more. However, there is a second effect: because project failure is a less informative signal of shirking for low type workers, they actually earn higher rents. When the machine improves, such

[4]This could be done across industries/plants using measures of management practices (Bloom *et al.*, 2019). In the case of finance, Efing *et al.* (2022) have employee-level data on bonuses in banks.

[5]There is a less surprising effect at the extensive margin: agency rents make the worker more expensive, and if they are too high the financier may decide to not rely on the worker at all. The machine is then over-used.

[6]In fact, a hypothesis for why the potential for automation is high in finance is precisely that the presence of rents has slowed down the adoption of previous waves of automation.

workers capture a higher fraction of the surplus created by machine improvement. Depending on the balance between these two effects, machine improvement may have a stronger impact on the performance of high type financiers, or low type financiers. This impact in turn will lead investors to reallocate capital from one type to the other.

To understand the balance between these two effects, we consider two special cases of machine improvement:

In the case of a "core" improvement, only tasks that both types of financiers allocate to the machine improve. An example would be advances in data analysis. Then the success probability increases by the same amount for both types of financiers, but high types leave a lower fraction of the surplus as rents to the worker. As a result, capital is reallocated towards high type financiers. As high type financiers use more labor and have lower rents, the capital reallocation effect is positive for labor and negative for rents.

In the case of a "frontier" improvement, only tasks that low type financiers allocate to the machine improve. An example would be advances in the use of AI for tasks usually allocated to humans. Then the success probability increases only for low type financiers, who attract more capital. The capital reallocation effect is then negative for labor but positive for rents.

Finally, we compute total welfare in our model as the sum of payoffs to investors, financiers, and workers. Despite machine improvement expanding the production set, this is not necessarily welfare improving in the presence of rents. We show in particular that welfare decreases if a machine improvement leads some financiers to stop using the worker. Indeed, if the machine improves sufficiently, a financier can offer a better return to investors by relying only on the machine and reducing the worker's rents to zero. However, the total return of the financier's project, which includes the part appropriated by the worker, inefficiently decreases. In fact, in the absence of moral hazard, it is never optimal to use the machine only.

The last section of the paper develops two extensions.

First, we modify the production function of the financier so that the machine and the worker are perfect complements on every task. We show that the free-riding effect still obtains in that setup, implying that substitutability is not a necessary condition for machine improvement to increase workers' rents. However, we also show that complementarity makes rents lower, which diminishes the impact of machine improvement. Indeed, agency rents derive from the worker being able to shirk and still succeed with a high probability. Under complementarity, not exerting effort also decreases the productivity of the machine, which decreases the success probability more than under subtitutability. This suggests that machine improvement can increase workers' rents more under substitutability than under complementarity, the opposite of what happens in the absence of moral hazard.

Second, we consider an extension in which the allocation of tasks is controlled by the worker, and is not observable by the financier. This would correspond for instance to a covert use of generative AI by employees. In this case, workers earn even higher rents, because they need to be incentivized to work on tasks instead of only relying on the machine. We show that the free-riding effect is still present, so that machine improvement can still increase workers' rents. However, there is also a second effect: as the machine improves, it is in the financier's interest that the machine is used on more tasks. Hence, it is less necessary to disincentivize the worker's covert use of the machine, so that his compensation and rents decrease. This result suggests that empirical research should ideally separate tasks and occupations where the use of machines is controlled by the firm from cases where it is not.

The remainder of this paper is organized as follows. Section 1 positions our paper in the literature. Section 2 describes the model, which we then solve in Section 3, both in the baseline version and in a benchmark version without moral hazard. Section 4 compares these two versions to identify the impact of moral hazard and Section 5 derives comparative statics results on machine improvement. Section 6 develops the two extensions mentioned above. Section 7 concludes. The figures and proofs omitted from the main text are in the appendix.

1 Literature Review

Our paper contributes to the theoretical literature on the determinants of pay in the financial sector, more specifically models relating the finance wage premium to the severity of moral hazard problems in finance (in particular Axelson and Bond (2015), Biais and Landier (2020), and Hoffmann *et al.* (2022)). Other explanations for high remunerations include the screening of the most talented workers (Bond and Glode, 2014), and excessive competition to attract talent into an industry with negative externalities (Thanassoulis, 2012) or fixed-sum surplus (Glode and Lowery, 2016).

We are aware of only two papers in this literature that consider the impact of technology. In Glode and Ordoñez (2024), firms can use technology either to create more surplus or to appropriate the surplus created by other firms. Surprisingly, as technology improves, more resources will (inefficiently) be allocated to appropriation, regardless of the type of improvement. We share the result that even a surplus-improving technology can increase rents, but our mechanism is different and gives predictions on the impact of technological change within a given activity. Tena (2020) studies a dynamic model of moral hazard and shows that the mere anticipation of future automation forces the principal to give high rents to the agent ex ante. These rents then decrease when automation becomes feasible. Our model does not have these interesting dynamics, but it has several tasks and firms. In particular, we show that partial automation can increase rents ex post, which is not possible in a single-task model.

At a broader level, our paper contributes to a growing literature that studies the automation of tasks and its impact on labor market outcomes. We refer the reader to, e.g., Autor (2015), Acemoglu and Restrepo (2018a), and Agrawal *et al.* (2019) for overviews of this literature. A common theme is that machines can substitute humans on some tasks, but humans then become more productive in complementary tasks.[7] Assuming standard competitive markets for goods and labor, whether workers benefit from automation depends on the relative strengths of the substitution and complementarity effects.[8] Our contribution to this literature is to integrate moral hazard and capital allocation in an otherwise simple task-based model of automation.[9] We obtain very different predictions for the impact of ma-

[7]Brynjolfsson *et al.* (2018) measure the task content of occupations and point out that few occupations only consist in tasks suitable to machine learning.

[8]Empirically, the evidence is mixed. Gregory *et al.* (2021), Babina *et al.* (2024), and Aghion *et al.* (2025) report evidence consistent with a positive impact of machines or specifically AI on aggregate labor demand. However, Acemoglu *et al.* (2022), Kogan *et al.* (2023), Bonfiglioli *et al.* (2024), or Eisfeldt *et al.* (2025) also provide evidence of a strong substitution effect. In line with our point that a competitive framework may not tell the whole story, Jiang *et al.* (2025) show empirically that even in the case of complementarity, AI does not necessarily benefit workers, who work longer and are not compensated in proportion. Part of the explanation is market power on the firm's side, and the possibility to use AI to monitor workers and reduce their agency rents.

[9]Acemoglu (2021) allows for a simpler form of worker rents in a task framework by assuming the presence

chine improvement, suggesting that it may be worthwhile for empirical research to separate occupations not only according to their task content, but also to the extent of moral hazard they are exposed to.

Our paper also relates to a growing literature on how technological advancements affect the asset management industry. Farboodi and Veldkamp (2020) explore how new technologies shift asset managers' effort allocation between acquiring information about fundamental values and anticipating others' demand. Dugast and Foucault (2025) and Bonelli and Foucault (2025) examine how big data influences quantitative and discretionary managers disproportionately. Sheng *et al.* (2024) show that hedge funds have started adopting generative AI, with a positive impact on their performance.[10] Zhang (2024) shows that AI adoption can help fund managers identify investment opportunities. Much of this literature either abstracts from agency considerations, which are central to our paper, or focuses on specialized technologies (e.g., alternative data) accessible only to a subset of managers.

A growing literature documents the impact of AI and more broadly technological change on other segments of the financial industry.[11] Bertomeu *et al.* (2025) provide evidence that financial analysts use ChatGPT, which makes them more accurate. Grennan and Michaely (2020) show that, several years before, analysts were already displaced by AI, with some exiting the industry and others focusing on different tasks. Dessaint *et al.* (2024) find evidence that analysts allocate more effort to tasks where they are more complementary with new technology. Jiang *et al.* (Forthcoming) provide evidence of reallocation for finance jobs most exposed to innovations in the FinTech sector. Finally, Eisfeldt and Schubert (2025) systematically measure the exposure of all finance occupations to generative AI. Our model offers new predictions, more specific to finance, that can be tested in this literature. In particular, an important prediction is that workers should earn higher rents when technology improves, provided that they are not entirely displaced by AI.

Finally, while the main application of our model is the financial industry, our insights can be applied to other sectors and more broadly to any employees or managers who earn

of a binding minimum wage.

[10] Interestingly, non hedge funds are not impacted and the disparity in returns between the two groups increases. In our model this would lead to a capital reallocation effect.

[11] Central banks and financial regulators are impacted too, see Aldasoro *et al.* (2024).

rents due to moral hazard. This relates our paper to a broader literature in economics and management on how to organize the collaboration between humans and machines. Boyacı *et al.* (2024) and de Véricourt and Gurkan (Forthcoming) for example explore theoretically how a human decision-maker can best exploit a machine on a prediction task. This literature in general does not consider moral hazard frictions. An important exception is Athey *et al.* (2020), who study a delegation game in which a principal can delegate a decision to either a human agent or an AI, but cannot use transfers to align incentives with the agent. A better machine increases the agent's rent and is thus not always preferred by the principal.[12] With transfers and heterogeneous agents, we show that the principal always benefits from machine improvement, while total rents may decline due to capital reallocation.

Several papers in this literature use experiments (conducted either in the laboratory or in actual firms) to study the impact of machines on employee performance and the potential obstacles to adoption (e.g., Fugener *et al.* (2020), Allen and Choudhury (2022), Wang *et al.* (2024), Brynjolfsson *et al.* (2025)). A recurrent empirical result is that some employees, in particular more senior ones, are more skeptical of the benefits of AI and more reluctant to use it. Our model suggests a rational explanation based on the concern of losing existing rents.[13] Organizing work and compensation in a way that alleviates such concerns could then lead to quicker AI adoption.

[12]Zhong (2025) extends the problem of Athey *et al.* (2020) to a multi-stage decision problem, focusing on the optimal design of the decision process that integrates humans and AI to minimize decision errors. In a competitive framework with hierarchical labor structures, Ide and Talamas (2024) and Ide and Talamas (Forthcoming) show that, due to the complementarity between the knowledge of solvers (more knowledgeable) and workers (less knowledgeable), AI and its improvement with different capabilities have different implications for labor outcomes. In contrast, the heterogeneous effect results in our model arises from differences in informational rents, a feature common in many knowledge-intensive occupations. Also but less related is Alonso and Câmara (2024), who consider the problem of a firm having to take a decision based on data analysis, when the manager in charge of the analysis has the possibility to tamper with the results.

[13]An interesting paper in this direction is Ye (2025), which shows that AI assistance has made it more difficult for programmers to signal their skills by sharing their code, prompting programmers to reallocate their effort towards other tasks with a higher signaling content.

2 Setup

2.1 Economic environment

We consider a two-period economy with a risk-free asset, risky projects, a continuum of investors, and a continuum of financiers who can invest in the projects on the investors' behalf.

Investors have one unit of capital each, and form a continuum of total mass K, the total capital in this economy. At $t = 1$, each investor can either invest his capital in a risk-free asset or allocate it to a financier, according to a contract defined below. The risk-free asset pays a net return normalized to zero in $t = 2$.

There is a continuum $[0, N_L + N_H]$ of financiers. The total capital allocated to financier n is denoted k_n. Each financier invests her capital k_n in a project at $t = 1$, and incurs some management costs $k_n\chi(k_n)$, with $\chi'(k_n) \geq 0$, $\chi(0) = 0$, and $\lim_{k_n \to +\infty} \chi(k_n) = +\infty$. Thus, the total management costs are an increasing and convex function of k_n. The project can either succeed or fail in $t = 2$. A successful investment brings a gross payoff of $R > 1$ per unit of capital invested, whereas a failed investment brings a gross payoff of 0.

The probability that the investment is successful depends on the completion of a continuum $x \in [0, 1]$ of tasks. The weight of task x is $f(x)$, where f is an integrable function, $F(x) = \int_0^x f(z)dz$, and $F(1) = 1$.

Each task can be performed by a human (the "worker") working for the financier, or by a machine, where $a_n(x)$ denotes the fraction of the task allocated the machine, $h_n(x)$ the fraction allocated to the worker, and $a_n(x) + h_n(x) \leq 1$. The productivity of the machine on task x is $q_M(x)$. The worker associated with financier n has a productivity equal to $e_n q_n$, where $e_n \in \{0, 1\}$ is the financier's effort. The worker associated with financier n has $q_n = q_L$ if $n \in [0, N_L]$ ("type L" financiers and workers), and $q_n = q_H$ if $n \in [N_L, N_L + N_H]$ ("type H" financiers and workers). We assume $0 < q_L < q_H < 1$.[14]

These ingredients determine the probability $y_n(e_n)$ that financier n's investment is suc-

[14] Our modeling of the financial sector is close to Bond and Glode (2014), with the addition of tasks and a machine, but the limitation that workers' allocation to financiers is fixed.

cessful, with:

$$y_n(a_n, h_n, e_n) = \int_0^1 \left[a_n(x)q_M(x) + h_n(x)e_n q_n\right] f(x)dx. \tag{1}$$

Without loss of generality we assume that $q'_M(x) \leq 0$: tasks are ranked from more to less machine-friendly. In particular, we think of AI as a machine such that q_M decreases slowly and is thus high for many tasks. In contrast, the worker is equally good at all tasks. Moreover, we assume $q_M(0) > q_H$ (the machine is more productive than the worker for the first task), and $q_M(1) = 0$ (the machine is useless on the last task). Finally, we denote $Q_M(\bar{x})$ the average productivity of the machine for tasks below $\bar{x}$:

$$Q_M(\bar{x}) = \int_0^{\bar{x}} \frac{q_M(x)f(x)}{F(\bar{x})}dx. \tag{2}$$

The success probability cannot be larger than 1. Moreover, we assume that the financier could offer a positive net return equal to $RQ_M(1) - 1$ by using the machine only:

$$\frac{1}{R} < Q_M(1) < 1. \tag{3}$$

Exerting effort costs $C(e_n, h_n, k_n) = ck_n e_n \int_0^1 h_n(x)f(x)dx$ to the worker: it is proportional to the number of tasks allocated to him and to the total capital allocated. The cost of using the machine is assumed to be zero, or negligible relative to the cost of effort. We assume that effort is economically efficient:

$$c < Rq_L. \tag{4}$$

All agents in the economy are risk-neutral. Each investor wants to maximize the expected return on his savings, and each financier wants to maximize the total profit of her project, net of the worker's compensation and the return offered to investors. A financier can choose to not receive any capital and obtain a utility of zero. Similarly, a worker can choose to not work for the financier and obtain a utility of zero.

2.2 Contracting environment

Financiers competitively offer contracts to investors. An investor who accepts a contract allocates his unit of capital to the financier, the financier invests this capital into her project, and will then return the proceeds of the investment to the investor, minus a compensation for the worker and a fee for the financier.

We assume that the type of each financier is known to investors. Moreover, in the baseline model the allocation of tasks (the functions a_n and h_n) is observable too. However, the worker's effort is not, which creates a moral hazard or agency problem, in which the financier is the principal and the worker is the agent.

As contracts are offered competitively, in equilibrium they need to maximize the return to the investor (under constraints) and the financiers' fees are zero. As the worker is risk-neutral, standard arguments imply that to solve the moral hazard problem it is sufficient to look at contracts specifying a payment $w_n k_n$ to the worker if the project is successful, and the worker receives no payment in case of failure.

Thus, the contract offered by financier n is fully characterized by $\mathcal{C}_n = \{q_n, a_n, h_n, w_n\}$. The investor accepting this contract knows that the financier is of type q_n, allocates a fraction $a_n(x)$ of task x to the machine and $h_n(x)$ to the worker, and the worker will receive a compensation w_n per unit of capital in case the project is successful.

2.3 Equilibrium definition

An equilibrium of the model is fully characterized by the contracts $\mathcal{C}_n = \{q_n, a_n, h_n, w_n\}$ offered by each financier and the capital k_n allocated to her by the investors, for $n \in [0, N_L + N_H]$. The contracts and the capital allocation form an equilibrium if and only if the following conditions are satisfied: (i) incentive compatibility constraints of workers; (ii) participation constraints of workers; (iii) each contract maximizes the return to investors; (iv) investors' capital allocation is privately optimal.

(i) If the worker exerts effort, he suffers a disutility $C(1, h_n, k_n)$ and receives $w_n k_n$ with probability $y_n(a_n, h_n, 1)$, whereas if he does not he suffers no disutility but receives $w_n k_n$ with probability $y_n(a_n, h_n, 0)$ only. This gives us the incentive compatibility constraint:

$$e_n = 1 \text{ only if } k_n y_n(a_n, h_n, 1)w_n - C(1, h_n, k_n) \geq k_n y_n(a_n, h_n, 0)w_n. \tag{IC}$$

(ii) If the worker exerts effort, the contract has to give him a positive utility. This gives the following participation constraint:

$$e_n = 1 \text{ only if } k_n y_n(a_n, h_n, 1)w_n - C(1, h_n, k_n) \geq 0. \tag{PC}$$

(iii) The return per unit of capital allocated to financier n writes as:

$$\alpha_n(k_n, a, h, e, w) = r_n(a, h, e, w) - \chi(k_n), \tag{5}$$

$$\text{with } r_n(a, h, e, w) = y_n(a, h, e)(R - w) - 1. \tag{6}$$

Note that we decompose the total return to the investor α_n as a return gross of management costs r_n, minus management costs $\chi(k_n)$. This decomposition will be useful below.

The contract offered by the financier needs to maximize α_n, which gives:

$$(a_n, h_n, e_n, w_n) \in arg \max_{a,h,e,w} \alpha_n(k_n, a, h, e, w) \text{ s.t. } (IC), (PC), \forall x \in [0, 1],\ a(x) + h(x) \leq 1. \tag{MAX}$$

(iv) Denote $\bar{\alpha} = \max_n \alpha_n(k_n, a_n, h_n, e_n, w_n)$ the maximum return offered across all financiers. Optimal allocation by the investors implies that all financiers receiving a positive amount of capital must offer the same return. Moreover, if investors allocate some capital to the risk-free asset in equilibrium, then this return has to be zero. This gives us the following optimal allocation conditions:

$$\forall n \in [0, N_L + N_H], k_n > 0 \text{ only if } \quad \alpha_n(k_n, a_n, h_n, e_n, w_n) = \bar{\alpha},$$

$$\text{and } k_n = 0 \text{ only if } \quad \alpha_n(k_n, a_n, h_n, e_n, w_n) \leq \bar{\alpha}, \tag{OA1}$$

$$\bar{\alpha} > 0 \text{ only if } \quad \int_0^{N_L+N_H} k_n dn = K. \tag{OA2}$$

3 Equilibrium

3.1 Complete information benchmark

As a benchmark, we first solve the model assuming that the worker's effort is observable, so that there is no moral hazard and we remove the constraint (IC). We denote $\mathcal{C}_n^* = \{q_n, a_n^*, h_n^*, w_n^*\}$ the equilibrium contracts in this benchmark model, and k_n^* the equilibrium capital allocation.

We conjecture that any financier n with $k_n > 0$ will choose $e_n = 1$, so that the worker exerts effort and the participation constraint (PC) will be binding. We can then solve for the worker's compensation as:

$$w_n^* = \frac{c \int_0^1 h_n^*(x) f(x) dx}{y_n(a_n^*, h_n^*, 1)}. \tag{7}$$

Financier n will then maximize the return for investors, and (MAX) becomes:

$$(a_n^*, h_n^*) \in arg \max_{a,h} R \int_0^1 [a(x) q_M(x) + h(x) q_n] f(x) dx - c \int_0^1 h(x) f(x) dx \tag{8}$$
$$\text{s.t. } \forall x \in [0,1], a(x) + h(x) \leq 1.$$

This program has a simple solution. Define x_n^* the threshold task such that using the worker or the machine is equally profitable:

$$R[q_M(x_n^*) - q_n] + c = 0. \tag{9}$$

The solution to (8) is then to choose $a_n^*(x) = 1$ and $h_n^*(x) = 0$ if $x \leq x_n^*$, and $a_n^*(x) = 0$ and $h_n^*(x) = 1$ otherwise. Note that our assumptions on q_M and q_n guarantee the existence of a unique x_n^* in $(0, 1)$. Moreover, as a solution with $e_n = 0$ would be equivalent to setting $x_n^* = 1$, the result that $x_n^* < 1$ validates our conjecture that $e_n = 1$. We can then compute the probability of success and the return gross of management costs as:

$$\begin{aligned} y_n^* &= \int_0^{x_n^*} q_M(x)f(x)dx + \int_{x_n^*}^1 f(x)dx \\ &= F(x_n^*)Q_M(x_n^*) + [1 - F(x_n^*)]q_n, \end{aligned} \tag{10}$$

$$r_n^* = Ry_n^* - c[1 - F(x_n^*)] - 1. \tag{11}$$

We now turn to the optimal allocation of capital by investors. First, note that r_n^* will be the same for all n with the same type. Second, as $\alpha_n^* = r_n^* - \chi(k_n^*)$, (OA1) implies that all financiers of a given type receive the same capital, possibly null. Hence, we denote $r_n^* = r_L^*, k_n^* = k_L^*$, and $\alpha_n^* = \alpha_L^*$ for $n \in [0, N_L]$, and $r_n^* = r_H^*, k_n^* = k_H$, and $\alpha_n^* = \alpha_H$ for $n \in [N_L, N_L + N_H]$.

Note that we necessarily have $r_H^* > r_L^*$. Hence, we cannot have $k_H^* < k_L^*$, since this would imply $\alpha_L^* < \alpha_H^*$ and yet $k_L^* > 0$, contradicting (OA1). This implies that there are three possible types of equilibrium allocation by investors: (i) they allocate K to the type H financiers only; (ii) they allocate K to the type L and type H financiers only; (iii) they allocate less than K to the type H and type L financiers, and the rest to the risk-free asset.[15] Which type of equilibrium obtains depends on the total supply of capital as follows:

Proposition 1. *Define the following thresholds:*

$$K_0(r_L, r_H) = N_H\chi^{-1}(r_H - r_L) \tag{12}$$

$$K_1(r_L, r_H) = N_H\chi^{-1}(r_H) + N_L\chi^{-1}(r_L). \tag{13}$$

We have $0 < K_0(r_L^*, r_H^*) < K_1(r_L^*, r_H^*)$ and: (i) for $K < K_0(r_L^*, r_H^*)$, all capital is allocated to type H financiers, $N_H k_H^* = K$, $k_L^* = 0$, and $\alpha_H^* > \alpha_L^* > 0$; (ii) for $K \in [K_0(r_L^*, r_H^*), K_1(r_L^*, r_H^*))$, all capital is allocated to both types of financiers, $N_H k_H^* + N_L k_L^* = K$, and $\alpha_L^* = \alpha_H^* > 0$; (iii) for $K \geq K_1(r_L^*, r_H^*)$, capital is allocated to both types of financiers and to the risk-free asset, $N_H k_H^* + N_L k_L^* < K$, and $\alpha_L^* = \alpha_H^* = 0$.

Figure 1 below plots the different equilibrium regions as a function of c and K. When

[15] It is easy to see that investors cannot allocate to type H financiers and the risk-free asset, but not to type L financiers. If this were the case, (OA2) would imply that $r_H^* - \chi(k_H^*) = 0$. Type L financiers would offer a return $\alpha_L^* = r_L^* - \chi(0) > 0$, which contradicts (OA1).

there is little capital K to allocate, all of it can be allocated to the type H financiers, and they can offer a higher return than the type L. As K increases, this return is diminished by the costs $\chi(k_H^*)$, and at some point the type L financiers can offer an equally good return if k_L^* is sufficiently small, so that both types receive capital. As K increases further, the management costs increase and eventually both types of financiers can only offer a zero return. Any extra capital is then allocated to the risk-free asset.

[Insert Figure 1 here.]

3.2 Equilibrium with moral hazard

We now turn to the main model with moral hazard, and solve for the equilibrium contract $\mathcal{C}_n^{**} = \{q_n, a_n^{**}, h_n^{**}, w_n^{**}\}$ and capital allocation k_n^{**}.

For a given contract $\{q_n, a, h, w\}$ and capital k, denote $U_n(a, h, w, e)$ the worker's utility from exerting effort e for financier n, per unit of capital, so that his total utility is $kU_n(a, h, w, e)$. We have:

$$U_n(a, h, w, e) = w \int_0^1 [q_M(x)a(x) + q_n e h(x)] f(x)dx - ce \int_0^1 h(x)f(x)dx. \tag{14}$$

The worker will optimally exert effort if and only if $U_n(a, h, 1, w) \geq U_n(a, h, 0, w)$, which reduces to $w \geq \tilde{w}_n$, with:

$$\tilde{w}_n = \frac{c}{q_n}. \tag{15}$$

Hence, the optimal contract will be to either offer $\tilde{w}_n$, or to not rely on the worker at all. In this section, for any variable of interest X, we denote $\tilde{X}_n$ its value at the optimal solution under the assumption that financier n uses the worker. Similarly, $\tilde{X}_M$ denotes the value at the optimal solution under the assumption that the worker is not used, where "M" stands for "machine only" (note that this solution does not depend on the financier's type). Finally, $X_n^{**} \in \{\tilde{X}_n, \tilde{X}_M\}$ denotes the value at the actual optimal solution.

If financier n relies on the worker, she maximizes the return for investors according to:

$$(\tilde{a}_n, \tilde{h}_n) \in arg \max_{a,h} \left(R - \frac{c}{q_n} \right) \int_0^1 [a(x)q_M(x) + h(x)q_n] f(x)dx \tag{16}$$
$$\text{s.t. } \forall x \in [0,1], a(x) + h(x) \leq 1.$$

The solution is to choose $\tilde{a}(x) = 1, \tilde{h}(x) = 0$ for $x \leq \tilde{x}_n$ and $\tilde{a}(x) = 0, \tilde{h}(x) = 1$ otherwise, with $\tilde{x}_n$ such that:

$$q_M(\tilde{x}_n) = q_n. \tag{17}$$

This gives us the probability of success and return gross of management costs:

$$\tilde{y}_n = F(\tilde{x}_n)Q_M(\tilde{x}_n) + [1 - F(\tilde{x}_n)]q_n, \tag{18}$$
$$\tilde{r}_n = \left(R - \frac{c}{q_n} \right) \tilde{y}_n - 1. \tag{19}$$

Alternatively, financier n can choose to not use the worker. In that case, given that using the machine is free, it is optimal to choose $\tilde{a}_M(x) = 1$ for every $x \leq \tilde{x}_M = 1$. Thus:

$$\tilde{y}_M = Q_M(1), \tag{20}$$
$$\tilde{r}_M = RQ_M(1) - 1. \tag{21}$$

Financier n finds it optimal to use the worker if and only if $\tilde{r}_n \geq \tilde{r}_M$, which writes as:

$$\underbrace{R \int_{\tilde{x}_n}^1 [q_n - q_M(x)] f(x)dx}_{\text{Worker's added value}} \geq \underbrace{c[1 - F(\tilde{x}_n)]}_{\text{Effort cost}} + \underbrace{\frac{c}{q_n} F(\tilde{x}_n)Q_M(\tilde{x}_n)}_{\text{Rents}}. \tag{22}$$

This equation states that it is optimal to use the worker rather than only the machine if the value of increasing the success probability (on the left-hand side) is larger than the total expected compensation given to the worker, which is the sum of his effort cost and informational rents due to moral hazard. Note that, due to assumption (4), in the absence of rents, (22) would always hold, as it does in the complete information benchmark. Instead, in the presence of moral hazard, whether it is optimal to use the worker will depend on the cost of incentivizing him. To see that, define:

$$\hat{c}_n = Rq_n \frac{\int_{\tilde{x}_n}^1 [q_n - q_M(x)]f(x)dx}{q_n[1 - F(\tilde{x}_n)] + F(\tilde{x}_n)Q_M(\tilde{x}_n)} = Rq_n \frac{\tilde{y}_n - Q_M(1)}{\tilde{y}_n}. \tag{23}$$

It is easy to see that $\hat{c}_H > \hat{c}_L$, which corresponds to the fact that the type H worker is more productive than the type L. We then deduce from (22) the following result:

Lemma 1. *Type H and type L financiers allocate tasks as follows: (i) if $c < \hat{c}_L$, $x_n^{**} = \tilde{x}_n$ for any $n \in [0, N_L + N_H]$: both types of financiers use the worker; (ii) if $c \in [\hat{c}_L, \hat{c}_H)$, $x_n^{**} = 1$ for $n \in [0, N_L]$ and $x_n^{**} = \tilde{x}_n$ for $n \in [N_L, N_L + N_H]$: only type H financiers use the worker; (iii) if $c \geq \hat{c}_H$, $x_n^{**} = 1$ for any $n \in [0, N_L + N_H]$: both types of financiers use only the machine.*

We can finally solve for the equilibrium capital allocation. The reasoning is the same as in the complete information benchmark (Proposition 1). It also has to be the case that k_n^{**} is equalized for all $n \in [0, N_L]$ and denoted k_L^{**}, and similarly k_H^{**} for all $n \in [N_L, N_L + N_H]$. Only the definition of the returns offered by the financiers is different:

Proposition 2. *For $n \in \{L, H\}$, we have $r_n^{**} = \tilde{r}_n$ if $c < \hat{c}_n$ and $r_n^{**} = \tilde{r}_M$ otherwise. Define $K_0^{**} = K_0(r_L^{**}, r_H^{**})$ and $K_1^{**} = K_1(r_L^{**}, r_H^{**})$.*

Then (i) for $K < K_0^{**}$ all capital is allocated to type H financiers, $N_H k_H^{**} = K$, $k_L^{**} = 0$, $\alpha_H^{**} > \alpha_L^{**} > 0$; (ii) for $K \in [K_0^{**}, K_1^{**})$, all capital is allocated to both types of financiers, $N_H k_H^{**} + N_L k_L^{**} = K$, and $\alpha_L^{**} = \alpha_H^{**} > 0$; (iii) for $K \geq K_1^{**}$, capital is allocated to both types of financiers and to the risk-free asset, $N_H k_H^{**} + N_L k_L^{**} < K$, and $\alpha_L^{**} = \alpha_H^{**} = 0$.

Note that in the case where $c > \hat{c}_H$ we have $r_L^{**} = r_H^{**} = \tilde{r}_M$ and hence $K_0(r_L^{**}, r_H^{**}) = 0$. Thus, in the case where c is so high that even type H financiers do not use the worker, both types necessarily receive capital in equilibrium (since they offer the same returns). Also note that $\hat{c}_L < Rq_L$ but $\hat{c}_H$ can be higher than Rq_L, in which case Assumption (4) implies that the case $c > \hat{c}_H$ is empty. Figure 2 below plots the different equilibrium regions as a function of c and K.

[Insert Figure 2 here.]

4 Equilibrium Impact of Moral Hazard

The motivation for our model is that the extent of moral hazard makes finance occupations special, which may change how they are affected by new technologies. To better understand the impact of moral hazard in the model, this section compares the outcomes obtained in the equilibrium with moral hazard and in the complete information benchmark.

We first derive a simple observation from Sections 3.1 and 3.2:

Corollary 1. *For a given capital k_n allocated to financier n: (i) if $c < \hat{c}_n$, $x_n^{**} < x_n^* < 1$: moral hazard leads to under-using the machine; (ii) if $c \geq \hat{c}_n$, $x_n^* < x_n^{**} = 1$: moral hazard leads to over-using the machine.*

Two different forces are at play in these two cases. When c is small, tasks are allocated both to the worker and to the machine in the presence of moral hazard. The worker derives rents, shown in (22), and is compensated based on success and not based on the effort spent, as the participation constraint is slack. This implies that the worker does not need additional compensation for working on a marginal task. Hence, from the perspective of the financier, allocating more tasks to the worker is free at the margin, whereas under complete information the cost of effort c would be taken into account. This implies that the worker will be allocated tasks x he is better at ($q_n > q_m(x)$), but not tasks on which he is more efficient (which would be $Rq_n - c > Rq_m(x)$). Hence, the worker is allocated too many tasks in the presence of moral hazard, and the machine too few.

The second force at play is more obvious: moral hazard makes the worker more expensive overall, and more so when c is higher. For a high enough c, the financier prefers to not use the worker at all and to rely on the machine only. While this necessarily decreases the success probability y_n, the financier saves on the rents she would have to give to incentivize the worker.

Second, we analyze how moral hazard affects the equilibrium allocation of capital. This depends in turn on how the returns offered by both types of financiers are affected. We derive an important preliminary result:

Proposition 3. *There exists $\tilde{c} \in (\hat{c}_L, \hat{c}_H)$ such that $r_H^{**} - r_L^{**} \geq r_H^* - r_L^*$ if and only if $c \leq \tilde{c}$.*

This Proposition states that, when the cost of effort c is low, moral hazard makes the performance of both types of financiers more unequal. This is true in particular when both types of financiers use the worker. Conversely, when c is high, moral hazard makes the performance of both types of financiers more equal. This is true in particular when no type of financier uses the worker.

The intuition for the case $c > \hat{c}_H$ is clear: in that case both types of financiers use the machine only and offer exactly the same return. For the opposite case $c < \hat{c}_L$, note that $r_H^{**} - r_L^{**} \geq r_H^* - r_L^*$ is equivalent to $r_L^* - r_L^{**} \geq r_H^* - r_H^{**}$. The latter condition means that type L financiers are more negatively affected by moral hazard than type H financiers. This is true when both types use the worker: by exerting effort on task x a type L worker increases the success probability by $q_L f(x)dx$, against $q_H f(x)dx$ for a type H, with $q_H > q_L$. Hence, success is a less informative signal about effort for a type L worker, and incentivizing this type is more expensive. Mathematically, (15) shows that $\tilde{w}_n$ decreases in q_n.

We can use Proposition 3 to study how moral hazard affects which types of financiers receive capital in equilibrium:

Corollary 2. *Suppose $\chi(k)$ is linear in k. $k_H^{**} - k_L^{**} \geq k_H^* - k_L^*$ if and only if $c \leq \tilde{c}$: moral hazard makes the allocation of capital more unequal between type H and type L financiers for low c, and more equal otherwise.*

When c is low, the two types of financiers offer more different returns under moral hazard, which leads to a more unequal capital allocation. The opposite is true when c is high. The condition that $\chi(k)$ is linear is required only when $K > K_1^{**}$.

We conclude this section by studying how moral hazard affects the equilibrium level of labor. For a given set of contracts $C_n = \{q_n, a_n, h_n, w_n\}$ and capital allocations k_n, we define total labor L as the total effort exerted by workers across all n:

$$L = \int_0^{N_L+N_H} k_n e_n \left(\int_0^1 h_n(x) f(x) dx \right) dn. \tag{24}$$

We denote L^* the total labor in equilibrium under complete information, and L^{**} under moral hazard. We have:

$$L^* = N_H k_H^* [1 - F(x_H^*)] + N_L k_L^* [1 - F(x_L^*)], \tag{25}$$

$$L^{**} = N_H k_H^{**} [1 - F(x_H^{**})] + N_L k_L^{**} [1 - F(x_L^{**})]. \tag{26}$$

The difference between L^* and L^{**} is driven by three effects: (i) $x_n^{**} \leq x_n^*$: each financier uses the worker for more tasks in the presence of moral hazard (Corollary 1); (ii) $k_H^{**} - k_L^{**} \geq k_H^* - k_L^*$ if and only if $c \leq \tilde{c}$: for a given total capital allocated to financiers, moral hazard shifts the allocation towards the more labor-intensive type H financiers if c is low, and towards the less labor-intensive type L financiers otherwise (Corollary 2); (iii) $N_H k_H^{**} + N_L k_L^{**} \leq N_H k_H^* + N_L k_L^*$: there is weakly less capital allocated to the financiers under moral hazard.

The combination of these three effects can in general go in either direction. In particular, there is a case in which (i) and (ii) go in the same direction, and (iii) is null, such that counter-intuitively the presence of moral hazard leads to more labor in equilibrium:

Corollary 3. *Assume $c < \hat{c}_L$ and $K < K_1^{**}$. Then equilibrium labor is higher under moral hazard than under complete information: $L^{**} > L^*$.*

The corollary implies more generally that workers can benefit from the existence of moral hazard. Not only the workers earn rents, which they do not under complete information, but despite labor being more costly the demand for labor can actually be higher. When this happens, both the number of employees and their remuneration increase due to the presence of moral hazard.

5 Machine Improvement

We now study how an improvement in the machine affects the equilibrium in the presence of moral hazard. We assume that machine productivity can be written as:

$$q_M(x) = \tilde{q}_M(x) + \epsilon \eta(x), \text{ with } \eta(x) \geq 0. \tag{27}$$

We will focus on marginal improvements, that is, a small increase of ϵ from a starting value of 0.[16] This specification allows us to consider a wide range of possible improvements, parameterized by the function η. For some results, we will focus on two special cases:

$$\text{Core improvement} \quad : \quad \eta(x) > 0 \Leftrightarrow x < x_H^{**}, \tag{28}$$

$$\text{Frontier improvement} \quad : \quad \eta(x) > 0 \Leftrightarrow x_H^{**} < x < x_L^{**}. \tag{29}$$

A machine improvement is "core" if it only affects tasks that both H and L financiers are already allocating to the machine. In reality, this may be the case of improvements in standard datasets, in computing power (e.g., cloud services), in statistical methodologies (e.g., adoption of machine learning). In contrast, a machine improvement is "frontier" if it affects tasks that some financiers, but not all, are allocating to the machine. Improvements in general purpose AI may belong to that category.[17]

It is immediate that an improvement in the machine (weakly) leads to some tasks being reallocated from the worker to the machine. Moreover, using (17), we have:

$$\frac{\partial \tilde{x}_n}{\partial \epsilon}|_{\epsilon=0} = \frac{\eta(\tilde{x}_n)}{|q'_M(\tilde{x}_n)|} \geq 0. \tag{30}$$

Note in particular that the impact may be stronger for H financiers or for L financiers. Moreover, in the special cases of a "core" or "frontier" improvement, there is no impact on task allocation, as $\eta(\tilde{x}_n) = 0$.

We can use (16) and the envelope theorem to study how machine improvement affects the returns that different financiers can offer to investors:

$$\frac{\partial \tilde{r}_n}{\partial \epsilon}|_{\epsilon=0} = \left(R - \frac{c}{q_n}\right) \int_0^{\tilde{x}_n} \eta(x) f(x) dx, \tag{31}$$

$$\frac{\partial \tilde{r}_M}{\partial \epsilon}|_{\epsilon=0} = R \int_0^1 \eta(x) f(x) dx. \tag{32}$$

[16] Note that this implies that $\tilde{q}_M$ satisfies the same same properties as q_M.

[17] More generally, we will see below that the impact of machine improvement can depend on the distribution of this improvement across tasks. This distribution can be measured empirically. In particular, Hampole *et al.* (2025) measure the dispersion of AI exposure of tasks within an occupation. In our model, a higher dispersion would mean that $\eta(x)$ is more flat over $[0, 1]$.

We deduce from these expressions the following Lemma:

Lemma 2. *The relative impacts of machine improvement on returns are as follows:*

$$\frac{\partial \tilde{r}_M}{\partial \epsilon}|_{\epsilon=0} \quad > \quad \max\left(\frac{\partial \tilde{r}_H}{\partial \epsilon}|_{\epsilon=0}, \frac{\partial \tilde{r}_L}{\partial \epsilon}|_{\epsilon=0}\right) \geq 0, \tag{33}$$

$$\frac{\partial \tilde{r}_H}{\partial \epsilon}|_{\epsilon=0} > \frac{\partial \tilde{r}_L}{\partial \epsilon}|_{\epsilon=0} \quad \Leftrightarrow \quad \left(\frac{c}{q_L} - \frac{c}{q_H}\right) \int_0^{\tilde{x}_H} \eta(x) f(x) dx > \left(R - \frac{c}{q_L}\right) \int_{\tilde{x}_H}^{\tilde{x}_L} \eta(x) f(x) dx. \tag{34}$$

In particular, (34) is true in the case of a core improvement, and false for a frontier improvement.

Condition (34) decomposes the relative impact of machine improvement on $\tilde{r}_H$ and $\tilde{r}_L$ into two effects. On the left-hand side, we have the machine improvement on core tasks. Such improvement increases success probability in the same way for both type H and type L financiers. However, type L financiers are more exposed to the worker's moral hazard and their investors capture less of the surplus, so that improvements on core tasks benefit the type H financiers more. On the right-hand side, we have the machine improvement on frontier tasks. As type H financiers do not use the machine for such tasks, such improvement only benefits type L financiers. Finally, the impact of machine improvement on the return that can be offered by only using the machine is always the highest: all tasks are allocated to the machine so the improvement benefits all tasks, and there is no moral hazard since the worker is not used, so none of the surplus is dissipated into rents.

An implication of the first part of Lemma 2 is that a small increase in ϵ can lead either type L or type H financiers to switch from a solution with worker involvement to a solution without, but not the opposite. Formally, differentiating (23) shows that $\hat{c}_L$ and $\hat{c}_H$ are both decreasing in ϵ. In what follows, we first study the consequences of machine improvement when both types of financiers use the worker. We then briefly discuss the other cases, including when the improvement leads one type to switch to a solution with the machine only.

5.1 Both types of financiers use the worker: $c < \hat{c}_L$

We first study the case $c < \hat{c}_L$, in which both types of financiers use the worker in parallel with the machine and offer returns gross of management costs $r_H^{**} = \tilde{r}_H$ and $r_L^{**} = \tilde{r}_L$, respectively. We assume that c is sufficiently far from $\hat{c}_L$ and K from K_0^{**} and K_1^{**} so that machine improvement does not change the type of equilibrium. The impact of machine improvement on returns then has the following implication:

Proposition 4. *When $c < \hat{c}_L$, an infinitesimal increase in ϵ from $\epsilon = 0$ affects capital allocation as follows: (i) If $K < K_0^{**}$, $k_H^{**} = K/N_H$ and $k_L^{**} = 0$ are unaffected; (ii) If $K \in (K_0^{**}, K_1^{**})$, if (34) holds then k_H^{**} increases and k_L^{**} decreases, otherwise the opposite obtains; (iii) If $K > K_1^{**}$, both k_H^{**} and k_L^{**} increase.*

Cases (i) and (iii) are straightforward: when $K < K_0^{**}$ all capital is allocated to type H financiers already, so that machine improvement does not change the allocation. When $K > K_1^{**}$ some capital is allocated to the risk-free asset. When the machine improves, both types of financiers can offer higher gross returns and hence capital flows from the risk-free asset to the financiers.

Case (ii) is more subtle: capital is only allocated to both types of financiers, and machine improvement leads to a reallocation between them. Both types of financiers can offer a higher gross return after the machine improves, so capital flows from the financiers with the lower increase in returns to the financiers with the higher increase. Which type of financier benefits more is then given by Lemma 2. In particular, type H financiers can surprisingly attract more capital, for instance in the case of a core improvement, even though they use the machine less than type L financiers. This is because moral hazard is less severe for type H financiers, so that less of the machine improvement goes to rents for the worker. Conversely, for a frontier improvement, capital flows from type H to type L financiers.

We can now study how the machine improvement affects the total labor in equilibrium L^{**}. Using (26), we have:

$$
\begin{aligned}
\frac{\partial L^{**}}{\partial \epsilon}|_{\epsilon=0} &= -\frac{\partial x_H^{**}}{\partial \epsilon}|_{\epsilon=0} f(x_H^{**}) N_H k_H^{**} - \frac{\partial x_L^{**}}{\partial \epsilon}|_{\epsilon=0} f(x_L^{**}) N_L k_L^{**} \\
&+ \frac{\partial k_H^{**}}{\partial \epsilon}|_{\epsilon=0} N_H [1 - F(x_H^{**})] + \frac{\partial k_L^{**}}{\partial \epsilon}|_{\epsilon=0} N_L [1 - F(x_L^{**})]. \qquad (35)
\end{aligned}
$$

The first line is necessarily negative due to (30): as the machine improves, necessarily each financier uses the machine more and fewer tasks are allocated to the worker. However, it cannot be the case that both types of financiers lose capital, so the sign of the second line is ambiguous. There are several cases in which we can sign the total impact on labor:

Corollary 4. *When $c < \hat{c}_L$, an infinitesimal increase in ϵ from $\epsilon = 0$ has a negative impact on total labor L^{**} if $K < K_0^{**}$ or if $K \in (K_0^{**}, K_1^{**})$ and (34) is false. Moreover:*
Following a core improvement, L^{**} decreases if $K < K_0^{**}$ and increases otherwise.
Following a frontier improvement, L^{**} decreases if $K < K_1^{**}$ and increases otherwise.

In particular, this corollary shows that machine improvement can surprisingly lead to a total increase in labor, despite the machine and the worker being substitutes. This happens when the negative effect due to the reallocation of tasks is dominated by the allocation of more capital to financiers. When $K > K_1^{**}$, investors reallocate capital from the risk-free asset to both types of financiers. If the reallocation of tasks is limited (it is null in the case of core or frontier improvements), then total labor increases. When $K \in (K_0^{**}, K_1^{**})$ and (34) is true, capital is reallocated from the less labor-intensive type L financiers to the more labor-intensive type H. Again, with little reallocation of tasks, total labor increases. If instead (34) is false, then all effects go towards less labor.

We can go further and analyze whether machine improvement benefits workers. Using (14), we derive the equilibrium utility of the worker associated with financier n as:

$$
U_n^{**} = U_n(a_n^{**}, h_n^{**}, w_n^{**}, e_n^{**}) = c e_n^{**} \frac{F(x_n^{**}) Q_M(x_n^{**})}{q_n}. \qquad (36)
$$

We then define the total workers' rents in equilibrium as:

$$
U^{**} = N_H k_H^{**} U_H^{**} + N_L k_L^{**} U_L^{**}, \qquad (37)
$$

so that the impact of machine improvement on rents is:

$$
\begin{aligned}
\frac{\partial U^{**}}{\partial \epsilon}|_{\epsilon=0} &= N_H \frac{ce_H^{**}k_H^{**}}{q_H} F(x_H^{**}) \int_0^{x_H^{**}} \eta(x) f(x) dx + N_H ce_H^{**} k_H^{**} f(x_H^{**}) \frac{\partial x_H^{**}}{\partial \epsilon}|_{\epsilon=0} \\
&+ N_L \frac{ce_L^{**}k_L^{**}}{q_L} F(x_L^{**}) \int_0^{x_L^{**}} \eta(x) f(x) dx + N_L ce_L^{**} k_L^{**} f(x_L^{**}) \frac{\partial x_L^{**}}{\partial \epsilon}|_{\epsilon=0} \\
&+ N_H \frac{ce_H^{**}}{q_H} \frac{\partial k_H^{**}}{\partial \epsilon}|_{\epsilon=0} F(x_H^{**}) Q_M(x_H^{**}) + N_L \frac{ce_L^{**}}{q_L} \frac{\partial k_L^{**}}{\partial \epsilon}|_{\epsilon=0} F(x_L^{**}) Q_M(x_L^{**}). \quad (38)
\end{aligned}
$$

The first and second lines are the impact of machine improvement on workers' rents for a given capital allocation, and they consist only of positive terms. Machine improvement increases the probability of success, so that the worker is more likely to receive payment. This "free-riding effect" is the main difference between a moral hazard model and a competitive model: effort is not paid at the worker's marginal productivity. Instead, the worker is rewarded for success, even if the success mostly comes from the machine. This explains that rents can surprisingly increase when the machine improves.

The third line is a capital reallocation effect. It is clear from (36) that rents per unit of capital are higher for type L workers than for type H. This is because the effort of type H workers has a bigger impact on success probability, and hence success is a stronger signal that effort was provided. As a result, if both types of financiers attract more capital, or if capital is reallocated from type H to type L, then machine improvement also increases rents via the capital reallocation effect. If instead capital is reallocated from type L to type H financiers, then the capital reallocation effect reduces rents and the total impact of machine improvement is ambiguous. The next corollary summarizes these results:

Corollary 5. *When $c < \hat{c}_L$, an infinitesimal increase in ϵ from $\epsilon = 0$ has a positive impact on total workers' rents U^{**}, unless $K \in (K_0^{**}, K_1^{**})$ and (34) holds, in which case the impact has the sign of (38), which is ambiguous.*

We can finally study the impact of machine improvement on total welfare, which we denote W^{**}. Total welfare is the sum of gains for investors and workers' rents,[18] which writes:

$$W^{**} = N_H k_H^{**}(\alpha_H^{**} + U_H^{**}) + N_L k_L^{**}(\alpha_L^{**} + U_L^{**}). \quad (39)$$

[18] Financiers are competitive and always earn zero surplus in the model.

Similarly to what we did for workers' rents, we can decompose the impact of machine improvement on welfare into an efficiency gain effect and a capital reallocation effect:

$$\begin{aligned}\frac{\partial W^{**}}{\partial \epsilon}|_{\epsilon=0} &= N_H k_H^{**} \frac{\partial[\alpha_H^{**}+U_H^{**}]}{\partial \epsilon}|_{\epsilon=0} + N_L k_L^{**} \frac{\partial[\alpha_L^{**}+U_L^{**}]}{\partial \epsilon}|_{\epsilon=0} \\ &+ N_H \frac{\partial k_H^{**}}{\partial \epsilon}|_{\epsilon=0}[\alpha_H^{**}+U_H^{**}] + N_L \frac{\partial k_L^{**}}{\partial \epsilon}|_{\epsilon=0}[\alpha_L^{**}+U_L^{**}]. \end{aligned} \tag{40}$$

On the first line we have the impact of machine improvement via a higher efficiency for a given capital allocation. We already know that U_n^{**} increases after the machine improves, and it is easy to see that α_n^{**} does too, so this impact is always positive.

On the second line we have a capital reallocation effect. Remember that if k_H^{**} and k_L^{**} are both positive then $\alpha_H^{**} = \alpha_L^{**}$. As $U_H^{**} < U_L^{**}$, this implies that $\alpha_H^{**} + U_H^{**} < \alpha_L^{**} + U_L^{**}$: if both are active, a type H financier offers the same return as a type L financier, but creates less total surplus. Hence, when capital flows from type H to type L financiers the capital reallocation effect increases welfare, but otherwise it decreases welfare. We thus obtain a corollary similar to Corollary 5:

Corollary 6. *When $c < \hat{c}_L$, an infinitesimal increase in ϵ from $\epsilon = 0$ has a positive impact on total welfare W^{**}, unless $K \in (K_0^{**}, K_1^{**})$ and (34) holds, in which case the impact has the sign of (40), which is ambiguous.*

5.2 Other cases

We now briefly consider the cases where $c \geq \hat{c}_L$, so that type L financiers no longer use the worker.

The case $c > \hat{c}_H$ is straightforward: both types of financiers use the machine only, and offer the same return to investors. When the machine improves, this return increases, and all financiers weakly attract more capital. Workers are never used, hence their total labor and their rents are unaffected. Total welfare weakly increases when the machine improves, but the effect is null if $K > K_1(\tilde{r}_M, \tilde{r}_M)$, as the productivity increase is compensated by higher costs $\chi(k_n)$.

The case $c \in (\hat{c}_L, \hat{c}_H)$ is qualitatively similar to the case $c < \hat{c}_L$ when only type H

financiers are active, or when investors allocate capital to the risk-free asset. The only case that is slightly different is when both financiers are active and share K, that is, $K \in (K_0(\tilde{r}_M, r_H^{**}), K_1(\tilde{r}_M, r_H^{**}))$. Indeed, we know from Lemma 2 that machine improvement increases $\tilde{r}_M$ more than r_H^{**}. Thus, a reasoning similar to the one of Proposition 4 implies that k_H^{**} decreases and k_L^{**} increases. The capital reallocation effect then implies that total labor L^{**} decreases. Contrary to the case $c < \hat{c}_L$, rents are now higher in type H financiers than in type L financiers (where they are null). Hence, the impact on workers' rents and on total welfare is ambiguous.

Finally, we can consider what happens if c is just below $\hat{c}_n$, so that machine improvement leads type n financiers to stop using the worker, for $n \in \{L, H\}$. Importantly, as the threshold is crossed, r_n^{**} changes from $\tilde{r}_n$ to $\tilde{r}_M$, but there is no discontinuity. As a result, k_n^{**} and α_n^{**} move continuously too. In contrast, the labor of type n workers and their rents jump discontinuously downwards. We deduce the following:

Corollary 7. *For $n \in \{L, H\}$, consider a machine improvement from $\epsilon = 0$ to $\epsilon = d\epsilon$, such that $c < \hat{c}_n$ before improvement and $c > \hat{c}_n$ after. Then, for $d\epsilon$ small enough, improving the machine discontinuously decreases total labor L^{**}, rents U^{**}, and welfare W^{**}.*

This corollary reflects the fact that the switch to no longer using the worker depends on a maximization program that aims at maximizing returns for investors, but neglects the worker's rents. At $\hat{c}_n$, an infinitesimal increase in returns is obtained at the cost of a discrete decrease in rents for the worker. This implies that machine improvement leads to a drop in total welfare.

5.3 Discussion

Machine improvement affects workers' rents and total welfare via multiple effects, the directions of which depend on the type of improvement and on the sign of the capital reallocation effect. Table 1 summarizes all the possibilities studied in this section. Moreover, Figure 3 shows in a numerical example the regions in the (K, c) space where $k_H^{**} - k_L^{**}$, L^{**}, U^{**}, and W^{**} are affected positively or negatively.

[Insert Table 1 and Figure 3 here.]

We briefly summarize and discuss the results that are specific to our setup with moral hazard, and would typically not obtain in a traditional competitive setup.

First, machine improvement can sometimes lead to a higher demand for labor, despite labor and machines being perfect substitutes. Note that, for a given capital allocation, each financier uses the worker on fewer tasks when the machine improves, which is due to the substitutability. However, there is an additional capital allocation effect. When capital is abundant, all financiers receive more capital as the machine improves, and this increases the total demand for labor. More surprisingly, when capital is less abundant and all capital is allocated to type H and type L financiers, machine improvement can benefit the type H more than the type L financiers, despite type H financiers using the machine less. This is due to the presence of moral hazard: type L financiers use the machine more but also need to leave a higher proportion of the surplus as rents to the workers. Then, capital can flow from type L financiers to the type H financiers, which increases the total demand for labor.

Second, machine improvement can actually increase the total utility (rents) of workers, and this even when the demand for labor decreases. The reason is that moral hazard implies that workers have to be paid conditionally on success. When the machine improves, success is more likely, independently of whether the worker exerts effort. Thus, the worker is paid not only for his own marginal productivity, but also for the productivity of the machine. This is no longer true when the machine becomes so productive that the financier no longer uses the worker, in which case rents fall to zero.

This last effect implies that a machine improvement, which expands the technical possibilities of the economy, can actually decrease welfare. Moral hazard creates rents for workers, but competition between financiers is based on returns offered to investors, net of the worker's compensation. An improvement in the machine can lead financiers to completely stop using the worker, which reduces welfare. Under complete information, it is instead never optimal to stop using the worker.

Finally, the type of machine improvement matters. It is natural to interpret the development of general-purpose AI as a "frontier improvement", as it affects tasks that are currently done by workers at least in some firms. An interesting prediction from the model is that such an improvement will benefit more the financiers that are already more machine-oriented (the L types), so that they will attract more capital. However, these financiers are also those whose workers earn the largest rents. The result of reallocating capital from more worker-oriented to more machine-oriented financiers is then that total labor decreases, but the rents of the remaining workers increase, and total rents increase as well. The prediction here is thus one of fewer jobs in the financial sector, but larger rents for each job.

6 Robustness to Extensions

In this section we consider two variations of the model: (i) the worker and the machine are complements instead of substitutes; (ii) the worker chooses how much to use the machine, and this choice is unobservable to the financier.

In both cases we focus on studying whether machine improvement still increases the worker's rents in those cases. We abstract away from the capital reallocation effect, which we already studied in Section 5, and thus set $k_n = 1$. This also allows us to simplify notations and drop the subscript n.

6.1 Worker and machine are complements

We modify the model by assuming that the worker and the machine are perfect complements instead of perfect substitutes. Formally, we replace (1) with:

$$y(a, h, e) = \int_0^1 \min\{a(x)q_M(x), h(x)[(1-\gamma) + \gamma e]q\} f(x)dx. \tag{41}$$

Notice that we also allow the worker to be productive even without exerting effort: the worker's total productivity is still q, and shirking reduces this productivity by a factor $\gamma \leq 1$. The previous specification (1) was assuming $\gamma = 1$. The reason for this assumption will become clear below.

If the worker exerts effort ($e = 1$), it is trivial that for any x it is optimal to have:

$$\tilde{a}(x)q_M(x) = q\tilde{h}(x). \tag{42}$$

The worker's incentive compatibility constraint is then:

$$w \int_0^1 q\tilde{h}(x) f(x) dx - c \int_0^1 \tilde{h}(x) f(x) dx \geq w \int_0^1 (1-\gamma) q \tilde{h}(x) f(x) dx \tag{43}$$

$$\Leftrightarrow w \geq \tilde{w} = \frac{c}{\gamma q}. \tag{44}$$

It is immediate to check that if $w \geq \tilde{w}$ then the worker's participation constraint is also satisfied. The financier chooses $\tilde{h}$ so as to maximize the return r_n to investors, which gives:

$$\tilde{h} = \arg\max_{h(\cdot)} \left(R - \frac{c}{\gamma q}\right) \int_0^1 q h(x) f(x) dx - 1 \tag{45}$$
$$\text{s.t.} \quad \forall x \in [0,1], a(x) q_M(x) = q h(x), 0 \leq a(x), 0 \leq h(x), a(x) + h(x) \leq 1.$$

We obtain that $\tilde{h}(x)$ is characterized by $\tilde{a}(x) q_M(x) = q\tilde{h}(x)$ and $\tilde{a}(x) + \tilde{h}(x) = 1$, which gives:

$$\tilde{h}(x) = \frac{q_M(x)}{q + q_M(x)} \tag{46}$$

$$\tilde{r}_n = \left(R - \frac{c}{\gamma q}\right) \int_0^1 \frac{q q_M(x)}{q + q_M(x)} f(x) dx - 1. \tag{47}$$

An alternative solution for the financier is to offer $w = 0$, so that the worker does not exert effort. Denoting this solution with a subscript M, a similar reasoning to the above gives us that $\tilde{h}_M(x)$ is characterized by $\tilde{a}_M(x) q_M(x) = q(1-\gamma)\tilde{h}_M(x)$ and $\tilde{a}_M(x) + \tilde{h}_M(x) = 1$, so that:

$$\tilde{h}_M(x) = \frac{q_M(x)}{(1-\gamma) q + q_M(x)} \tag{48}$$

$$\tilde{r}_M = R \int_0^1 \frac{(1-\gamma) q q_M(x)}{(1-\gamma) q + q_M(x)} f(x) dx - 1. \tag{49}$$

The financier chooses a solution with positive effort by the worker if and only if $\tilde{r} \geq \tilde{r}_M$, which is equivalent to $c \leq \hat{c}$ with:

$$\hat{c} = \gamma^2 q R \frac{\int_0^1 \frac{q_M(x)^2}{(q+q_M(x))((1-\gamma)q + q_M(x))} f(x) dx}{\int_0^1 \frac{q_M(x)}{q + q_M(x)} f(x) dx}. \tag{50}$$

Note that necessarily $\hat{c} > 0$. From now on, we focus on the case $c < \hat{c}$ so that the worker exerts effort and receives a positive wage. The worker's rent in that case is:

$$\begin{aligned} U^{**} &= \tilde{w} \int_0^1 q\tilde{h}(x)f(x)dx - c\int_0^1 \tilde{h}(x)f(x)dx \\ &= \frac{1-\gamma}{\gamma} c \int_0^1 \frac{q_M(x)}{q+q_M(x)} f(x)dx. \end{aligned} \tag{51}$$

We can finally compute the impact of a marginal improvement in the machine:

$$\frac{\partial U^{**}}{\partial \epsilon}|_{\epsilon=0} = \frac{1-\gamma}{\gamma} c \int_0^1 \frac{q\eta(x)}{[q+q_M(x)]^2}\eta(x)f(x)dx > 0. \tag{52}$$

Thus, our result that machine improvement increases the worker's rents is robust to a setting in which the worker and the machine are complements. This is for two reasons. First, as in the case with substitutability, when the machine improves the success probability increases, even if the worker shirks. Second, as would happen also on a competitive market without moral hazard, the improvement in the machine increases the demand for labor and $\tilde{h}$ increases.

Finally, note that the rents and how much they increase when the machine improves are scaled by $\frac{1-\gamma}{\gamma}$. In particular, if $\gamma = 1$ as in the baseline model of Section 2, then the rents are null. Indeed, in that case, complementarity implies that if the worker shirks, then the success probability is zero. This logic extends to lower values of γ: in the case of complementarity, when the share of the worker's output that is subject to shirking increases, success becomes more correlated with effort, and the worker's rents decrease. This suggests that, for occupations affected by moral hazard, machine improvement may be less favorable to workers complementary to machines than to workers that are substitutes (but not replaced). This is in sharp contrast to the competitive case without moral hazard.

6.2 Unobservable use of the machine

We now consider a variant of the setup of Section 2 in which the task allocation is chosen by the worker rather than by the financier, and the worker's choice is not observable. This reflects current concerns in companies that workers may use AI to perform certain tasks without the approval of their management.[19]

The worker's utility is still given by (14). For a given wage w, the worker will choose a, h, and e so as to maximize this quantity. This program is similar to the financier's program, except that the worker receives w in case of success instead of R. Assuming that the worker chooses $e = 1$, the optimal allocation of tasks is to choose $\tilde{a}(x) = 1$ if $x < \tilde{x}$ and $\tilde{h}(x) = 1$ if $x \geq \tilde{x}$, with:

$$\tilde{w}[q_M(\tilde{x}) - q] + c = 0. \tag{53}$$

Inverting this expression, if the financier wants the worker to complete tasks $x \geq \bar{x}$, she will have to offer a wage $\tilde{w}(\bar{x})$ equal to:

$$\tilde{w}(\bar{x}) = \frac{c}{q - q_M(\bar{x})}. \tag{54}$$

Note that it is impossible to incentivize the worker to complete tasks $x < x^{**}$, as $q_M(x^{**}) = q$ (equation (17)). Implementing the same task allocation as in the case where this allocation is observable is infinitely costly. Instead, the financier will trade off letting the worker use the machine too much (relative to x^{**}) against the cost of incentivizing him to complete the tasks himself.

A second constraint the financier needs to satisfy is that the worker must be better off exerting effort than not exerting effort and letting the machine complete all the tasks. This incentive compatibility constraint writes:

$$w[F(\bar{x})Q_M(\bar{x}) + (1 - F(\bar{x}))q] - c(1 - F(\bar{x})) \geq wQ_M(1). \tag{55}$$

This condition can be rewritten as:

$$w \int_{\bar{x}}^{1} [q - q_M(x)] f(x) dx \geq c(1 - F(\bar{x})). \tag{56}$$

[19]See Yang *et al.* (2025) for experimental evidence on the "covert" use of generative AI, and the impact of companies' disclosure policies on AI use.

Using (54), if we replace w by $\tilde{w}(\bar{x})$ in this expression we obtain:

$$\int_{\bar{x}}^{1} \frac{[q - q_M(x)]}{1 - F(\bar{x})} f(x)dx \geq q - q_M(\bar{x}), \tag{57}$$

which is true as $q_M(.)$ is decreasing. Hence, if the financier chooses $w = \tilde{w}(\bar{x})$ then the worker necessarily exerts effort. The financier's problem then reduces to choosing the optimal $\bar{x}$ to maximize the return $\tilde{r}$, which we denote:

$$\tilde{x} = arg \max_{\bar{x} \geq x^{**}} (R - \tilde{w}(\bar{x}))\,[F(\bar{x})Q_M(\bar{x}) + (1 - F(\bar{x}))q]. \tag{58}$$

Solving this program gives us the following result:

Proposition 5. *Assume* $\frac{q'_M(x)}{f(x)}$ *is increasing, and* $c < \hat{c}$*, where* $\hat{c} > 0$ *is defined in Appendix A.2.14. Then, the wage optimally offered by the financier is* $\tilde{w}(x^u) > 0$*, with* $x^u \in (x^{**}, 1)$*. The worker completes tasks* $x > x^u$ *and exerts effort,* $e^u = 1$*.*

The assumption that $\frac{q'_M(x)}{f(x)}$ is increasing is sufficient to ensure the existence of a unique solution, whereas $c < \hat{c}$ ensures that the financier asks the worker to exert effort instead of allocating all tasks to the machine. From now on, we focus on the solution derived in Proposition 5. We can compute the worker's rents as:

$$U^u = \frac{c}{q - q_M(x^u)}[F(x^u)Q_M(x^u) + (1 - F(x^u))q_M(x^u)] - c[1 - F(x^u)]. \tag{59}$$

We want to study whether machine improvement has a positive impact on the worker's rents. Denoting $\frac{\partial U^u}{\partial \epsilon}$ the derivative of U^u with respect to ϵ but keeping x^u constant, we have:

$$\frac{dU^u}{d\epsilon}|_{\epsilon=0} = \frac{\partial U^u}{\partial x^u}\frac{\partial x^u}{\partial \epsilon}|_{\epsilon=0} + \frac{\partial U^u}{\partial \epsilon}|_{\epsilon=0}. \tag{60}$$

We can sign the different terms as follows:

Corollary 8. *We have* $\frac{\partial U^u}{\partial \epsilon}|_{\epsilon=0} \geq 0$ *and* $\frac{\partial U^u}{\partial x^u} \leq 0$*.* $\frac{\partial x^u}{\partial \epsilon}|_{\epsilon=0}$ *is positive if and only if* $\eta'(x^u) \leq \hat{\eta}$*, where* $\hat{\eta} > 0$ *is defined in the Appendix A.2.15.*

The improvement of the machine has several effects on the worker's rents. The first one is $\frac{\partial U^u}{\partial \epsilon}|_{\epsilon=0} \geq 0$: this is again the free-riding effect, which is thus robust to the task allocation being unobservable.

However, there is now a second effect. When the machine improves, and unless $\eta'(x^u)$ is too large, the financier targets a task allocation with more tasks allocated to the machine. This is intuitive, since the machine is more efficient. However, when machine use is not observable, the worker's rents derive from the financier having to incentivize the worker to exert effort instead of free-riding on the machine. As the financier needs the worker less, his compensation and rents are lower.

This Corollary implies that, perhaps surprisingly, machine improvement is expected to increase the worker's rents less when the worker has full discretion on how to improve the machine. His rents are higher than if the use of the machine were observable, but they are less sensitive to machine improvement. Figure 4 gives an example in which this happens. Intuitively, as the machine improves, the financier wants to use it more, which is exactly what the worker is tempted to do. Hence, the conflict between the two decreases, and incentive pay decreases.

[Insert Figure 4 here.]

This result suggests that empirical studies on which occupations are more affected by machine improvement should not only distinguish occupations according to how much they are exposed to moral hazard (as the comparison between the complete information benchmark and the model with moral hazard suggests), but also according to whether the deployment of the machine is controlled, or at least observed, by the management.

7 Conclusion

A striking feature of recent advances in automation is that they are affecting well-paid knowledge workers, whereas previous waves of automation typically affected manual workers (Kogan *et al.*, 2023). Such workers are more likely to receive agency rents. Thus, models designed to analyze the automation of manual work and relying on competitive labor markets may not tell the whole story. Indeed, our model shows that workers' rents can actually increase with automation, even in the case of perfect substitutability, through both a free-riding effect and a capital reallocation effect.

Due to the pervasiveness of agency rents in the financial sector, we argue that understanding the impact of new technologies on the pay of finance workers requires to enrich existing models of automation with a moral hazard component. Our paper is only a first step, and there are several interesting avenues for future research. First, our model endogenizes the allocation of capital, but the availability and types of workers are exogenous. In reality, the presence of rents in the financial sector may attract too much talent that may be better allocated elsewhere, a phenomenon that according to our model could be exacerbated by AI. Second, our model is static. If a large share of occupations currently impacted by AI are characterized by agency rents, including such rents in a dynamic model may be important to understand the long-term impact of AI on economic growth and innovation (Acemoglu and Restrepo (2018b), Gans (2025)) and income distribution (Bond and Kremens, 2024).

References

Abis, S. and Veldkamp, L. (2023). The changing economics of knowledge production. *The Review of Financial Studies*, **37** (1), 89–118. 1

Acemoglu, D. (2021). *Harms of AI*. Working paper. 6

—, Autor, D., Hazell, J. and Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics*, **40** (S1), S293–S340. 6

— and Restrepo, P. (2018a). Modeling automation. *AEA Papers and Proceedings*, **108**, 48–53. 6

— and — (2018b). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, **108** (6), 1488–1542. 1, 35

Aghion, P., Bunel, S., Jaravel, X., Mikaelsen, T., Roulet, A. and Søgaard, J. (2025). How different uses of ai shape labor demand: Evidence from france. *AEA Papers and Proceedings*, **115**, 62–67. 6

Agrawal, A., Gans, J. S. and Goldfarb, A. (2019). Artificial intelligence: The ambiguous labor market impact of automating prediction. *Journal of Economic Perspectives*, **33** (2), 31–50. 6

Aldasoro, I., Doerr, S., Gambacorta, L., Notra, S., Oliviero, T. and Whyte, D. (2024). Generative artificial intelligence and cyber security in central banking. *Journal of Financial Regulation*, **11** (1), 119–128. 7

Allen, R. and Choudhury, P. R. (2022). Algorithm-augmented work and domain experience: The countervailing forces of ability and aversion. *Organization Science*, **33** (1), 149–169. 8

Alonso, R. and Câmara, O. (2024). Organizing data analytics. *Management Science*, **70** (5), 3123–3143. 8

Athey, S. C., Bryan, K. A. and Gans, J. S. (2020). The allocation of decision authority to human and artificial intelligence. In *AEA Papers and Proceedings*, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, vol. 110, pp. 80–84. 8

Autor, D. H. (2015). Why are there still so many jobs? the history and future of workplace automation. *Journal of Economic Perspectives*, **29** (3), 3–30. 6

Axelson, U. and Bond, P. (2015). Wall street occupations. *The Journal of Finance*, **70** (5), 1949–1996. 5

Babina, T., Fedyk, A., He, A. and Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, **151**, 103745. 6

Bertomeu, J., Lin, Y., Liu, Y. and Ni, Z. (2025). The impact of generative ai on information processing: Evidence from the ban of chatgpt in italy. *Journal of Accounting and Economics*, p. 101782. 7

Biais, B. and Landier, A. (2020). Endogenous agency problems and the dynamics of rents. *The Review of Economic Studies*, **87** (6), 2542–2567. 5

Bick, A., Blandin, A. and Deming, D. (2024). *The Rapid Adoption of Generative AI.* Working paper. 1

Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Patnaik, M., Saporta-Eksten, I. and Van Reenen, J. (2019). What drives differences in management practices? *American Economic Review*, **109** (5), 1648–83. 3

Bond, P. and Glode, V. (2014). The labor market for bankers and regulators. *The Review of Financial Studies*, **27** (9), 2539–2579. 6, 9

— and Kremens, L. (2024). *Income and Inequality under Asymptotically Full Automation.* Working paper. 35

Bonelli, M. (2025). *Data-driven Investors.* Working paper. 1

— and Foucault, T. (2025). *Does Big Data Devalue Traditional Expertise? Evidence from Active Funds.* Working paper. 7

Bonfiglioli, A., Crinò, R., Gancia, G. and Papadakis, I. (2024). Artificial intelligence and jobs: evidence from US commuting zones. *Economic Policy*, **40** (121), 145–194. 6

Boyaci, T., Canyakmaz, C. and de Véricourt, F. (2024). Human and machine: The impact of machine input on decision making under cognitive limitations. *Management Science*, **70** (2), 1258–1275. 8

Brynjolfsson, E., Li, D. and Raymond, L. (2025). Generative AI at Work. *The Quarterly Journal of Economics*, **140** (2), 889–942. 8

—, Mitchell, T. and Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, **108**, 43–47. 6

Campello, M., Cong, L. W. and Zhou, L. (2025). *AlphaManager: A Data-Driven-Robust-Control Approach to Corporate Finance*. Working paper. 1

Cao, S., Jiang, W., Wang, J. and Yang, B. (2024). From Man vs. Machine to Man + Machine: The art and AI of stock analyses. *Journal of Financial Economics*, **160**, 103910. 1

Célérier, C. and Vallée, B. (2019). Returns to talent and the finance wage premium. *The Review of Financial Studies*, **32** (10), 4005–4040. 1

de Véricourt, F. and Gurkan, H. (Forthcoming). Is your machine better than you? you may never know. *Management Science*, p. null. 8

Dessaint, O., Foucault, T. and Frésard, L. (2024). Does alternative data improve financial forecasting? the horizon effect. *The Journal of Finance*, **79** (3), 2237–2287. 7

Dugast, J. and Foucault, T. (2025). Equilibrium data mining and data abundance. *The Journal of Finance*, **80** (1), 211–258. 7

Efing, M., Hau, H., Kampkötter, P. and Rochet, J.-C. (2022). Bank bonus pay as a risk sharing contract. *The Review of Financial Studies*, **36** (1), 235–280. 3

Eisfeldt, A. and Schubert, G. (2025). *Generative AI and Finance*. Working paper. 7

—, —, Zhang, M. B. and Taska, B. (2025). *Generative AI and Firm Values*. Working paper. 6

Farboodi, M. and Veldkamp, L. (2020). Long-run growth of financial data technology. *American Economic Review*, **110** (8), 2485–2523. 7

Fugener, A., Grahl, J., Gupta, A. and Ketter, W. (2020). Will humans-in-the-loop become borgs? merits and pitfalls of working with ai. *MIS Quarterly*, **45** (3), 1527–1556. 8

Gans, J. (2025). *Growth in AI Knowledge*. Working paper. 35

Glode, V. and Lowery, R. (2016). Compensating financial experts. *The Journal of Finance*, **71** (6), 2781–2808. 6

— and Ordoñez, G. (2024). Technological progress and rent seeking. *The Review of Financial Studies*, **38** (4), 1259–1289. 6

Gregory, T., Salomons, A. and Zierahn, U. (2021). Racing with or against the machine? evidence on the role of trade in europe. *Journal of the European Economic Association*, **20** (2), 869–906. 6

Grennan, J. and Michaely, R. (2020). *Artificial Intelligence and High-Skilled Work: Evidence from Analysts*. Working paper. 7

Hampole, M., Papanikolaou, D., Schmidt, L. and Seegmiller, B. (2025). *Artificial Intelligence and the Labor Market*. Working paper. 21

Hoffmann, F., Inderst, R. and Opp, M. (2022). The economics of deferral and clawback requirements. *The Journal of Finance*, **77** (4), 2423–2470. 5

Ide, E. and Talamas, E. (2024). *The Turing Valley: How AI Capabilities Shape Labor Income*. Working paper. 8

— and — (Forthcoming). Artificial intelligence in the knowledge economy. *Journal of Political Economy*. 8

Jiang, W., Park, J., Xiao, R. and Zhang, S. (2025). *AI and the Extended Workday: Productivity, Contracting Efficiency, and Distribution of Rents*. Working paper. 6

—, Tang, Y., Xiao, R. and Yao, V. (Forthcoming). Surviving the Fintech Disruption. *Journal of Financial Economics*. 7

Kogan, L., Papanikolaou, D., Schmidt, L. and Seegmiller, B. (2023). *Technology and Labor Displacement: Evidence from Linking Patentswith Worker-Level Data*. Working paper. 6, 34

Lax, P. D. (2014). *Functional analysis*. John Wiley & Sons. 49

Lopez-Lira, A. and Tang, Y. (2024). *Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models*. Working paper. 1

Philippon, T. and Reshef, A. (2012). Wages and Human Capital in the U.S. Finance Industry: 1909-2006. *The Quarterly Journal of Economics*, **127** (4), 1551–1609. 1

SHENG, J., SUN, Z., YANG, B. and ZHANG, A. (2024). *Generative AI and Asset Management*. Working paper. 7

TENA, V. (2020). *Will Asset Managers Survive the Advent of Robots? An Optimal Contracting Approach*. Working paper. 6

THANASSOULIS, J. (2012). The case for intervening in bankers' pay. *The Journal of Finance*, **67** (3), 849–895. 6

WANG, W., GAO, G. G. and AGARWAL, R. (2024). Friend or foe? teaming between artificial intelligence and workers with variation in experience. *Management Science*, **70** (9), 5753–5775. 8

YANG, C., RESTREPO AMARILES, D., ALLEN, L. and TROUSSEL, A. (2025). *GPT Adoption and the Impact of Disclosure Policies*. Working paper. 32

YE, S. (2025). *Task Efficiency and Signaling in the Age of GenAI: Effort Reallocation and Firm Value Effects*. Working paper. 8

ZHANG, Y. (2024). *Do Mutual Funds Benefit from the Adoption of AI Technology?* Working paper. 7

ZHONG, H. (2025). *Optimal Integration: Human, Machine, and Generative AI*. Working paper. 8

Appendix

A.1 Figures and Tables

Unless mentioned otherwise, the figures use the following parameters, meant for illustration only: $R = 3$, $q_L = 0.5$, $q_H = 0.75$, $N_H = 1$, $N_L = 1$, $f(x) = 1$, $q_M(x) = 0.5\left(\frac{1}{\sqrt{x}} - 1\right)$, $\chi(k) = 0.5k$, $\eta(x) = q_M(x)$.

Figure 1: Equilibrium in the complete information benchmark

This figure plots which type of equilibrium obtains under complete information, in the (K, c) space: (i) all capital is allocated to type H financiers (blue); (ii) all capital is allocated to type H and type L financiers (yellow); (iii) capital is allocated to type H and type L financiers, and to the risk-free asset (green). The solid black lines correspond to $K_0(r_L^{**}, r_H^{**})$ and $K_1(r_L^{**}, r_H^{**})$.

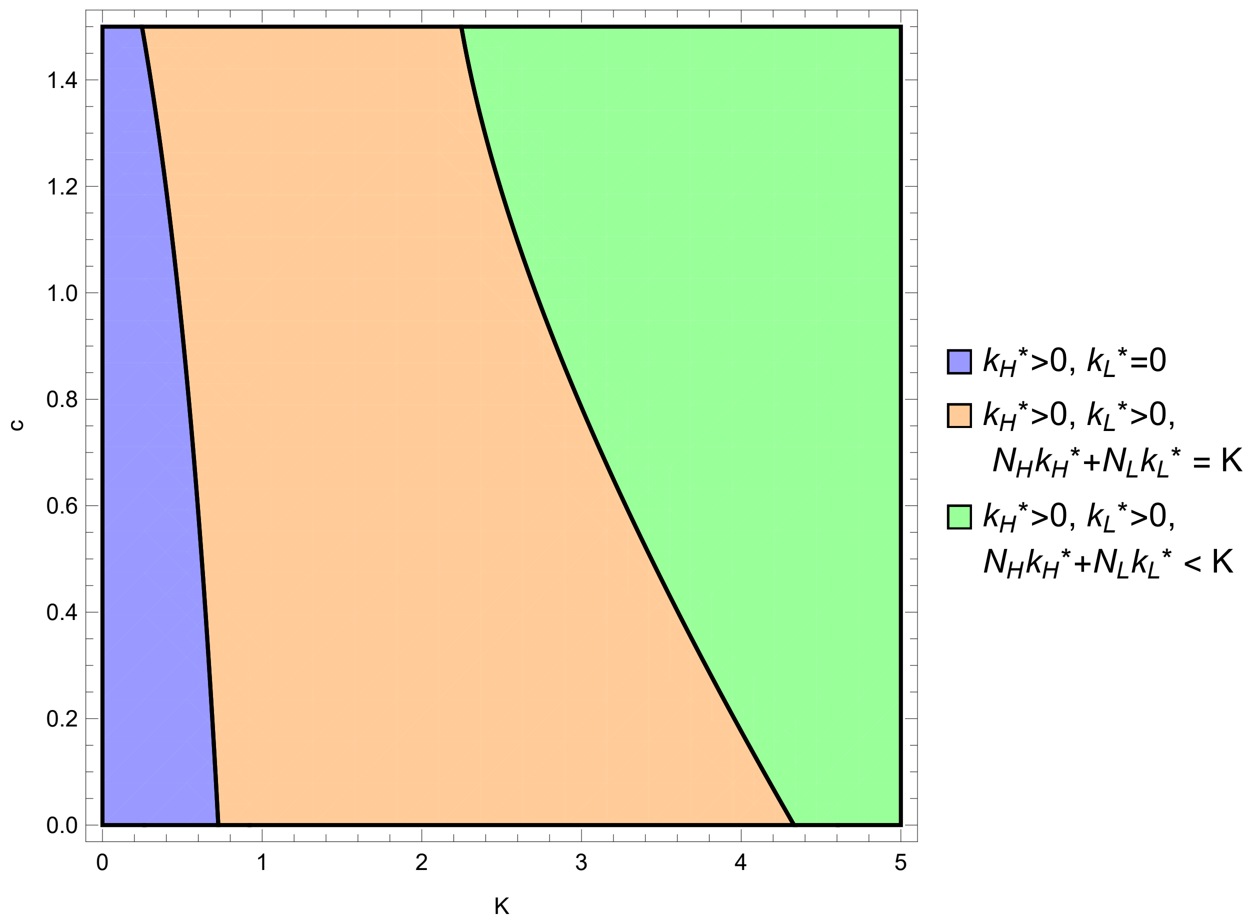


Figure 2: Equilibrium under moral hazard

This figure plots which type of equilibrium obtains under moral hazard, in the (K, c) space. The solid black lines correspond to $\hat{c}_L$, $\hat{c}_H$, K_0^{**} and K_1^{**}. Capital is allocated only to type H financiers in blue regions, to only to both types of financiers in orange regions, and to both types of financiers and the risk-free asset in green regions. For $c < \hat{c}_L$ both types of financiers use the worker, for $c \in (\hat{c}_L, \hat{c}_L)$ only type H financiers do, and for $c > \hat{c}_L$ both types of financiers use the machine only.

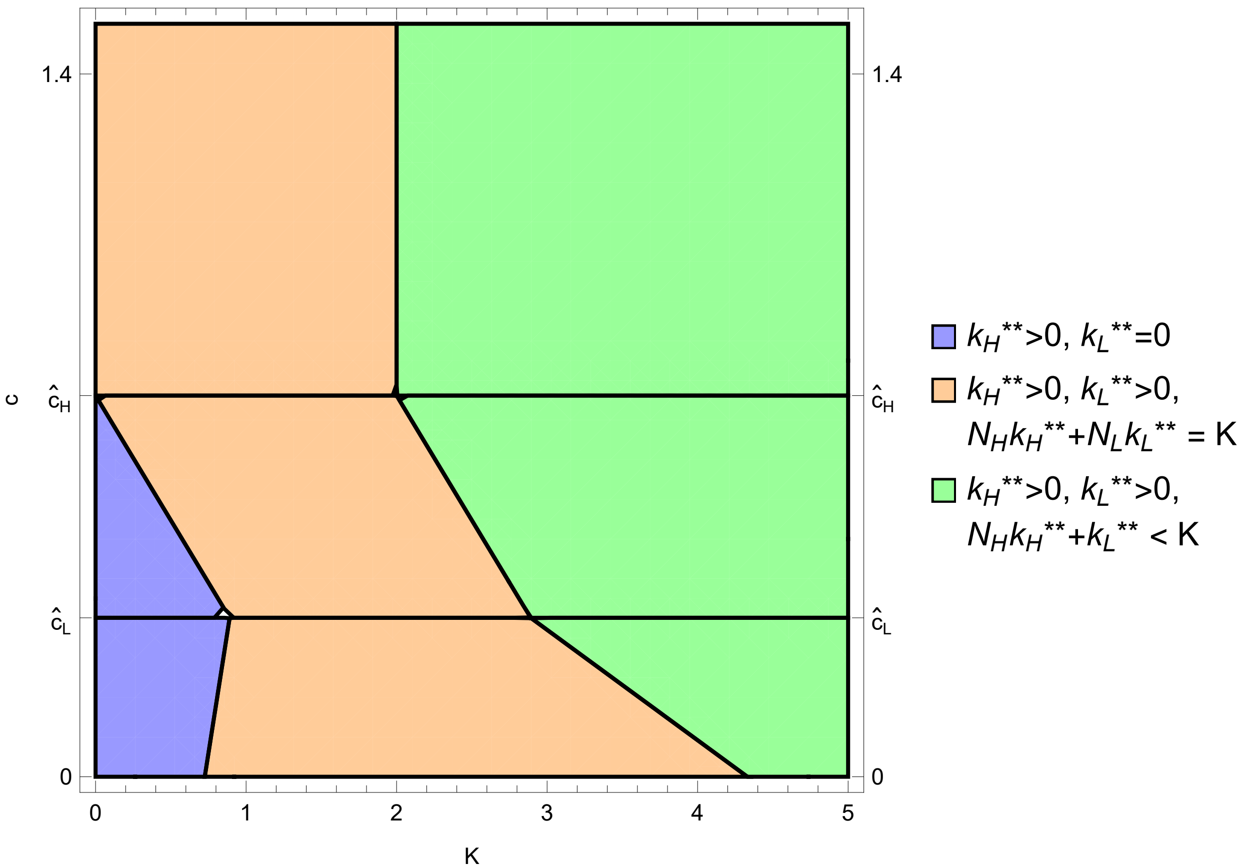


Table 1: Impact of machine improvement on capital allocation, labor, rents, and welfare

This table summarizes the signs of the impact of a marginal improvement in machine performance on capital allocation k_H^{**} and k_L^{**}, total labor L^{**}, worker rents U^{**}, and total welfare W^{**}, for different regions of the parameters c and K. The corresponding results were shown in Proposition 4, Corollaries 4 to 7, and Section 5.2. $c \simeq \hat{c}_H$ and $c \simeq \hat{c}_L$ denote cases in which c is just below the threshold, so that machine improvement makes the corresponding type of financier stop using the worker. A $+/-$ entry means that the sign is ambiguous, and depends on a condition identified in the proof of the associated proposition or corollary.

	k_H^{**}	k_L^{**}	L^{**}	U^{**}	W^{**}
$c > \hat{c}_H$					
$K < K_1^{**}$	0	0	0	0	+
$K > K_1^{**}$	+	+	0	0	+
$c \simeq \hat{c}_H$					
Any K	0	0	−	−	−
$\hat{c}_L < c < \hat{c}_H$					
$K < K_0^{**}$	0	0	−	+	+
$K_0^{**} < K < K_1^{**}$	−	+	−	+/−	+/−
$K > K_1^{**}$	+	+	+/−	+	+
$c \simeq \hat{c}_L$					
$K < K_0^{**}$	0	0	0	0	0
$K > K_0^{**}$	0	0	−	−	−
$c < \hat{c}_L$					
$K < K_0^{**}$	0	0	−	+	+
$K_0^{**} < K < K_1^{**}$ and (34)	+	−	+/−	+/−	+/−
$K_0^{**} < K < K_1^{**}$ and $\neg$ (34)	−	+	−	+	+
$K > K_1^{**}$	+	+	+/−	+	+

Figure 3: Impact of machine improvement

The four panels below illustrate the impact of machine improvement on capital reallocation, labor, workers' rents, and total welfare. We consider an improvement that increases $q_M(x)$ by 10% for every x. For each variable, we plot in red the region where the variable is affected negatively, in green the region where it is affected positively, and in white the region where it is unaffected. The solid black lines correspond to $\hat{c}_L$, $\hat{c}_H$, K_0^{**} and K_1^{**}, before machine improvement.

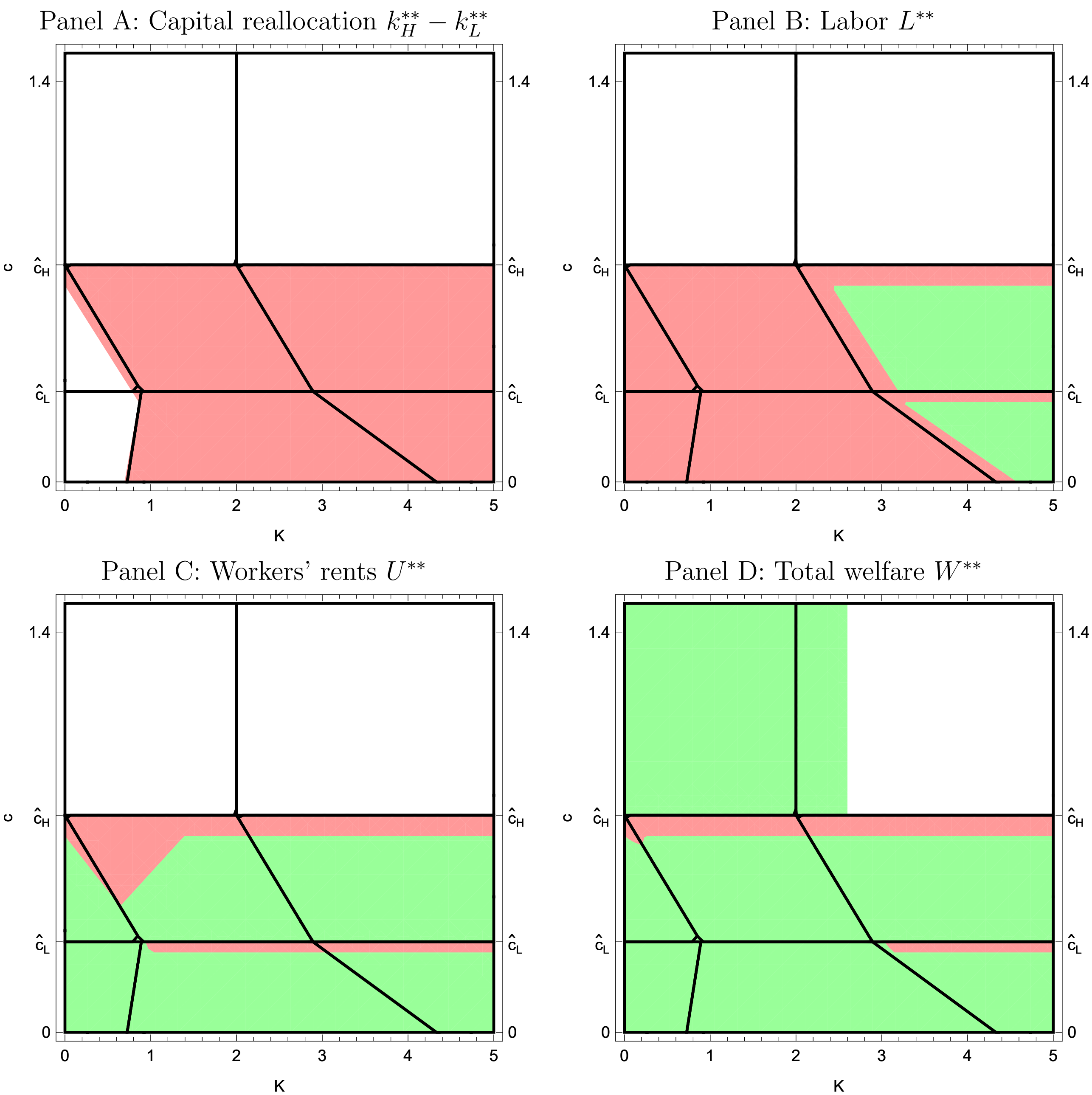
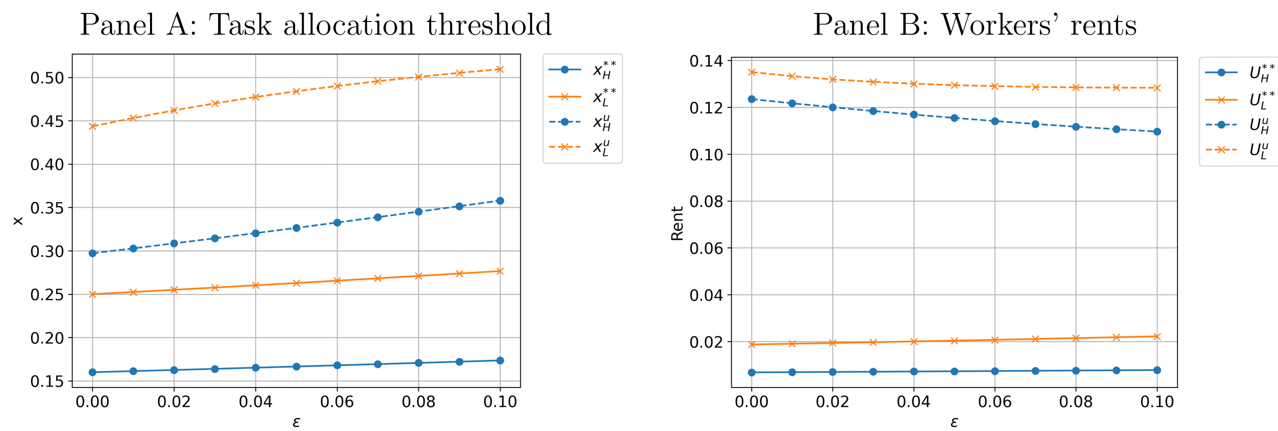


Figure 4: Impact of machine improvement when machine use is not observable

The figure shows the impact of improving the machine when machine use is not observable. Panel A plots the task allocation threshold x^u_H (for $q = q_H$) and x^u_L (for $q = q_L$), and compares them to the thresholds x^{**}_H and x^{**}_L in the baseline model. Panel B plots the workers' rents U^u both for $q = q_H$ and $q = q_L$, and compares them to U^{**}_H and U^{**}_L in the baseline model. The parameters are the same as in the previous figures, except that $\eta(x) = 0.5/[1 + exp(20(x - 0.5))]$. K is set to 1 and c to 0.1.



A.2 Proofs

A.2.1 Proof of Proposition 1

We first consider point (i). If all capital is allocated to type H financiers then $k^*_H = K/N_H$ and $\alpha^*_H = r^*_H - \chi(K/N_H)$. The type L financiers have no capital, $k^*_L = 0$, and can offer $\alpha^*_L = r^*_L$. Assumption (3) guarantees that $r^*_L > 0$. Such an equilibrium obtains if and only if $\alpha^*_H > \alpha^*_L$, which gives:

$$K \leq K_0(r^*_L, r^*_H) = N_H \chi^{-1}(r^*_H - r^*_L). \quad \text{(A.1)}$$

Point (ii): As all capital is allocated to both type H and type L financiers, we have $N_H k^*_H + N_L k^*_L = K$, and (OA1) gives us $r^*_H - \chi(k^*_H) = r^*_L - \chi(k^*_L)$. We thus have:

$$\chi\left(\frac{K - N_L k^*_L}{N_H}\right) - \chi(k^*_L) = r^*_H - r^*_L. \quad \text{(A.2)}$$

The left-hand side of this equation is decreasing in k^*_L and is negative when all $k^*_L = K/N_L$. We thus have a solution with a positive k^*_L if and only if the left-hand side is greater than $r^*_H - r^*_L$ when $k^*_L = 0$, which gives $\chi(K/N_H) \geq r^*_H - r^*_L$, which is equivalent to $K \geq K_0(r^*_L, r^*_H)$.

We then need to have $\alpha_L^* = r_L^* - \chi(k_L^*) > 0$. For $K = K_0(r_L^*, r_H^*)$ we have $\alpha_L^* = r_L^*$ and hence this is true. As K increases, k_L^* increases (using (A.2)) and α_L^* decreases. As K goes to infinity, so does k_L^*, so that necessarily α_L^* is negative for K large enough. This implies that there exists a unique $K_1 > K_0(r_L^*, r_H^*)$ such that $\alpha_L^* > 0$ if and only if $K < K_1$. This threshold value is such that (A.2) holds and $r_L^* = \chi(k_L^*)$, which gives $k_L^* = \chi^{-1}(r_L^*)$ and:

$$K_1(r_L^*, r_H^*) = N_H \chi^{-1}(r_H^*) + N_L \chi^{-1}(r_L^*). \tag{A.3}$$

Point (iii): For this last possibility we need $r_H^* - \chi(k_H^*) = r_L^* - \chi(k_L^*) = 0$, which gives $k_H^* = \chi^{-1}(r_H^*)$ and $k_L^* = \chi^{-1}(r_L^*)$. Thus $N_H k_H^* + N_L k_L^* = N_H \chi^{-1}(r_H^*) + N_L \chi^{-1}(r_L^*) = K_1(r_L^*, r_L^*)$. This equilibrium obtains if and only if $N_H k_H^* + N_L k_L^* \leq K$, which gives $K \geq K_1(r_L^*, r_H^*)$.

A.2.2 Proof of Lemma 1

The proof follows directly from isolating c in (22).

A.2.3 Proof of Proposition 2

Lemma 1 gives us that type H financiers offer the return (gross of costs) r_H^{**} and type L financiers offer the return r_L^{**} as defined in the proposition. For $c < \hat{c}_H$ we have $r_H^{**} > r_L^{**}$ and one can use the proof of Proposition 1 in Section A.2.1, simply replacing r_H^* and r_L^* with r_H^{**} and r_L^{**}, α_H^* and α_L^* with α_H^{**} and α_L^{**}, and k_H^* and k_L^* with k_H^{**} and k_L^{**}.

When $c \geq \hat{c}_H$ we have $r_L^{**} = r_H^{**} = \tilde{r}_M$. As all financiers offer the same return to investors, they must all receive the same amount of capital in equilibrium. Hence, there are only two possible types of equilibrium:

(i) no capital is allocated to the risk-free asset, so that every financier receives $\frac{K}{N_L+N_H}$. This requires that for every n we have $\alpha_n^{**} = \tilde{r}_M - \chi\left(\frac{K}{N_L+N_H}\right) > 0$, which is equivalent to $K < (N_L + N_H)\chi^{-1}(\tilde{r}_M) = K_1(r_L^{**}, r_H^{**})$.

(ii) some capital is allocated to the risk-free asset, which requires that $\alpha_n^{**} = \tilde{r}_M - \chi\left(k_n^{**}\right) = 0$, or equivalently $k_n^{**} = \chi^{-1}(\tilde{r}_M)$. We need to have $(N_L + N_H)k_n^{**} \leq K$, which is equivalent to $K \geq K_1(r_L^{**}, r_H^{**})$.

Note that $K_0(r_L^{**}, r_H^{**}) = N_H \chi^{-1}(\tilde{r}_M - \tilde{r}_M) = 0$ when $c \geq \hat{c}_H$, so the set $K < K_0(r_L^{**}, r_H^{**})$ is empty.

A.2.4 Proof of Corollary 1

The corollary follows direction from the definitions of x_n^* and $\tilde{x}_n$ in (9) and (17), and from Lemma 1.

A.2.5 Proof of Proposition 3

When $c > \hat{c}_H$ we have $r_H^{**} - r_L^{**} = \tilde{r}_M - \tilde{r}_M = 0$, whereas $r_H^* > r_L^*$. Thus, $r_H^{**} - r_L^{**} < r_H^* - r_L^*$ in this region.

When $c < \hat{c}_L$, we can write:

$$r_n^* = \max_{\bar{x}} R[F(\bar{x})Q_M(\bar{x}) + (1 - F(\bar{x}))q_n] - c(1 - F(\bar{x})) - 1, \tag{A.4}$$

$$r_n^{**} = \max_{\bar{x}} \left(R - \frac{c}{q_n}\right)[F(\bar{x})Q_m(\bar{x}) + (1 - F(\bar{x}))q_n] - 1. \tag{A.5}$$

Using the envelope theorem, we have:

$$\frac{\partial[r_n^* - r_n^{**}]}{\partial q_n} = R(1 - F(x_n^*)) - \left(R - \frac{c}{q_n}\right)(1 - F(\tilde{x}_n)) - \frac{c}{q_n^2}\tilde{y}_n \tag{A.6}$$

$$= -R(F(x_n^*) - F(\tilde{x}_n)) - \frac{c}{q_n^2}[\tilde{y}_n - q_n(1 - F(\tilde{x}_n))]. \tag{A.7}$$

Corollary 1 implies that $F(x_n^*) - F(\tilde{x}_n) \geq 0$, and (18) implies that $\tilde{y}_n \geq q_n(1 - F(\tilde{x}_n))$. Hence, $r_n^* - r_n^{**}$ decreases in q_n, which implies that $r_H^* - r_H^{**} < r_L^* - r_L^{**}$, or equivalently $r_H^* - r_L^* < r_H^{**} - r_L^{**}$.

Finally, when $c \in [\hat{c}_L, \hat{c}_H]$, we write:

$$r_H^{**} = \max_{\bar{x}} \left(R - \frac{c}{q_H}\right)[F(\bar{x})Q_m(\bar{x}) + (1 - F(\bar{x}))q_H] - 1, \tag{A.8}$$

$$r_L^{**} = RQ_M(1) - 1, \tag{A.9}$$

$$r_n^* = \max_{\bar{x}} R[F(\bar{x})Q_M(\bar{x}) + (1 - F(\bar{x}))q_n] - c(1 - F(\bar{x})) - 1. \tag{A.10}$$

Denote $\Delta = [r_H^{**} - r_L^{**}] - [r_H^* - r_L^*]$. Δ is continuous in c at $c = \hat{c}_L$ and $c = \hat{c}_H$: indeed, $\hat{c}_n$ is defined by $r_n^{**} = \tilde{r}_M$. Hence, we know that Δ is negative for $c = \hat{c}_H$ and positive for $c = \hat{c}_L$. We only need to show that Δ decreases in c for $c \in [\hat{c}_L, \hat{c}_H]$ to complete the proof.

Using (A.8) to (A.10) and the envelope theorem, we have:

$$\frac{\partial \Delta}{\partial c} = -\frac{\tilde{y}_H}{q_H} - F(x_H^*) + F(x_L^*) \tag{A.11}$$

$$= -\frac{1}{q_H}\left(F(\tilde{x}_H)Q_M(\tilde{x}_H) + q_H[1 - F(x_L^*) + F(x_H^*) - F(\tilde{x}_H)]\right). \tag{A.12}$$

Corollary 1 implies that $F(x_H^*) \geq F(\tilde{x}_H)$, so that Δ decreases in c, which concludes the proof.

A.2.6 Proof of Corollary 2

We first prove the following Lemma:

Lemma 3. *We have $K_1(r_L^{**}, r_H^{**}) \leq K_1(r_L^*, r_H^*)$: if some capital is allocated to the risk-free asset under complete information, then this is also the case under moral hazard.*

If $c \leq \tilde{c}$ we have $K_0(r_L^{**}, r_H^{**}) \geq K_0(r_L^*, r_H^*)$: if capital is only allocated to type H financiers under complete information, then this is also the case under moral hazard.

If $c \geq \tilde{c}$ we have $K_0(r_L^{**}, r_H^{**}) \leq K_0(r_L^*, r_H^*)$: if capital is allocated to both type of financiers under complete information, then this is also the case under moral hazard.

We first compare $K_1(r_L^{**}, r_H^{**})$ and $K_1(r_L^*, r_H^*)$. Consider $K = K_1(r_L^*, r_H^*)$. Then we have $r_H^* - \chi(k_H^*) = r_L^* - \chi(k_L^*) = 0$, and hence $k_H^* = \chi^{-1}(r_H^*)$ and $k_L^* = \chi^{-1}(r_L^*)$. Moreover, we have $K = N_H k_H^* + N_L k_L^* = N_H \chi^{-1}(r_H^*) + N_L \chi^{-1}(r_L^*)$.

By contradiction, assume that $K = K_1(r_L^*, r_H^*) < K_1(r_L^{**}, r_H^{**})$. Then we must have $r_H^{**} - \chi(k_H^{**}) = r_L^{**} - \chi(k_L^{**}) > 0$, and $N_H k_H^{**} + N_L k_L^{**} = K$. The first property implies that $k_H^{**} < \chi^{-1}(r_H^{**})$ and $k_L^{**} < \chi^{-1}(r_L^{**})$. Moreover, we have $r_H^{**} < r_H^*$ and $r_L^{**} < r_L^*$ and χ^{-1} is increasing. Hence, we have $k_H^{**} < \chi^{-1}(r_H^*) = k_H^*$ and $k_L^{**} < \chi^{-1}(r_L^*) = k_L^*$.

To conclude, we obtain that $k_H^{**} < k_H^*$ and $k_L^{**} < k_L^*$, and yet $N_H k_H^{**} + N_L k_L^{**} = N_H k_H^* + N_L k_L^* = K$, a contradiction. This proves that $K_1(r_L^*, r_H^*) \geq K_1(r_L^{**}, r_H^{**})$.

We then compare $K_0(r_L^{**}, r_H^{**})$ and $K_0(r_L^*, r_H^*)$. Consider $K = K_0(r_L^*, r_H^*)$. We have $k_H^* = \frac{K}{N_H}$ and $r_H^* - \chi(k_H^*) = r_L^* - \chi(0) = r_L^*$.

We have $K < K_0(r_L^{**}, r_H^{**})$ if and only if at K all capital is allocated to type H financiers under moral hazard. This is equivalent to $r_H^{**} - \chi(K/N_H) \geq r_L^{**}$. As $K/N_H = \chi(k_H^*) = r_H^* - r_L^*$, this condition is equivalent to $r_H^{**} - r_L^{**} \geq r_H^* - r_L^*$, which is equivalent to $c \leq \tilde{c}$ using Proposition 3. This concludes the proof of Lemma 3.

We now turn to proving Corollary 2. We prove the corollary by distinguishing the following four different (and exhaustive) possibilities:

(i) The type L financier receives no capital both under complete information and under moral hazard. Then we have $k_H^{**} = k_H^* = K/N_H$ and $k_L^{**} = k_L^*$ and the result is (weakly) true, regardless of c.

(ii) The type L financier receives capital under moral hazard, but not under complete information. Hence we have $k_H^* = K/N_H$ and $k_L^* = 0$, whereas $k_H^{**} \leq K/N_H$ and $k_L^{**} \geq 0$, so that $k_H^{**} - k_L^{**} \leq k_H^* - k_L^*$. Moreover this case requires $K \geq K_0(r_L^{**}, r_H^{**})$ and $K \leq K_0(r_L^*, r_H^*)$, hence $K_0(r_L^{**}, r_H^{**}) \leq K_0(r_L^*, r_H^*)$. Lemma 3 shows that this is true if and only if $c \geq \tilde{c}$. The corollary thus holds in this case.

(iii) The type L financier receives capital under complete information, but not under moral hazard. This is the converse of case (ii): We have $k_H^{**} = K/N_H$ and $k_L^{**} = 0$, whereas $k_H^* \leq K/N_H$ and $k_L^* \geq 0$, so that $k_H^{**} - k_L^{**} \geq k_H^* - k_L^*$. Moreover this case requires $K \leq K_0(r_L^{**}, r_H^{**})$ and $K \geq K_0(r_L^*, r_H^*)$, hence $K_0(r_L^{**}, r_H^{**}) \geq K_0(r_L^*, r_H^*)$. Lemma 3 shows that this is true if and only if $c \leq \tilde{c}$. The corollary thus holds in this case.

(iv) The type L financier receives capital both under complete information and under moral hazard. There are three subcases.

(iv-a) There is no allocation to the risk-free asset both under complete information and under moral hazard. Then we can write:

$$r_H^{**} - r_L^{**} = \phi(k_H^{**}), \tag{A.13}$$

$$r_H^* - r_L^* = \phi(k_H^*), \tag{A.14}$$

$$\text{with } \phi(k) = \chi(k) - \chi\left(\frac{K - N_H k}{N_L}\right). \tag{A.15}$$

We have $\phi'(k) \geq 0$. Thus, $k_H^{**} \geq k_H^*$ if and only if $r_H^{**} - r_L^{**} \geq r_H^* - r_L^*$, which using Proposition 3 is true if and only of $c \leq \tilde{c}$. Moreover, as $N_H k_H^{**} + N_L k_L^{**} = N_H k_H^* + N_L k_L^* = K$, $k_H^{**} \geq k_H^*$ is equivalent to $k_L^{**} \leq k_H^*$ and hence $k_H^{**} - k_L^{**} \geq k_H^* - k_L^*$.

(iv-b) There is no allocation to the risk-free asset under complete information, but there is under moral hazard. In this case, we have $k_H^{**} = \chi^{-1}(r_H^{**})$, $k_L^{**} = \chi^{-1}(r_L^{**})$, and r_H^* and r_L^* satisfy (A.14). Fixing $r_H^* - r_L^*$, by (A.14) and (A.15), we have

$$\frac{\partial k_H^*}{\partial K} = \frac{\chi'(k_L^*)}{N_L \chi'(k_H^*) + N_H \chi'(k_L^*)} \tag{A.16}$$

Notice that

$$\begin{aligned} \frac{\partial}{\partial K}(k_H^* - k_L^*) &= \frac{\partial}{\partial K}\left(\left(1 + \frac{N_H}{N_L}\right) k_H^* - \frac{K}{N_L}\right) \\ &= \frac{1}{N_L}\left(\frac{(N_L + N_H)\chi'(k_L^*)}{N_L \chi'(k_H^*) + N_H \chi'(k_L^*)} - 1\right). \end{aligned} \tag{A.17}$$

If $\chi(k)$ is linear in k, then $\frac{\partial}{\partial K}(k_H^* - k_L^*) = 0$ and $k_H^* - k_L^*$ is a constant. In particular, at $K = K_1(r_H^*, r_L^*)$, we have $k_H^* - k_L^* = \chi^{-1}(r_H^*) - \chi^{-1}(r_L^*)$. Thus, by continuity, we have $k_H^* - k_L^* = \chi^{-1}(r_H^*) - \chi^{-1}(r_L^*)$ for $K_0(r_H^*, r_L^*)) \leq K \leq K_1(r_H^*, r_L^*)$, and we only need to compare $\chi^{-1}(r_H^{**}) - \chi^{-1}(r_L^{**})$ with $\chi^{-1}(r_H^*) - \chi^{-1}(r_L^*)$, which is reduced to case (iv-c) that we discuss below.

(iv-c) There is a positive allocation to the risk-free asset both under complete information and under moral hazard. In this case, we have $k_H^* = \chi^{-1}(r_H^*)$, $k_L^* = \chi^{-1}(r_L^*)$, $k_H^{**} = \chi^{-1}(r_H^{**})$, and $k_L^{**} = \chi^{-1}(r_L^{**})$. Thus $k_H^{**} - k_L^{**} > k_H^* - k_L^*$ is equivalent to $\chi^{-1}(r_H^{**}) - \chi^{-1}(r_L^{**}) > \chi^{-1}(r_H^*) - \chi^{-1}(r_L^*)$. By Proposition 3, to prove this corollary, we need to show that $\chi^{-1}(r_H^{**}) - \chi^{-1}(r_L^{**}) > \chi^{-1}(r_H^*) - \chi^{-1}(r_L^*)$ is equivalent to $r_H^{**} - r_L^{**} > r_H^* - r_L^*$, which essentially requires $\chi^{-1}(\cdot)$ to be a distance-preserving mapping (isometry). By the Mazur-Ulam theorem (see Lax (2014) pp. 49, Theorem 12): every bijective isometry between real normed spaces is affine, this result is true if and only if $\chi(k)$ is linear in k, as assumed.

Note that it cannot be the case that there is no allocation to the risk-free asset under moral hazard, but there is under complete information. This would violate Lemma 3.

A.2.7 Proof of Corollary 3

The corollary follows directly from (25), (26), and Corollaries 1 to 2. The conditions of the corollary isolate a case in which all capital is allocated to the financiers, so that effect (iii) is null. Moreover, as $c \leq \hat{c}_L$, moral hazard shifts the allocation towards the type H. Since $x_H^{**} < x_L^{**}$, $x_H^{**} < x_H^*$, and $x_L^{**} < x_L^*$, this shift increases the demand for labor via effect (ii). Effect (i) always increases the demand for labor.

A.2.8 Proof of Lemma 2

The proof follows directly from (31) and (32).

A.2.9 Proof of Proposition 4

Case (i) is obvious since $k_H^{**} = K/N_H$ and $k_L^{**} = 0$ in that case. In Case (iii), k_H^{**} and k_L^{**} are determined by $r_H^{**} = \chi(k_H^{**})$ and $r_L^{**} = \chi(k_L^{**})$, then Lemma 2 implies that k_H^{**} and k_L^{**} both increase.

In Case (ii), Lemma 2 characterizes when r_H^{**} increases more than r_L^{**} or when the opposite obtains. Then, one can follow the exact same proof as in case (iv-a) of the proof of Corollary 2 (Section A.2.6) to show that k_H^{**} increases and k_L^{**} decreases if and only if r_H^{**} increases more than r_L^{**}.

A.2.10 Proof of Corollary 4

When $K < K_0(r_L^{**}, r_H^{**})$, $k_H^{**} = K/N_H$ is fixed and hence the second line of (35) is null, while the first line is necessarily negative. This implies that an increase in ϵ has a negative impact on L^{**}.

When $K \in (K_0(r_L^{**}, r_H^{**}), K_1(r_L^{**}, r_H^{**}))$, k_H^{**} is defined by (A.13) and (A.15). Using the implicit function theorem, we have:

$$\frac{\partial k_H^{**}}{\partial \epsilon}|_{\epsilon=0} = N_L \frac{\frac{\partial r_H^{**}}{\partial \epsilon}|_{\epsilon=0} - \frac{\partial r_L^{**}}{\partial \epsilon}|_{\epsilon=0}}{N_L \chi'(k_H^{**}) + N_H \chi'(k_L^{**})}. \tag{A.18}$$

Moreover, as $N_L k_L^{**} + N_H k_H^{**} = K$, we have $N_H \frac{\partial k_H^{**}}{\partial \epsilon}|_{\epsilon=0} = -N_L \frac{\partial k_L^{**}}{\partial \epsilon}|_{\epsilon=0}$. We can then rewrite

(35) as:

$$\begin{aligned}\frac{\partial L^{**}}{\partial \epsilon}|_{\epsilon=0} &= -\frac{\partial x_H^{**}}{\partial \epsilon}|_{\epsilon=0} f(x_H^{**}) N_H k_H^{**} - \frac{\partial x_L^{**}}{\partial \epsilon}|_{\epsilon=0} f(x_L^{**}) N_L k_L^{**} \\ &+ N_H N_L \frac{\frac{\partial r_H^{**}}{\partial \epsilon}|_{\epsilon=0} - \frac{\partial r_L^{**}}{\partial \epsilon}|_{\epsilon=0}}{N_L \chi'(k_H^{**}) + N_H \chi'(k_L^{**})} [F(x_L^{**}) - F(x_H^{**})]. \end{aligned} \tag{A.19}$$

We necessarily have $F(x_L^{**}) - F(x_H^{**}) \geq 0$, so the sign of the second line depends only on (34). When (34) is false, both lines are negative and $\frac{\partial L^{**}}{\partial \epsilon}|_{\epsilon=0}$ is negative. This is in particular the case for a frontier improvement. With a core improvement, the first line is null and the second one positive, so that instead $\frac{\partial L^{**}}{\partial \epsilon}|_{\epsilon=0}$ is positive.

Finally, when $K > K_1(r_L^{**}, r_H^{**})$, k_n^{**} is determined by $r_n^{**} = \chi'(k_n^{**})$, so that:

$$\frac{\partial k_n^{**}}{\partial \epsilon}|\epsilon = 0 = \frac{\frac{\partial r_n^{**}}{\partial \epsilon}|\epsilon = 0}{\chi'(k_n^{**})}. \tag{A.20}$$

We can then rewrite (35) as:

$$\begin{aligned}\frac{\partial L^{**}}{\partial \epsilon}|_{\epsilon=0} &= -\frac{\partial x_H^{**}}{\partial \epsilon}|_{\epsilon=0} f(x_H^{**}) N_H k_H^{**} - \frac{\partial x_L^{**}}{\partial \epsilon}|_{\epsilon=0} f(x_L^{**}) N_L k_L^{**} \\ &+ N_H \frac{\partial r_H^{**}}{\partial \epsilon}|_{\epsilon=0} \frac{1 - F(x_H^{**})}{\chi'(k_H^{**})} + N_L \frac{\partial r_L^{**}}{\partial \epsilon}|_{\epsilon=0} \frac{1 - F(x_L^{**})}{\chi'(k_L^{**})}. \end{aligned} \tag{A.21}$$

The second line is necessarily positive. With a core or a frontier improvement, the first line is null, so that necessarily $\frac{\partial L^{**}}{\partial \epsilon}|_{\epsilon=0}$ is positive.

A.2.11 Proof of Corollary 5

The first two lines of (38) are always positive. If $K < K_0(r_L^{**}, r_H^{**})$, as k_H^{**} does not change, the third line is null. If $K > K_1(r_L^{**}, r_H^{**})$, we have $\frac{\partial k_H^{**}}{\partial \epsilon}|_{\epsilon=0} \geq 0$ and $\frac{\partial k_L^{**}}{\partial \epsilon}|_{\epsilon=0} \geq 0$, so that the third line is positive too.

If $K \in (K_0(r_L^{**}, r_H^{**}), K_1(r_L^{**}, r_H^{**}))$, we use again that $N_H k_H^{**} = N_L k_L^{**}$ and hence $N_H \frac{\partial k_H^{**}}{\partial \epsilon}|_{\epsilon=0} = -N_L \frac{\partial k_L^{**}}{\partial \epsilon}|_{\epsilon=0}$ to rewrite the third line of (38), the capital reallocation effect, as:

$$c N_H \frac{\partial k_H^{**}}{\partial \epsilon}|_{\epsilon=0} \left(\frac{F(x_H^{**}) Q_M(x_H^{**})}{q_H} - \frac{F(x_L^{**}) Q_M(x_L^{**})}{q_L} \right) \tag{A.22}$$

We have $F(x_H^{**})Q_M(x_H^{**}) < F(x_L^{**})Q_M(x_L^{**})$ and $q_H > q_L$, hence the term in brackets is negative. We know from Proposition 4 that $\frac{\partial k_H^{**}}{\partial \epsilon}|_{\epsilon=0} \geq 0$ if and only if (34) holds, which concludes the proof.

A.2.12 Proof of Corollary 6

When $K < K_0(r_L^{**}, r_H^{**})$ we have $\alpha_H^{**} = r_H^{**} - \chi(K/N_H)$. As r_H^{**} increases after machine improvement, so does α_H^{**}. We know from the previous corollary that U_H^{**} also increases, so the first line of (40) is positive. As k_H^{**} is fixed, the second line is null.

When $K > K_1(r_L^{**}, r_H^{**})$, $\alpha_H^{**} = \alpha_L^{**} = 0$. Total welfare is then equal to U^{**}, and we can simply apply Corollary 5.

When $K \in (K_0(r_L^{**}, r_H^{**}), K_1(r_L^{**}, r_H^{**})$, we can follow the proof of Corollary 5 to write the second line of (40) as:

$$cN_H \frac{\partial k_H^{**}}{\partial \epsilon}|_{\epsilon=0} \left(\alpha_H^{**} + U_H^{**} - \alpha_L^{**} - U_H^{**}\right). \tag{A.23}$$

Noting that $\alpha_H^{**} = \alpha_L^{**}$ and $U_H^{**} < U_L^{**}$ and using Proposition 4 gives us the result.

A.2.13 Proof of Corollary 7

This corollary is proved in the main text.

A.2.14 Proof of Proposition 5

Differentiating (58) with respect to $\bar{x}$ and rearranging, we obtain that the return increases in $\bar{x}$ if and only if $\varphi(\bar{x}) \geq 0$, with:

$$\varphi(x) = -c\frac{q_M'(x)}{f(x)}[F(x)Q_M(x) + (1 - F(x))q] - (R(q - q_M(x)) - c)(q - q_M(x))^2. \tag{A.24}$$

Note that $\varphi(x^{**}) > 0$ and as $x > x^{**}$ we necessarily have $q - q_M(x) > 0$ and $F(x)Q_M(x) + (1 - F(x))q > 0$. Hence, a sufficient condition for φ to be decreasing in x is to have $\frac{q_M'(x)}{f(x)}$ increasing in x. The proposition assumes this condition to be true. Then, to have an interior solution $x^u \in (x^{**}, 1)$, we need $\varphi(1) < 0$. This gives us:

$$c \leq \hat{c}_1 = \frac{f(1)Rq^3}{f(1)q^2 - q'_M(1)Q_M(1)}. \tag{A.25}$$

Note in particular that $\hat{c}_1 > 0$. We then need to ensure that solution x^u gives a higher profit than offering a wage of zero and making the worker rely entirely on the machine. This gives us the following condition:

$$\left(R - \tfrac{c}{q - Q_M(\tilde{x})}\right)[F(\tilde{x})Q_M(\tilde{x}) + (1 - F(\tilde{x}))q] \geq RQ_M(1) \tag{A.26}$$

$$\Leftrightarrow \quad c \leq \tfrac{R(q - q_M(\tilde{x}))\int_{\tilde{x}}^1 (q - q_M(x))f(x)dx}{F(\tilde{x})Q_M(\tilde{x}) + (1 - F(\tilde{x}))q} = \hat{c}_2. \tag{A.27}$$

Note that $\hat{c}_2 > 0$. Defining $\hat{c} = \min(\hat{c}_1, \hat{c}_2)$, we have $\hat{c} > 0$ and for $c < \hat{c}$ the unique solution to the financier's problem is to offer wage $\tilde{w}(x^u)$ to the worker, so that the worker exerts effort on all tasks above x^u, and uses the machine on all tasks below.

A.2.15 Proof of Corollary 8

We compute the following quantities:

$$\begin{aligned} \frac{\partial U^u}{\partial \epsilon}\Big|_{\epsilon=0} &= \frac{c}{[q - q_M(x^u)]^2}\Bigg(\eta(x^u)[F(x^u)Q_M(x^u) + (1 - F(x^u))q] \\ &\quad + (q - q_M(x^u))\int_0^{x^u} \eta(x)f(x)dx\Bigg) \geq 0 \qquad \text{(A.28)} \\ \frac{\partial U^u}{\partial x^u} &= \frac{cq'_M(x^u)}{[q - q_M(x^u)]^2}\left(F(x^u)Q_M(x^u) + (1 - F(x^u))q\right) \leq 0. \qquad \text{(A.29)} \end{aligned}$$

Finally, using the implicit function theorem, $\frac{\partial x^u}{\partial \epsilon}|_{\epsilon=0}$ is positive if and only if $\frac{\partial \varphi(x^u)}{\partial \epsilon}|_{\epsilon=0}$ is positive. This is equivalent to:

$$\begin{aligned} - &\quad \tfrac{c\eta'(x^u)}{f(x^u)}[F(x^u)Q_M(x^u) + (1 - F(x^u)q] - \tfrac{cq'_M(x^u)}{f(x^u)}\int_0^{x^u} \eta(x)f(x)dx \\ + &\quad R\eta(x^u)(q - q_M(x^u))^2 + 2\eta(x^u)(q - q_M(x^u))[R(q - q_M(x^u)) - c] \geq 0. \qquad \text{(A.30)} \end{aligned}$$

All the terms in this condition are positive except potentially the first one. Denoting $\hat{\eta} > 0$ the value of $\eta'(x^u)$ such that (A.30) holds with an equality, we have that $\frac{\partial x^u}{\partial \epsilon}|_{\epsilon=0} \geq 0$ if and only if $\eta'(x^u) < \hat{\eta}$.