Will Asset Managers Survive the Advent of Robots? An Optimal Contracting Approach

November 15, 2020

Abstract

In this paper, we study the adoption of automation technology in asset management. We build a principal-agent model in continuous time in which delegation of asset management to an agent is subject to moral hazard and will become automatable at an uncertain time. While the characteristics and the advent of the automation technology are exogenous and publicly observable, automation may not be as efficient as the agent. We derive an optimal long-term contract that adjusts the provision of incentives to the availability of such a technology so that automation impacts the agent since the contracting date. Our model suggests that the empirically observed layoffs that accompany the emergence of an automation technology may have a contractual foundation. Compared to the situation where automation is never feasible, we predict that (1) some poor performers are kept employed longer only to be instantaneously substituted at the technology advent, and (2) bonuses are front-loaded and then dropped once the technology becomes available.

I. Introduction

It is widely recognized that automation technologies are about to replace many high-skill workers in the years to come (see, e.g., Brynjolfsson and Mcafee, 2014, or Ford, 2016). However, there is currently "no clear consensus on how automation should be conceptualized and modeled" according to Acemoglu and Restrepo (2018b). We claim that contract theory can offer a comprehensive theoretical framework to address this concern, especially when the presence of agency friction and the opportunity to automate interact.

In this paper, we focus on the asset management industry, where automation's benefit is apparent. Indeed, both economists and practitioners acknowledge that strong incentives are necessary to motivate asset managers because this industry suffers from severe agency frictions (Ben-David, Birru, and Rossi (2020)). In this context, we examine the following fundamental economic questions: How to design an optimal contract that embeds the advent of the automation technology? How does such a contract incentivize the agent? When does the investor decide to automate?

To answer these questions, we build a principal-agent model in continuous time. We consider that a representative investor (hereafter, a principal) delegates the long-term management of an asset to an agent (hereafter, also referred to as an asset manager), and that such a task will become automatable by a technology (hereafter, a robot) at an uncertain time. A moral-hazard problem arises due to the unobservability of the agent's effort at managing the asset. To align the objectives of both parties, the principal offers to the agent a long-term contract under full commitment.

As it is standard in dynamic contracting models, the principal provides incentives to the agent both (i) by deferring compensation, so lump-sum payments are given only when the agent has performed sufficiently well, and (ii) by the threat of terminating the contract after the observation of too many bad outcomes. The novelty of our model is that at the contracting date, the advent of an automation technology (hereafter a robot) capable to replace the asset manager is foreseen. We assume that both parties take as given the characteristics and the arrival time of the robot, so the development of such a technology is not internalized.

Using this model, we show that the principal adjusts the provision of incentives according to the availability of automation, so that the contract is impacted by this opportunity even before it

emerges. It entails leniency for bad performance prior to the emergence of the technology, so it alleviates the risk of contract termination at this stage when no valuable alternative to the agent exists. The principal then reassesses delegation to the agent when the automation technology becomes available, which puts the agent at greater risk of termination in this new contractual environment. This leads to the agent's instantaneous automation if he has not performed sufficiently well until the technology emergence. It is noteworthy that such an adjustment of the provision of incentives increases the contract's overall efficiency but does not alter its incentive compatibility to induce effort. The foreseeable nature of the automation technology in our model lets us then generate multiple implications.

Our main prediction is the "performance-biased automation" of asset managers, i.e., the impact of the advent of technology on the contract depends on the agent's history of performance of managing the asset. This prediction offers a potential explanation of BlackRock's decision to substitute A.I. algorithms for 7 out of their 53 stock pickers in 2017.[1] According to its CEO Laurence D. Fink, BlackRock automated stock pickers because it had to undertake a necessary "change [of] ecosystem", yet (i) it seems puzzling why a significant part of the team was replaced by technology at once, and (ii) the criterion used to make the decision remain unknown. Our model suggests that Blackrock has lessen the risk of contract termination of the stock pickers prior the emergence of the technology, only to substitute technology in place of the agents who did not perform sufficiently well up until the technology advent. We also predict that the bonuses of the remaining stock pickers should decline afterward and that BlackRock did not substitute technology for 100% of stock pickers because technology was not yet as efficient as the agent at managing the asset. Furthermore, the 2018 *asset management compensation study* by Greenwich Associates documents a drop in the annual bonuses in the industry. Its authors claim that the large cost of investment in automation technologies is "eating out" of the compensation pool of asset managers. Our optimal contracting approach offers an alternative explanation. We show that even in absence of cash constraints, the adjustment of the provision of incentives that puts the agent more at threat of termination following the emergence of the technology lets the principal reduce the payments

[1] https://www.nytimes.com/2017/03/28/business/dealbook/blackrock-actively-managed-funds-computer-models.html

to the agent. Consistent with Philippon (2018), the emergence of *fintech* innovations (such as automation technologies) may then decrease the cost of financial intermediation, even when agents are remaining active.

We also derive several cross-sectional implications. If the automation technology is more profitable to the principal, then the agent must be more sensitive to its emergence. This worsens the decline in bonus following the advent of robots and accelerates the automation of asset management. Hence, the better adaptation of a firm in the asset management industry to the forthcoming automation technology – that makes automation more valuable – is associated with (1) a larger decline in compensation following its emergence and (2) a shorter implementation lag (the time interval between the emergence of the technology and its adoption). According to Brynjolfsson, Rock, and Syverson (2017) the implementation lag is the main reason we do not observe a more significant contribution of A.I. in the economy. Our testable predictions may help understand the difference in firms' responses to the emergence of automation technology in the asset management industry.

When no automation technology is foreseen, the baseline model is à la DeMarzo and Sannikov (2006) with effort[2]. Indeed, our optimal contract after the advent of robots is similar to that derived in the cited paper. To solve our problem at the contracting date, we adopt the method introduced by Sannikov (2008) in the context of principal-agent models and based on the martingale optimality principle. Both (i) the asset manager's continuation value (the total value the manager expects to obtain from a contractual relationship) and (ii) the process that accounts for the advent of the automation technology are relevant state processes. The principal controls the sensitivity of the agent to these processes to achieve optimality. As the technology emergence changes the profitability of terminating the contract, it is optimal to make the agent sensitive to the advent of a robot.

Unlike the case of standard dynamic contracting models without a persistent shock such as that of DeMarzo and Sannikov (2006), the implementation of an optimal contract cannot be achieved

[2]Although DeMarzo and Sannikov (2006) build a cash-flow diversion model, they extend the setting to a hidden binary effort choice in Section 3. The authors show that both models lead to the same optimal contract.

through standard securities, but rather through a point-based incentive programme that features the expiration of a number of points at the advent of technology. This number of points traces the agent's continuation value that (1) fluctuates with the asset's performance and (2) is sensitive to the advent of the automation technology. After sufficiently good performance, some points can be redeemed and thus converted into lump-sum payments. In accordance with the contract signed at date 0, a fraction of the points owned by the asset manager expires at the advent of the automation technology, and the contract is terminated as soon as the manager holds no more points in the programme. We note that compared to a scenario where the agent cannot be made sensitive to the advent of robots, such a programme makes more payments in the state of the world where there is no valuable alternative, i.e. before the advent of robots.

We acknowledge that it may be difficult to enforce a contract contingent on a forthcoming automation technology. On the one hand, its advent can be considered industry-wide and thus observable, and for instance Frey and Osborne (2017) rank occupations by their propensity to be automated. On the other hand, the value of implementing a robot-driven process is harder to assess because it is probably firm-specific[3], as suggested by Brynjolfsson et al. (2017). Consequently, it may not be verifiable by a third-party, e.g., a judge. Hence, we also discuss the situation where it is impossible to make the agent sensitive to the advent of technology, while such an exogenous event still occurs.

Finally, we extend our model to examine richer settings. First, we let the principal invest prior to the contracting date to increase the value of the forthcoming automation technology. This may be interpreted as a complementary investment and structural changes that are a prerequisite to an efficient usage of artificial intelligence according to Brynjolfsson et al. (2017). There also exists evidence that the asset management industry participates directly in the innovation process – that may lead to the advent of the automation technology – through investments in fintech[4]. We show that, if the automation technology is more efficient than the agent at managing the asset, then we can separate the investment problem and the contracting problem. Consequently, (1) the contracting problem does not distort the level of investment compared to the first-best investment

[3]We assume that the principal takes the value of automation as given in our main model.

[4]See Boston Consulting Group's study "Fintech in Capital Markets: A Land of Opportunity", 2016

scenario, and (2) both the investment decision and the contract do not interact.

Second, we extend our model and assume that after the technology emergence, the principal can either replace the agent as in the main model, or enhance the agent's productivity, both alternatives being mutually exclusive. This extension suggests the *hollowing out* of asset managers given their performance. Indeed, the continuation value of poor performers decreases to put them at a greater risk of termination and the one of the good performers increases to put them at a greater chance of enhancement at the technology advent. Consequently, the propensity of the advent of automation technology to polarize jobs may have a contractual foundation. In the literature, Autor and Dorn (2013) studies the skill-related polarization due to technological change, while Goos and Manning (2007) and Goos, Manning, and Salomons (2014) investigate the task-based approach, so routine jobs are more at risk of automation. Our theoretical framework would suggest to test for the performance-biased polarization of high-skill workers.

Our study is closely related to those of Hoffmann and Pfeil (2010), Demarzo, Fishman, He, and Wang (2012) and Li (2017) who theoretically investigate how an exogenous shock to the agent's profitability impacts the optimal compensation. They show that the agent's continuation value reacts instantaneously to lucky events and that *rewards for luck* are part of the optimal compensation scheme. This is consistent with the findings of Garvey and Milbourn (2006), Bertrand and Mullainathan (2001), and Francis, Hasan, John, and Sharma (2013) who show that exogenous events significantly impact the compensation of CEOs and VPs. In contrast, we study an exogenous shock to the contract's termination value, where the automation technology can be regarded as a real option held by the principal. Consequently, we also look for the optimal time to implement the technology. It turns out that it is optimal to wait for the manager's continuation value to reach the fixed termination boundary before implementing the robot. However, the principal may design the contract so that poor performers would see their continuation value drop to zero instantly with the advent of the automation technology.

He (2009) also builds on DeMarzo and Sannikov (2006) but specifies that the cash-flow process follows a geometrical Brownian motion, so changes in firm size generate incentives. Gryglewicz, Hartman-Glaser, and Zheng (2019) present a model where a firm is managed by a risk-averse agent

and where the principal holds a growth option. The agent's risk aversion implies compensation for bearing risk. The study of a growth option is relevant in the author's framework as the cash flow process follows a geometric Brownian motion, so the firm size matters. Grenadier and Wang (2005) study a model with an investment option that encompasses both a moral hazard and adverse selection. They show that the agency problem leads to a postponement of the investment decision made by the agent.

Finally, our study tries to bridge the gap between the literature on the contracting theory and that on labour economics, where the impact of automation on jobs is broadly investigated (see Acemoglu and Autor (2011) for an overview of this literature). Consistently with the model of Acemoglu and Restrepo (2018b), we capture the impact of automation "at the extensive margin" as a single task – namely, asset management in our model – becomes at risk of automation. While it is well known in agency theory that contracts are contingent on their environment, there is a consensus in labour economics on the prominent role of technological change in the shape of wages and in polarization of the job market. To the best of our knowledge, no study examines however the interaction between the presence of financial frictions (such as the firm's cash constraint or agency friction) and automation. A body of the literature investigates the propensity of different kinds of jobs to be automated (see, e.g., Frey and Osborne (2017), Arntz, Gregory, and Zierahn (2016)). Another stream of paper (e.g., Autor and Dorn (2013), Goos and Manning (2007), Goos et al. (2014)) adopts the task-based approach introduced by Autor, Levy, and Murnane (2003), and assumes the current substitutability of robots for routine tasks while in our paper we investigate the foreseen "computerisation in a non-routine task" (Frey and Osborne (2017))[5]. Among several studies, Acemoglu and Restrepo (2020) follow this approach and develop a general equilibrium model to estimate the impact of robots on wages and employment in the U.S. labour market.

This paper is divided into five sections. Section 2 provides an overview of the model and solves for the first-best benchmark. We solve for an optimal contract in Section 3. Section 4 presents an implementation of the optimal contract and empirical implications. In Section 5, we offer two extensions of our main model.

[5]Others approaches to model automation include a factor-augmenting technological change (e.g.,human-augmenting change as in Bessen (2017)) or capital-augmenting technologies in Graetz and Michaels (2018)).

II. Model

We build upon a principal-agent model in continuous time *à la* DeMarzo and Sannikov (2006) with effort. We investigate the problem of a representative investor (hereafter, the principal) who delegates the management of an asset to an agent (hereafter, also referred to as the asset manager) who may face the risk of automation with a forthcoming technology. While the relationship is subject to moral hazard, the implementation of automation technology (hereafter, the robot) irreversibly substitutes technology for the agent and thus terminates the agency friction. The principal has unlimited wealth, and the agent is protected by limited liability. Both parties fully commit to a long-term contract that characterizes the terms of the relationship. First, we present the agency problem where the agent's effort impacts the dynamics of the asset value. Second, we describe the technology that can replace the agent. Then, we formulate the principal's problem and solve for the first-best scenario. Finally, we describe the set of incentive-compatible contracts that satisfies the limited liability condition.

A. *Agency Problem*

If the principal delegates the management of the asset to the agent, a moral hazard problem arises due to the unobservability of the agent's effort $a_t \in \{0; \bar{a}\}$ applied to control the performance of the asset.[6] Specifically, the value of the asset under the agent's management, and which is observable by the principal, evolves according to the dynamics

$$dX_t = a_t \mu dt + dZ_t^a \tag{1}$$

where μ is a positive constant that accounts for the asset's specific parameter of profitability, and $(Z_t^a)_t$ is a Brownian motion. Details on the probabilistic background of the model are provided in the appendix B. Shirking by the agent obviously has a negative impact on the value of the asset under management, and lets the agent receive a private benefit Bdt, where B is a positive constant. In our setting, the principal is risk-neutral, and we denote by $r > 0$ the principal's discount rate.

[6] At each instant, the agent either exerts effort ($a_t = \bar{a}$) or shirks ($a_t = 0$).

B. Automation Problem

While we do not initially consider any competitor to the agent, both parties foresee prior to the contracting date the advent of the automation technology that can manage the asset and thus replace the agent. Such a technological change impacts the contractual environment,[7] and we anticipate based on the existing results in contracting theory[8] that the characteristics of the optimal contract are contingent on this shock.

Let us assume that the robot becomes available at a random time denoted by $\tilde{T}$, and the principal can then decide to automate the management of the asset at any time. The termination of the agent's contract is a prerequisite, and the substitution is assumed to be irreversible. Thus, the automation technology can be regarded as a real option. According to Acemoglu and Restrepo (2018b), modelling automation as an irreversible substitution device is "both descriptively realistic and leads to distinct and empirically plausible predictions".

We study automation at "the extensive margin"[9], so the single task in this model – asset management – will become automatable. We assume that the advent of technology able to replace the agent follows a single-jump process $N = \{N_t\}_{t \geq 0}$ with intensity λ.[10] Specifically, λdt is the probability of the automation technology arising during any time interval $(t, t + dt]$ $\forall t < \tilde{T}$, and such an event will make the process N jump from 0 to 1.[11] We denote by $\mathcal{H} = \{\mathcal{H}_t\}_{t \geq 0}$ the filtration generated by the single-jump process.

Once the principal automates asset management, the asset has a net perpetual profitability of

[7] There is a consensus in labour economics that technological change is a prominent cause of changes on the job market (see,e.g.,Acemoglu and Autor (2011)).

[8] see, e.g., the papers on reward for luck such as Hoffmann and Pfeil (2010) and the extension of Demarzo et al. (2012).

[9] Alternatively, we could model automation "at the intensive margin", where the ability of the robot to manage assets would improve over time. In this setting, a Poisson process could model the dates of technological change. While it makes the model more complex, we claim that our main results are robust to such a change.

[10] We claim that while routine tasks are already automatable (see,e.g.,Autor, Dorn, and Hanson (2013)), non-routine tasks such as asset management will be automatable in the future.

[11] If the agent was instead competing with other high-skill individuals on a scarce job market, then a birth-death process could model the presence of competitors(for a reference in Queuing Theory, see Brémaud (1981)).

$m\mu dt$ per unit of time.[12] While μ is the asset's specific parameter of profitability, m denotes the technology-specific parameter that accounts for the automation efficiency. By construction, the principal's expected value associated with a robot-driven asset and observed as of $s \geq \tilde{T}$ satisfies

$$M = \mathbb{E}_{s \geq \tilde{T}} \left[\int_{s}^{+\infty} e^{-r(t-s)} (m\mu dt + dZ_t) \right],$$

where we set $M = \frac{m\mu}{r}$. If the automation technology is available once the principal terminates the contract, it is implemented with no delay. At any date prior to $\tilde{T}$, the principal must wait for the technology's advent, and the principal obtains the following from an expected automation at $\tilde{T}$

$$\frac{\lambda}{\lambda + r} M = \mathbb{E}_{s < \tilde{T}} \left[\int_{\tilde{T}}^{+\infty} e^{-r(t-\tilde{T})} (rMdt + dZ_t) \right].$$

Consequently, the advent of the automation technology at $\tilde{T}$ makes the value of contract termination jump from $M_0 = \frac{\lambda}{\lambda+r} M$ to $M > M_0$.

In our setting, both parties take as given the characteristics and the arrival time of the robot. This describes a situation where the principal does not internalize the development of robots. Hence, the presence of the automation technology does not create a new source of information asymmetry, and thus contrasts with the unobservability of the agent's effort that leads to moral hazard.

C. *Formulation of the Principal's Problem*

On date 0, the principal offers a long-term contract under full commitment to the agent who may be replaced with a robot in the future. Such a contract $\Pi = \{U; \tau\}$ consists of payments $U = (U_t)_{t \leq \tau}$ and a date of termination τ, which are based on the observed performance of the asset

[12] Note that switching to a robot-driven asset management may encompass a positive sunk cost, but the principal has unlimited wealth and such a cost can be seen as deterring the perpetual profitability without loss of generality. Furthermore, switching to a robot-driven management does not impact the exposition to the Brownian component.

and depend on the advent of the automation technology. Formally, we regard it is as contingent on the total information set that is thus represented by the joint filtration $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$ available to the principal on any date t. Process U is $\mathcal{G}$-adapted, finite, non-decreasing (due to limited liability), and is measured in the same unit as the agent's private benefit. The termination date τ is a measurable $\mathcal{G}$-stopping time that can be infinite.[13] We assume that the agent is risk-neutral and that the agent discounts at $\gamma > r$, thus being more impatient than the principal. Consequently, the payments to the agent cannot be perpetually postponed in an optimal contract.

Next, we present the agent's value and the principal's value for a given contractual relationship. We fix an arbitrary contract Π and the agent's effort strategy $a = (a_t)_t$. We assume that at each instant t, the agent's decision on a_t is made prior to the observation of N_t and the realization of Z_t, so $(a_t)_t$ is $\mathcal{G}$-predictable. Then, the agent expects to extract from the contractual relationship a total expected payoff[14]

$$\mathbb{E}^a \left[\int_0^{\tau^-} e^{-\gamma t}(dU_t + \frac{B}{\bar{a}}(\bar{a} - a_t)dt) + e^{-\gamma \tau} \Delta U_\tau \right] \tag{2}$$

where $\frac{B}{\bar{a}}$ represents the severity of the agency problem in our model. An effort process $a = (a_t)_t$ generates a unique probability distribution, and we denote by $\mathbb{E}^a$ the associated expectation operator. The long-term contract embeds a terminal payment $\Delta U_\tau = U_\tau - U_{\tau^-} \geq 0$ to let the principal fulfil any remaining contractual promise to terminate the contract. Indeed, there is nothing that prevents the principal from terminating the contractual relationship on any date as long as the principal pays any amount owed to the agent. Afterwards, the agent benefits from an outside option, the value of which we normalize to 0.

[13] For instance, Section A3.3 in Daley and Vere-Jones (1989) defines the concepts of adaptability and measurability as well as that of predictability used hereafter for the effort process.

[14] For any two real numbers s and t, we denote by $\int_s^t$ an integral over $[s, t]$ and by $\int_s^{t^-}$ an integral over $[s, t)$.

The principal's expected value is

$$\mathbb{E}^a\left[\int_0^{\tau^-} e^{-rt}(a_t\mu dt - dU_t) + e^{-r\tau}(\tilde{M}_\tau - \Delta U_\tau)\right] \tag{3}$$

where $\tilde{M}_\tau = M_0 1_{N_\tau=0} + M 1_{N_\tau=1}$ is the value of either waiting for the advent of the automation technology if $N_\tau = 0$ or implementing the robot-driven asset management if $N_\tau = 1$.

An *incentive-compatible* effort process is the agent's best response to a given contract Π, so it satisfies

$$a^*(\Pi) = \arg\max_a \mathbb{E}^a\left[\int_0^{\tau^-} e^{-\gamma t}(dU_t + \frac{B}{\bar{a}}(\bar{a} - a_t)dt) + e^{-\gamma\tau}\Delta U_\tau\right] \tag{4}$$

Following the literature's standard approach, we extend this notion and we call Π an incentive-compatible contract if it induces the agent to exert effort according to an incentive-compatible effort process $a = (a_t^*(\Pi))_t$.

Hence, the principal's problem is to determine an optimal contract (if it exists), i.e., an incentive-compatible contract that maximizes the principal's value on date 0 and that delivers to the agent the agent's reservation value noted $w_0 > 0$. Formally, this can be formulated as

$$\sup_{\Pi \text{ I.C.}} \mathbb{E}^{a^*(\Pi)}\left[\int_0^{\tau^-} e^{-rt}(a_t^*(\Pi)\mu dt - dU_t) + e^{-r\tau}(\tilde{M}_\tau - \Delta U_\tau)\right] \tag{5}$$

such that

$$\mathbb{E}^{a^*(\Pi)}\left[\int_0^{\tau^-} e^{-\gamma t}(dU_t + \frac{B}{\bar{a}}(\bar{a} - a_t^*(\Pi))dt) + e^{-\gamma\tau}\Delta U_\tau\right] \geq w_0. \tag{6}$$

Inequality (6) is called the agent's participation constraint.

D. First-Best Case : Optimal Contract without Information Asymmetry

In this section, we focus on the so-called first-best framework, where we relax the information asymmetry. Here, the principal chooses directly the effort process $a^{FB} = (a_t^{FB})_t$ exerted by the agent, and designs a first-best contract $\Pi^{FB} = ((U_t)_t, \tau)$[15] that maximizes the principal's value while providing to the agent at least the agent's reservation value. Hence, this problem can be formulated as

$$\sup_{\Pi^{FB}, a^{FB}} \mathbb{E}\left[\int_0^{\tau^-} e^{-rt}(a_t^{FB}\mu dt - dU_t) + e^{-r\tau}(\tilde{M}_\tau - \Delta U_\tau)\right] \tag{7}$$

such that

$$\mathbb{E}\left[\int_0^{\tau^-} e^{-\gamma t}(dU_t + \frac{B}{\bar{a}}(\bar{a} - a_t^{FB})dt) + e^{-\gamma\tau}\Delta U_\tau\right] \geq w_0. \tag{8}$$

As postponing payments is costly and there is no need to generate incentives in the first-best framework, all payments are front-loaded. We have $U_0 = w_0$ together with $(U_t)_{t>0} = 0$, so the participation constraint is binding.

In the baseline model of DeMarzo and Sannikov (2006) that does not embed the advent of the automation technology, the first-best contract always terminates at $\tau = +\infty$. Here, the principal may be better off terminating the agent's contract if the agent does not perform as well as the robot at managing the asset, which is the case if $m > \bar{a}$, or equivalently $M > \frac{\bar{a}\mu}{r}$. Therefore, we derive a first-best contract that stochastically terminates at the advent of the automation technology at $\tilde{T}$ if $M > \frac{\bar{a}\mu}{r}$. The following proposition summarizes our results.

Proposition 1: *In absence of agency conflict, the principal's value is*

$$\begin{cases} \frac{\bar{a}\mu}{r} - w_0 & \text{if } M < \frac{\bar{a}\mu}{r} \\ \frac{1}{\lambda+r}\bar{a}\mu + \frac{\lambda}{\lambda+r}M - w_0 & \text{otherwise} \end{cases} \tag{9}$$

[15]A first-best contract consists of a stream of payment $(U_t)_t$ and a stopping-time τ that satisfy the usual conditions provided in section II.C.

We note that the principal only offers a contract to the agent if it generates a positive pledgeable income, i.e., only if $\frac{\bar{a}\mu}{r} > w_0$ provided that $M < \frac{\bar{a}\mu}{r}$, and if $\frac{1}{\lambda+r}\bar{a}\mu > w_0$ otherwise.

E. *Incentive-Compatible Contract with Limited Liability*

In this section, we follow the techniques introduced by Sannikov (2008) in the context of principal-agent models and apply those techniques to characterize the incentive-compatibility effort strategies in the presence of information asymmetry. For now, let us consider an incentive-compatible contract Π, so the effort strategy $a = (a_t)_t$ satisfies equation (4). Then, we define the agent's continuation value $W^\Pi = (W_t^\Pi)_t$ associated with contract Π as

$$W_t^\Pi = \mathbb{E}\left[\int_t^{\tau^-} e^{-\gamma(s-t)}(dU_s + \frac{B}{\bar{a}}(\bar{a} - a_s)ds) + e^{-\gamma(\tau-t)}\Delta U_\tau \mid \mathcal{G}_t\right] \quad \forall t \leq \tau. \tag{10}$$

The agent's continuation value represents the agent's expected earnings on any date. As the stream of payments $(dU_t)_t$ is non-negative and the private benefit is a positive constant, it also holds by construction that $W_t^\Pi \geq 0$ for all $t \leq \tau$. Additionally, we assume that the agent is protected by limited liability, so the contractual relationship has to stop the first time $W_t^\Pi = 0$. In the following lemma, we apply the martingale representation theorem to find a stochastic representation of the agent's continuation value.[16]

Lemma 1: ***Representation of the agent's continuation value as a jump-diffusion process*** *There exists a $\mathcal{G}$-predictable pair of processes $(\beta^\Pi, \delta^\Pi) = ((\beta_t^\Pi)_{t\leq\tau}, (\delta_t^\Pi)_{t\leq\tau})$ such that the agent's continuation value W^Π associated with an incentive-compatible contract Π evolves up to the $\mathcal{G}$-stopping time τ with the dynamics*

$$dW_t^\Pi = (\gamma W_t^\Pi - \frac{B}{\bar{a}}(\bar{a} - a_t))dt + \beta_t^\Pi dZ_t^a + \delta_t^\Pi(dN_t - \lambda dt)1_{t\leq\tilde{T}} - dU_t \qquad \textit{for } t \leq \tau \tag{11}$$

[16] All proofs are provided in the Appendix

Hereafter, β^{Π} is called the process of sensitivity to the asset realization, and $\delta^{\Pi}$is called the process of sensitivity to the advent of the automation technology for an incentive-compatible contract Π.

We interpret the agent's continuation value as a measure of the agent's historical performance at managing the asset for a given contract Π and contingent on the availability of a robot capable of replacing the agent. Indeed, Lemma 1 states that while the agent's continuation value grows steadily at the discount rate γ up to contract termination, it also depends on the asset realization and the robot's availability. This is a standard feature of dynamic agency models with a Brownian motion and persistent shocks[17]. Specifically, every change in the asset realization due to the Brownian motion dZ_t^a is amplified in the agent's continuation value by a factor of β_t^{Π}. Additionally, the drift of the continuation value increases by $-\delta_t^{\Pi}\lambda$ during any time interval $(t, t+dt]$ prior to the advent of the automation technology, while it jumps instantaneously by $\delta_{\hat{T}}^{\Pi}$ at such advent. We note that whenever $-\delta_t^{\Pi}\lambda > 0$, making the agent sensitive to the advent of the automation technology means *boosting* the agent's continuation value and that the term $\delta_t^{\Pi}(dN_t - \lambda dt)$ disappears from the continuation value's dynamics after the said advent. Furthermore, every time the agent enjoys a private benefit because the contract lets the agent shirk and every time the agent receives a payment, the dynamics of the continuation value decreases. We also remark that if the incentive-compatible contract induces the full-effort strategy $a = (\bar{a})_{t\leq\tau}$, the term $\frac{B}{\bar{a}}(\bar{a} - a_t)dt$ disappears.

Next, we characterize the set of contracts that are incentive compatible. To this end, we apply the martingale optimality principle as introduced by Sannikov (2008) in the context of principal-agent models. It enhances the principal able to enforce implicitly any effort strategy by controlling the pair of sensitivity processes $(\beta^{\Pi}, \delta^{\Pi})$. First, we consider a pair of $\mathcal{G}$-predictable processes (β, δ) such that

$$\mathbb{E}^a\left[\int_0^{+\infty} e^{-2rs}\beta_s^2 ds\right] < +\infty, \tag{12}$$

and we also consider a $\mathcal{G}$-adapted, finite and non-decreasing process $U = (U_t)_t$. This lets us define

[17]See Hoffmann and Pfeil (2010) and the extension of Demarzo et al. (2012) for models where the agent's profitability experiences a persistent shock.

the process $W^\alpha = (W_t^\alpha)_t$, where $\alpha = (\beta, \delta, U)$, with the following controlled stochastic differential equation under $\mathbb{P}^{\bar{a}}$

$$\begin{cases} dW_t^\alpha = (\gamma W_t^\alpha + f(\beta_t))dt + \beta_t dZ_t^{\bar{a}} + \delta_t(dN_t - \lambda dt)1_{t \leq \tilde{T}} - dU_t \\ W_0^\alpha \geq w_0 \end{cases} \tag{13}$$

The following lemma characterizes incentive-compatible contracts as a function of the process of sensitivity to the asset realization β that is controlled by the principal.

Lemma 2: ***Incentive compatible contract***
For a given pair of $\mathcal{G}$-predictable processes (β, δ) that satisfies equation (12) and for a given $\mathcal{G}$-adapted, finite and non-decreasing process $U = (U_t)_t$, a contract that induces effort on date t ($a_t = \bar{a}$) if $\beta_t \geq \frac{B}{\bar{a}}$ and lets the agent shirk ($a_t = 0$) otherwise is incentive-compatible. Hereafter, we denote by $\underline{\beta} := \frac{B}{\bar{a}}$ the lowest sensitivity to the asset realization that induces effort, so the incentive-compatible effort strategy is $a_t^*(\beta) = \bar{a}1_{\beta_t \geq \underline{\beta}}$.

It follows from Lemma 2 that setting $\beta_t \geq \underline{\beta} = \frac{B}{\bar{a}}$ induces the agent to exert effort (so $a_t = \bar{a}$) because it ensures that the agent obtains a greater expected value from exerting the effort than instantaneously from shirking. Thus, making the agent sensitive to the advent of a robot does not alter an incentive-compatible effort strategy because it does not depend on the controlled processes of sensitivity $(\delta_t)_{t \leq \tilde{T}}$.

For the process W^α associated with such an incentive-compatible contract to be the agent's continuation value, it must satisfy the limited liability constraint, so $W_t^\alpha \geq 0$ for all $t \leq \tau$. To this end, we first introduce

$$\tau_0^\alpha = inf\{t \geq 0 \mid W_t^\alpha = 0\} \tag{14}$$

and require the contract termination to occur whenever τ_0^α is reached, so $\tau \leq \tau_0^\alpha$. Second, as the agent's continuation value jumps instantaneously by $\delta_{\tilde{T}}$ at the stochastic advent of the automation technology while it must remain nonnegative, the limited liability condition also leads to a lower

boundary on the feasible control $(\delta_t)_{t\leq\tilde{T}}$.[18] To ensure that $W^{\alpha}_{\tilde{T}} \geq 0$, we require that for all t up to $\tau \wedge \tilde{T}$, $\delta_t \geq -W^{\alpha}_{t^-}$, where a $\mathcal{G}$-predictable process. $W^{\alpha}_{\cdot^-} = (W^{\alpha}_{t^-})_t$ is the left-hand limit of the continuation value process $(W^{\alpha}_t)_t$[19]

Hence, any contract such that (1) $\beta_t \geq \underline{\beta}$ up to τ (incentive compatibility) and (2) $\delta_t \geq -W^{\alpha}_{t^-}$ for all t up to $\tau \wedge \tilde{T}$ (limited liability) that implements $W^{\alpha}_0 \geq w_0$ (participation constraint) and that terminates before τ^{α}_0 is a candidate for solving (5)-(6). We derive an optimal contract in the following section[20].

III. Optimal Solution to the Principal's Problem

In this section, we heuristically derive the optimal contract and the associated principal's value function to establish some intuition. Our results are summarized in propositions, and the verification is provided in the appendix. We focus on the characterization of an optimal contract that induces the agent to follow the full effort strategy, so $a_t = \bar{a}$ up to contract termination. Such a restriction is standard in the literature (see, among others, the baseline model of DeMarzo and Sannikov (2006), its extension by Hoffmann and Pfeil (2010) with persistent shock, or Biais et al. (2010) in the case of the agent preventing the occurrence of a large risk), and necessary and sufficient conditions for the optimality of the full-effort strategy are provided in Appendix A.

The optimal contract we derive can be described with two state variables prior to the advent of the automation technology: the single-jump process that accounts on any date for the availability of such technology, and the agent's continuation value. Once such technology has emerged, the continuation value remains the only relevant state variable. As value functions are forward-looking processes, we use backward induction to solve the principal's problem. We will start by characterizing the optimal contract after the advent of the automation technology, and then the optimal contract prior to such advent.

[18] Such a restriction by limited liability is standard in models with Poisson shocks, whether they are controlled by the agent as in Biais, Mariotti, Rochet, and Villeneuve (2010)) or independent as in Hoffmann and Pfeil (2010).

[19] The left-hand limit of any process $(Y_t)_t$ is defined as $Y_{t^-} = \lim_{s\uparrow t} Y_s$ together with $Y_{0^-} = Y_0$.

[20] In the rest of the paper, we get rid of the superscript α

A. Heuristic Derivation

Let us denote by $V_n(w)$ the principal's value function, which refers to the highest value the principal can obtain from delegating to an agent with a current continuation value w and where $n = \{0; 1\}$ specifies whether the automation technology is available ($n = 1$) or not ($n = 0$). As the principal can provide at any moment a lump-sum payment ΔU and then proceed with the optimal contract and a remaining continuation value $w - \Delta U$, the following inequality holds

$$V_n(w) \geq V_n(w - \Delta U) - \Delta U, \qquad \text{with } n = \{0; 1\}. \tag{15}$$

It implies that $V_n'(w) \geq -1, \forall\, w$ and for $n = \{0; 1\}$, so the marginal benefit of providing incentives must remain greater than the marginal value of making a lump-sum payment to the agent. As equation (11) shows, making a payment reflects downwards the agent's continuation value, and deferring a payment mitigates the risk of termination that exists when the continuation value hits zero for the limited liability reason. The agent is more impatient than the principal, so deferring payments is costly and cannot be done in perpetuity. Consequently, we conjecture that the principal's value function is concave. We denote by $\bar{W}^0 = \inf\{w \mid V_0'(w) = -1\}$ and $\bar{W}^1 = \inf\{w \mid V_1'(W) = -1\}$ the lowest continuation values where equation (15) holds with an equality. These thresholds will play the role of the payment barrier in the optimal contract. Unlike the payment barrier in the baseline model of DeMarzo and Sannikov (2006), the payment barrier here is contingent on the state of variable $(N_t)_t$ that accounts for the availability of a robot. Then, the agent is employed in the interval $[0, \bar{W}^0]$ prior to the advent of the automation technology and $[0, \bar{W}^1]$ afterwards. Finally, the principal can terminate the relationship at any instant t only if he is better off paying any amount owed to the agent (the agent's current continuation value w) and obtaining the value $\tilde{M} = M_0 1_{N_t=0} + M 1_{N_t=1}$. Hence, $\forall w$, $V_n(w) \geq \tilde{M} - w$.

Therefore, one expects that the principal's value when delegating to the agent is given by the

following Hamilton-Jacobi-Bellman (HJB) equation:

$$max\left(\tilde{M}-.-v_n,\mathcal{L}_nv_n-rv_n,1+v_n'\right)=0,\text{for } n=\{0;1\} \tag{16}$$

and where $\mathcal{L}_nV_n(w)=$

$$\bar{a}\mu+(\gamma w-\lambda\delta 1_{n=0})\,V_n'(w)+\frac{1}{2}\beta^2V_n''(w)+\lambda(V_1\,(w+\delta)-V_0\,(w))1_{n=0}, \tag{17}$$

where the first term in equation (16) represents the value associated with the opportunity of terminating the contract to automate, the second refers to the value associated with delegating to the agent, and the third states that the slope of the principal's value is always greater than -1. We note that $\tilde{M}-.$ has a slope of -1, while $V_n'>-1\ \forall w$ below the payment boundary. Then, the marginal benefit of delegating to the agent is equal to or greater than that of automation for any strictly positive continuation value. Therefore, whenever the principal is in a contractual relationship with an agent, the principal always prefers to incentivize the agent rather than to pay any amount owed at once in order to automate. We infer that the principal always waits until reaching the stopping time $\tau=\tau_0$ to automate.

Note that whenever the principal offers a contract to the agent, β^2 appears as a multiplier of $V_0''(w)$ in the principal's value function, and $V_0''(w)\leq 0$ since the value function is concave. Hence, the principal sets the lowest sensitivity to the asset realization that induces the agent to exert effort, i.e. $\beta_t=\underline{\beta}$ up to $t=\tau$. The first-order condition with respect to the sensitivity to the advent of the automation technology leads to $V_1'\,(w+\delta)=V_0'\,(w)$ as long as it satisfies the limited liability condition that requires that $\delta\geq -w$.

The following propositions characterize the optimal contract prior to and after the advent of a robot; such a contract is discussed in Section III.B.

Proposition 2: ***Optimal contracting after the advent of the automation technology*** *Suppose that an automation technology is already available and can replace the agent at any moment. The agent's continuation value is the only relevant state variable in the contracting problem, and*

its dynamics under an incentive-compatible contract that implements the full-effort strategy satisfy

$$dW_t = \gamma W_t dt + \underline{\beta} dZ_t^{\bar{a}} - dU_t, \quad \forall \quad t \leq \tau_0 \tag{18}$$

until contract termination that occurs at τ_0 – the first time w hits 0.

The principal's value function is that derived in section III, Proposition 7 of the baseline model of DeMarzo and Sannikov (2006). It is concave and solves the following second-order differential equation

$$rV_1(w) = \bar{a}\mu + \gamma w V_1'(w) + \frac{\beta^2}{2} V_1''(w) \qquad \text{if } w \in [0; \bar{W^1}]; \tag{19}$$

together with $V_1(0) = M$ (the value matching condition); $V_1'(\bar{W^1}) = -1$ (the smooth pasting condition) ; and $V_1''(\bar{W^1}) = 0$ (the super contact condition). The value function extends linearly afterwards with slope -1.

Consequently, whenever the robot is more efficient than the agent at managing the asset, $\bar{W^1} = 0$ and the principal's value function satisfies

$$V_1(w) = M - w \quad \forall w \geq 0. \tag{20}$$

If the technology is already available, the agent's continuation value $(W_t)_t$ is the only relevant state variable, and Proposition 2 characterizes the optimal contract. $(W_t)_t$ evolves within an interval $[0, \bar{W^1}]$ and is subject to uncertainty arising through the Brownian component $(Z_t)_t$ that ties its dynamics with the asset's performance. On the one hand, if the robot is less efficient than the agent at managing the asset, the agent is paid if good performance leads W to reach $\bar{W^1} > 0$, and the contract only stops when poor performance makes W hit zero for the first time. On the other hand, if the robot is at least as good as the agent, the principal would not continue the delegation to an agent subject to a costly agency friction while the principal can automate. Then, $\bar{W^1} = 0$, so the agent is paid the agent's current continuation value at once, and the principal automates. Consequently, the principal who already delegates to an agent prefers to postpone automation if

and only if the available technology is less efficient.

Nevertheless, when the principal has to decide prior the contracting date to either automate from the outset or delegate to an agent, the principal may prefer automation even if the robot is less efficient. As illustrated in Figure 1, this occurs if (1) providing incentives is too costly or (2) the agent's reservation value is too large. First, the presence of agency friction lets the agent extract rent, so the principal cannot internalize the total surplus. Consequently, there exists a threshold $\underline{M} = inf\{M \geq 0 \mid \sup_w V_1(w) = V_1(0)\}$ such that $\forall M \geq \underline{M}$, the principal prefers to automate from the outset. This case is depicted in the left panel of Figure 1. Such threshold depends on how profitable automation is compared with the contract. It is noteworthy that the advantage of automating is greater if the asset intrinsic profitability μ is lower. Indeed, the severity of agency friction captured by $\frac{B}{\bar{a}}$ is independent of μ, so it is less profitable to delegate to an agent if μ is lower. Second, the principal may prefer to automate from the outset because the agent's reservation value is too large. we denote by $\tilde{w_0} = \{w \geq 0 \mid V_1(\tilde{w_0}) = M \quad \& \quad V_1'(\tilde{w_0}) \leq 0\}$, so $\forall w_0 \geq \tilde{w_0}$,[21] the agent is too *expensive* and the principal automates from the outset. The right panel of Figure 1 illustrates this scenario.

Proposition 3: ***Optimal contracting prior to the advent of the automation technology*** *Assume that both parties foresee the advent of the automation technology that occurs at an uncertain date. The dynamics of the agent's continuation value under an incentive-compatible contract that implements the full-effort strategy satisfy*

$$dW_t = \gamma W_t dt + \underline{\beta} dZ_t^{\bar{a}} + \delta(dN_t - \lambda dt) - dU_t, \quad \forall \quad t \leq \tau_0 \wedge \tilde{T}. \tag{21}$$

The principal's value function is concave and is given by the solution to the following second-order differential equation

$$\begin{aligned} &\forall w \in [0; \bar{W^0}], \quad (\lambda + r)V_0(w) = \\ &\qquad \bar{a}\mu + (\gamma w - \lambda\delta)V_0'(w) + \frac{1}{2}\underline{\beta}^2 V_0''(w) + \lambda V_1(w + \delta) \end{aligned} \tag{22}$$

[21] $\forall M \geq \underline{M}, \tilde{w_0} = 0$.

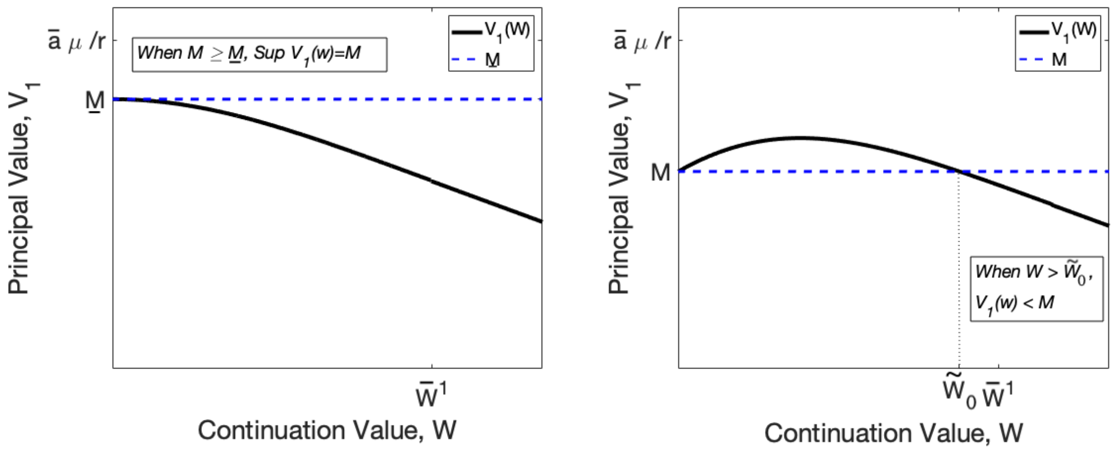

Figure 1. Principal's value function V_1 after the advent of the technology. Parameters are $r = 10\%, \gamma = 12\%, \lambda = 12.5\%, \bar{a} = 10, \mu = 1, B = 0.8, \underline{M} = 91$ (left panel), $M = 80$ (right panel). $\bar{W}^1$ is the payment barrier and $\tilde{W}_0$ is the largest reservation value such that the principal is better off automating from the outset rather than delegating first to the agent.

together with $V_0(0) = M_0$ (the value-matching condition), $V_0'(\bar{W}^0) = -1$ (the smooth-pasting condition), $V_0''(\bar{W}^0) = 0$ (the super-contact condition), and where the value function V_1 is given in Proposition 2. The value function extends linearly afterwards with slope -1. Process δ is given by the first-order condition on $V_0(w)$ if it satisfies the limited liability condition. Therefore, it is determined by $V_1'(W_{t^-} + \delta(W_{t^-})) = V_0'(W_{t^-})$ for all W_{t^-} such that $V_1'(W_{t^-}) \geq V_1'(0)$ and $\delta(W_{t^-}) = -(W_{t^-})$ otherwise (the limited liability condition).

B. *Analysis of the Optimal Contract*

We illustrate the value functions associated with the optimal contract and the optimal sensitivity to the advent of a robot in Figure 2 for $M < \frac{\bar{a}\mu}{r}$. In this case, the value function jumps from V_0 to $V_1 > V_0$ because asset management becomes more valuable in the state of the world where the contract is terminated. As the technology is implemented at the termination boundary $W = 0$, the difference in value decreases with W. At the payment barrier, the two value functions V_0 and V_1 are close because the probability of substituting

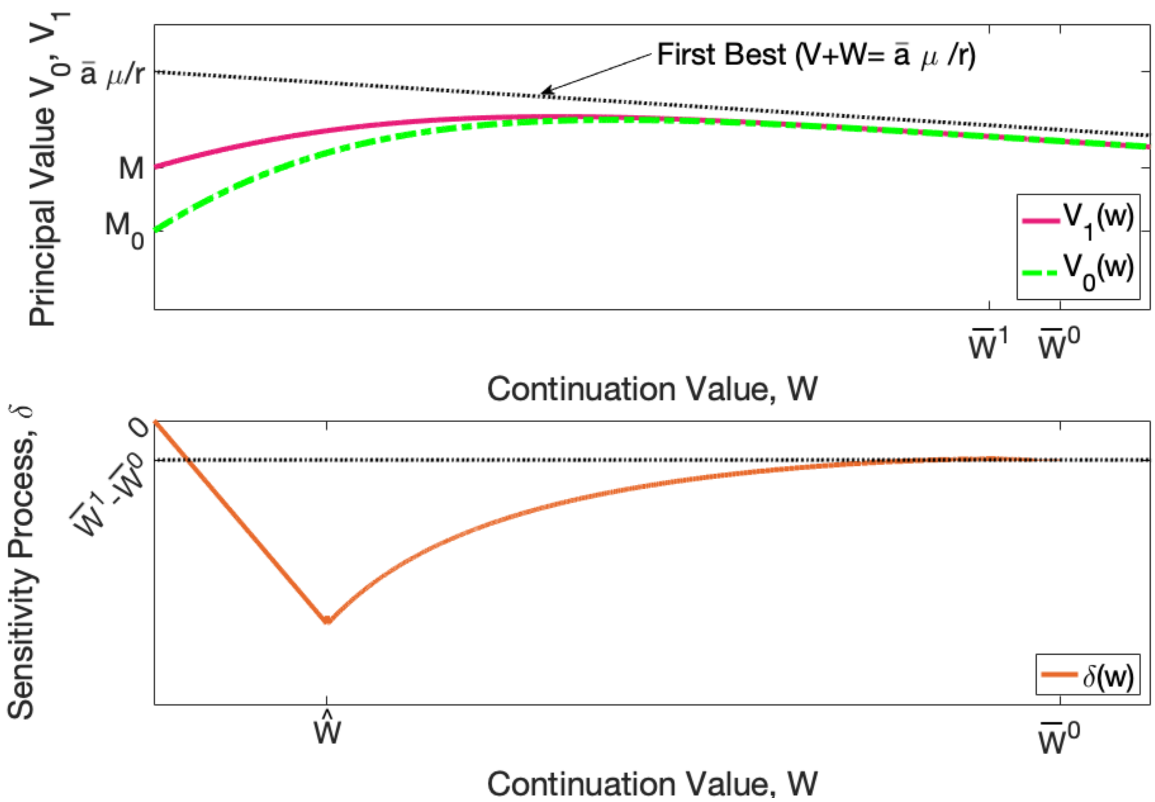

Figure 2. Principal's value functions V_0 and V_1 prior to and after the advent of the automation technology, and sensitivity process δ if $M < \frac{\bar{a}\mu}{r}$.
Parameters are $r = 10\%, \gamma = 12\%, \lambda = 12.5\%, \bar{a} = 10, \mu = 1, B = 0.8, M = 60$. $\bar{W}^0$ (resp., $\bar{W}^1$) is the payment barrier – a reflecting boundary – before (resp., after) the advent of the automation technology. The value function attaches the payment frontier with slope -1, and then extends linearly. For sufficiently large w, the limited liability constraint over $\delta(w)$ is not binding and the optimal controlled sensitivity keeps the marginal value of delegating to the manager constant before and after the advent of technology. Otherwise, such an advent triggers an instantaneous termination of the asset manager's contract and automation of asset management. We denote by $\hat{w}$ the largest value on w such that the limited liability constraint binds. We provide in Appendix Section D a description of our algorithm to simulate $(V_0, V_1, \bar{W}^0, \bar{W}^1, \delta)$.

the robot for the agent in the near future is small. Furthermore, we note that the first-best value of asset management that would be reached in the absence of agency friction is higher than V_0 and V_1, because the robot is less efficient than the agent at managing the asset.

Thus, automation would never occur in the first-best framework if $M < \frac{\bar{a}\mu}{r}$.

Next, the question raised is how the agent's continuation value reacts to the advent of the automation technology, i.e., how the jump from V_0 to V_1 occurs. Since $W_{\tilde{T}} = W_{\tilde{T}^-} + \delta(W_{\tilde{T}^-})$, it is the value of the sensitivity to the advent of the automation technology at $\tilde{T}$ that provides this information. The following proposition states that the jump is always negative according to the optimal contract, so the agent's continuation value decreases at the advent of the technology.

Proposition 4: *The optimal contract prior to the advent of the automation technology, as presented in the Proposition 3, implements a negative sensitivity to such an advent, so* $\forall t \leq \tilde{T}$, $\delta_t =\leq 0$, *where* $\delta_t = \delta(W_{t^-})$. *As the limited liability constraint imposed on the processes of sensitivity* δ *binds over the interval* $[0, \hat{W}]$, *the principal instantaneously substitutes the robot for the agent if* $W_{\tilde{T}^-} \leq \hat{W}$. *Furthermore, the payment boundary is sensitive to the advent of the automation technology and decreases from* $\bar{W}^0$ *to* $\bar{W}^1 < \bar{W}^0$ *after the advent of technology.*

Consequently, it is optimal to continuously *boost* the drift of the agent's continuation value by $-\lambda\delta(W) > 0$ prior to the advent of the technology, and then let the agent's continuation value decline instantly by $\delta(W_{\tilde{T}})$. Such a mechanism may substitute the robot for the agent instantaneously if the agent's continuation value at $\tilde{T}$ is below the threshold $\hat{W}$, defined by $V_0'(\hat{W}) = V_1'(0)$.

We note that because the advent of automation technology changes the future profitability of terminating the contract, it is optimal to make the agent sensitive to the advent of a robot. This contrasts with standard results in contracting theory, where the optimal contract does not rely on exogenous changes in the contractual environment, as in the seminal paper of Holmstrom (1979). Such a jump could also be interpreted as a *punishment for luck* to emphasize on the worsening of the agent's efficiency relative to the contract's environment. Indeed, an analogue mechanism is at play in Hoffmann and Pfeil (2010) or Demarzo et al.

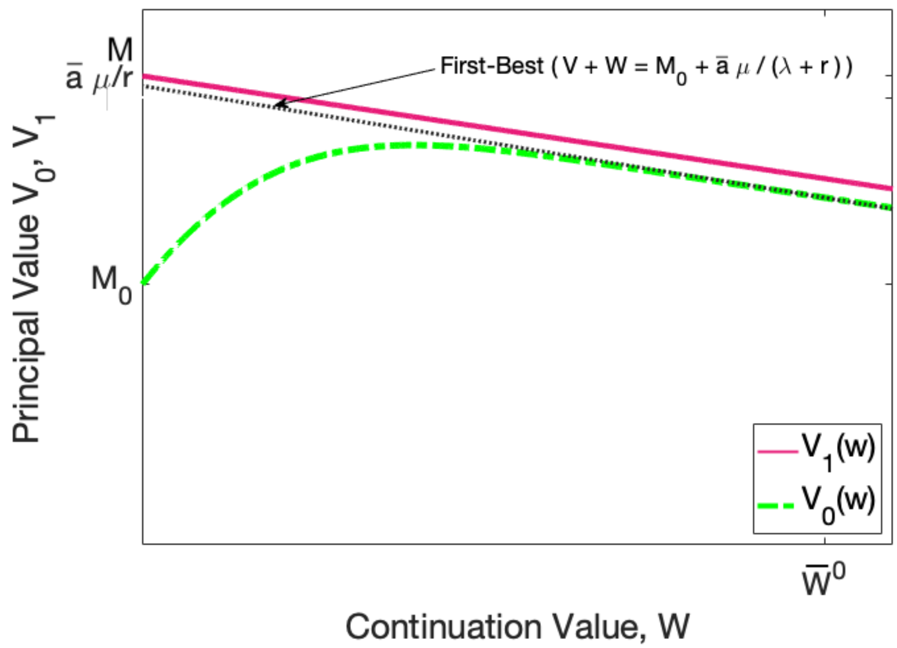

Figure 3. Principal's value functions prior to and after the advent of the automation technology if $M \geq \frac{\bar{a}\mu}{r}$.
Parameters are $r = 10\%, \gamma = 12\%, \lambda = 12.5\%, \bar{a} = 10, \mu = 1, B = 0.8, M = 105$. In this case, the limited liability constraint on the processes of sensitivity δ is always binding. Consequently, the principal instantaneously substitutes the robot for the agent at $\tilde{T}$.

(2012), where a *lucky event* changes the intrinsic profitability of the asset, captured here by parameter μ.

Next, we consider the case of $M \geq \frac{\bar{a}\mu}{r}$. Figure 3 shows the typical form of value functions in this case. The principal's value jumps instantaneously from V_0 to $V_1(0) = M$ at the advent of the technology because it is optimal to substitute the robot for the agent as soon as the robot becomes available. Indeed, $V_1'(0) = V_0'(\bar{W}^0) = -1$, so we obtain following the Proposition 4 that $\hat{W} = \bar{W}^0$. Furthermore, the first-best scenario that would consist here of delegating to an agent up to $\tilde{T}$ and switch to automation provides a lower value than V_1, where the robot is always in charge.

Contracting Upon the Forthcoming Automation Technology and Firm Value. We acknowledge that contracting upon the forthcoming automation technology may be difficult. It means that it is verifiable by a third-party, e.g., a judge. Following Frey and Osborne (2017) who rank occupations by their propensity to be automated, the availability of the robot may be assumed industry-wide, so would be λ observable. However, the value of implementing a robot-driven process is harder to assess because it is probably firm-specific, as suggested by Brynjolfsson et al. (2017). M would be then firm-specific and may even depend on the asset under management.

To examine formally the benefit of contracting upon the robot, we next compare the optimal contract derived in Propositions (2)-(3) with the suboptimal contract where the principal cannot make the agent sensitive to the advent of a robot, so we impose $\delta_t = 0$, $\forall t$. In this case, the dynamics of the continuation value if the agent who exerts full-effort satisfy

$$dW_t = \gamma W_t dt + \underline{\beta} dZ_t^{\bar{a}} - dU_t, \quad \forall \quad t \leq \tau_0. \tag{23}$$

The principal terminates the contract if and only if $\tau_0 = inf\{t \geq 0 \mid W_t = 0\}$ is reached. The principal's value V_0^s associated with the suboptimal contract satisfies :

$$\forall w \in [0; \bar{W_s^0}], \quad (\lambda + r)V_0^s(w) = \bar{a}\mu + \gamma w V_0^{s\prime}(w) + \frac{1}{2}\underline{\beta}^2 V_0^{s\prime\prime}(w) + \lambda V_1(w) \tag{24}$$

together with $V_0^s(0) = M_0$ (the value-matching condition), $V_0^{s\prime}(\bar{W}^0) = -1$ (the smooth-pasting condition), $V_0^{s\prime\prime}(\bar{W}^0) = 0$ (the super-contact condition), and where the value function V_1 is still that given in Proposition 2. Indeed, value functions are forward-looking processes, so imposing $\delta_t = 0$, $\forall t$ does not alter the value function V_1 once a robot is available.

Next, we compare $V_0(W)$ with $V_0^s(W)$. On the one hand, if we assume that the agent's reservation value is sufficiently low, then the agent is offered to start the contract with an initial continuation value $W_0^s = \arg\max_W V_0^s(W)$ in the suboptimal contract and $W_0 =$

$\arg\max_W V_0(W)$. As illustrated in Figure 4, the suboptimal contract provides a lower (respectively, larger) value to the principal (respectively, the agent). On the other hand, if the reservation value w_0 is such that $w_0 > max\left(\arg\max_W V_0^s(W), \arg\max_W V_0(W)\right)$, then the inefficiency of the suboptimal contract is lower because the difference between V_0^s and V_0 decreases with W.

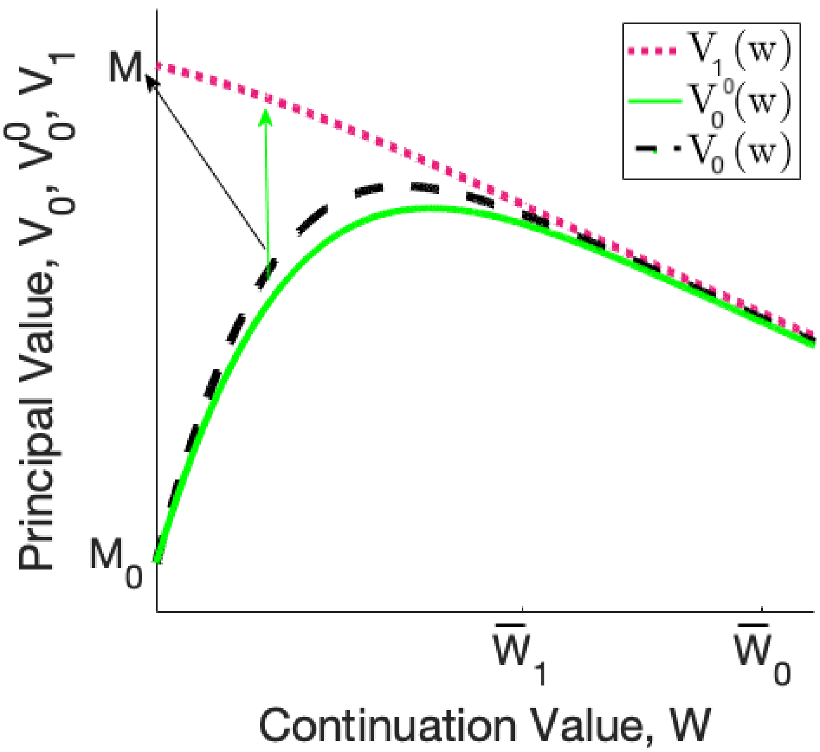

Figure 4. Value functions V_0 and V_1 derived from the optimal contract, and V_0^s derived from a suboptimal contract where we impose $\delta = 0$. Parameters are $r = 10\%, \gamma = 12\%, \lambda = 12.5\%, \bar{a} = 10, \mu = 1, B = 0.8, M = 95$.

IV. Implementation and Empirical Implications

A. Implementation

To implement the optimal contract presented in the Propositions (2)-(3), we design a *point-based incentive programme* where the number of points coincides with the asset manager's continuation value. On the one hand, whenever the point balance hits an upper payment boundary, the excess is converted into a lump sum of cash paid to the asset manager. On the other hand, the contract is terminated as soon as the agent has no more points, because the agent is then "too poor to be punished effectively" (Spear and Wang (2005)). At the advent of the automation technology, $\delta_{\tilde{T}}$ points expire and are removed from the programme. This mechanism may trigger termination due to the limited-liability condition if the number of points at $\tilde{T}$ was too low. Then and as long as the agent manages the asset, the points accumulate at a slower rate.

As in *He* (2009), the implementation of an optimal contract cannot solely use a cash balance à la DeMarzo and Sannikov (2006) that would mimic the dynamics of cash flows to trace the asset manager's continuation value. Otherwise it would be impossible to make it jump at the technological advent $\tilde{T}$.

B. Performance-Biased Automation of Asset Managers

We offer a theoretically grounded novel prediction of *performance-biased automation* of asset managers. Indeed, under an optimal contract and for a given automation technology, observing the agent's continuation value below a given threshold at the technology's advent is sufficient for deciding to instantaneously automate. This threshold is agent-specific and is determined by solving for the optimal contract (see Proposition 4). We interpret the agent's continuation value as a measure of the agent's performance because it traces the agent's history of success at managing the asset within a given contractual environment. Hence,

only an agent who has performed poorly up to the advent of the automation technology is instantaneously replaced with technology.

Furthermore, we claim that a striking illustration of performance-biased automation has been offered when BlackRock substituted algorithms for 7 of its 53 stock pickers in 2017. BlackRock's CEO Laurence D. Fink has justified such a decision by a "change [of] the ecosystem"[22], and we interpret this change as the advent of technology that can replace stock pickers. This justifies our approach of studying automation at the extensive margin, i.e., the extension of the set of automatable tasks (the task here being stock picking).

Finally, we infer that BlackRock's stock-picking algorithm is not yet as efficient as an agent would be in the absence of agency friction. Indeed, only a fraction of stock pickers have been replaced with technology. Our model predicts that otherwise every stock picker would have been replaced instantaneously, irrespective of the individual's performance history. Consequently, we predict that BlackRock has implemented a technology that is less efficient than the agent at picking stocks mainly because it is costly to provide incentives to stock pickers.

C. Other Testable Implications

Following the implementation of the optimal contract discussed in Section IV.A, the principal controls the dynamics of the points in the incentive programme – that mimics the dynamics of the agent's continuation value – to achieve optimality. While the number of points may not be observable by a third party at each instant, it then maps onto a stream of payments and a termination date that can be empirically observed. Consequently, we examine how the advent of the automation technology impacts (1) the payments made to the agent and (2) the duration of the contract. This analysis lets us make several testable predictions.

[22]https://www.nytimes.com/2017/03/28/business/dealbook/blackrock-actively-managed-funds-computer-models.html

First, we solve numerically for the optimal contract by the shooting method, and then perform Monte Carlo simulations[23]. Following the baseline model of DeMarzo and Sannikov (2006), we set the principal's discount rate to $r = 10\%$, and the instantaneous performance of the asset under full effort to $\bar{a}\mu = 10$. We fix the agent's discount rate to $\gamma = 12\%$, and assume that $B = 8$, so $\frac{B}{\bar{a}} = 0.8$, where $\frac{B}{\bar{a}} \in [0, 1]$ to simulate the presence of a severe agency friction observed in the finance industry. As to the automation technology, there is no consensus on what figure to use. According to Frey and Osborne (2017) who categorize jobs according to the probability of automation, "securities, commodities, and financial services sale agents" (SOC code 41-3031) such as asset managers are at slight risk of automation (probability of 1.6%, ranked the 74^{th} less likely occupation to be automated over 702 occupations). Nevertheless, BlackRock has already in 2017 substitutes stock pickers for algorithms. Thus, we fix the intensity of its advent to $\lambda = 0.125$, so the technology should become available on average in $\frac{1}{0.125} = 8$ years, and the probability of the technology arising in less than a year is $1 - e^{-0.125} = 11.75\%$. We consider the value of automating $M \in [60; 99]$, so the robot-managed asset reaches between 60% and 99% of the performance of agent-managed asset in the absence of agency friction, and the latter is given by $\frac{\bar{a}\mu}{r} = 100$ (see Proposition 1). Table I presents the parameter values used in the numerical analysis.

Expected Bonus Following the Advent of a Robot. Bonuses are easy to observe empirically. Thus, we consider here the sum of payments that have been received by the agent over a year, as it can be regarded as a proxy for an annual bonus. To examine whether the expiration of points at the advent of a robot leads to a decline in bonus, we then compare this to the counterfactual scenario of this event not occurring.

We fix the starting number of points in the incentive programme to $W_0 = w$, and consider

[23]For the initial continuation value, we take $W_0 \in [\hat{W}, \bar{W}^0]$, where $\hat{W}$ is the smallest value such that $V_1'(w + \delta(w)) = V_0'(w)$. For $W_0 < \hat{W}$, the contract would be instantaneously terminated if $N_0 = 1$, so the analysis of the bonus and the contract duration is irrelevant below that threshold.

Variable	Symbol	Parameter	Symbol	Value
Value of the robot	M	Principal's discount rate	r	0.10
Cumulative cash flows	X	Agent's discount rate	γ	0.12
Contract termination date	τ	Full effort	$\bar{a}$	10
Date of advent of technology	$\tilde{T}$	Asset's intrinsic performance	μ	1
Firm value prior to the advent of the automation technology	V_0	Private benefit	B	8
Firm value after the advent of the automation technology	V_1	Intensity of the advent of the automation technology	λ	0.125
Payment boundary prior to the advent of the automation technology	$\bar{W}^0$	Agent's reservation value	w_0	0
Payment boundary after the advent of the automation technology	$\bar{W}^1$			
Sensitivity to the asset performance	β			
Sensitivity to the advent of the automation technology	δ			

Table I
Parameter Values and Variables

the discounted sum of payments received over $[0, t \wedge \tau]$ just after the advent of the automation technology ($N_0 = 1$), and that we denote by ϕ_1^t. We have

$$\phi_1^t(w) = \mathbb{E}_w \left[\int_0^{t \wedge \tau} e^{-\gamma s} dU_s \mid N_0 = 1 \right], \tag{25}$$

where, according to Proposition (3), the agent's number of points jumps instantly from $W_{0^-} = w$ to $W_0 = w + \delta(w)$. In the counterfactual scenario, the discounted sum of payments

received over $[0, t \wedge \tau]$ if the automation technology does not emerge ($N_0 = 0$) is given by

$$\phi_0^t(w) = \mathbb{E}_w \left[\int_0^{t \wedge \tau} e^{-\gamma s} dU_s \mid N_0 = 0 \right]. \tag{26}$$

In this case, the automation technology can still emerge during the interval $]0; t \wedge \tau]$. We remind the reader that the contract characteristics $(U_t)_t$ and τ are adapted to the entirety of information available to the principal, so they differ in equations (25) and (26). We compare the cross-sectional estimates $\phi_1^t(w)$ and $\phi_0^t(w)$, and consider $t = 1$; our results are provided in Table II for $M = \{60, 75, 95, 99\}$.

We estimate that ϕ_0^t dominates ϕ_1^t. This is illustrated in Figure 5 for $M = 90$, and is consistent with the values provided in Table II for other values of M. Thus, we conclude that the annual bonus declines after the advent of the automation technology. This is consistent with the empirical observation made by the 2018 *Asset Management Compensation Study* by Greenwich Associates. Nevertheless, The authors claim that the large cost of investing in automation technologies reduces the incentive compensation pool of asset managers, while our optimal contracting approach offers a potential alternative channel based on the decline in the agency rent.

Furthermore, we also note that the annual bonus increases with the number of points in the incentive programme. This is intuitive because the latter can be interpreted as a measure of the agent's performance history. Figure 6 shows the severity of the decline in the annual bonus at the advent of a robot, which is given by $\frac{\phi_1^t - \phi_0^t}{\phi_0^t}$. As expected, the decline in the bonus is larger if M is higher.

The value of M may be firm-specific if it is interpreted as the efficiency of the firm in adopting the forthcoming robot. While it is in practice difficult to observe directly, Brynjolfsson et al. (2017) assert that firms adapt to technology, and that complementary investments and structural changes are important factors for efficient usage of innovations.

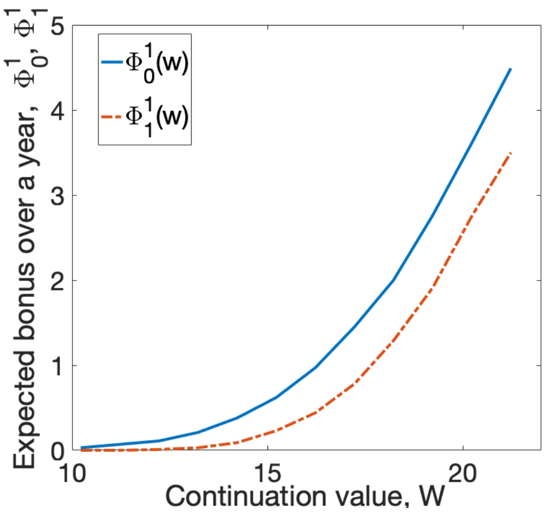

Figure 5. Expected Annual Bonus following the advent of technology Φ_1^1 and in the counterfactual scenario Φ_0^1.

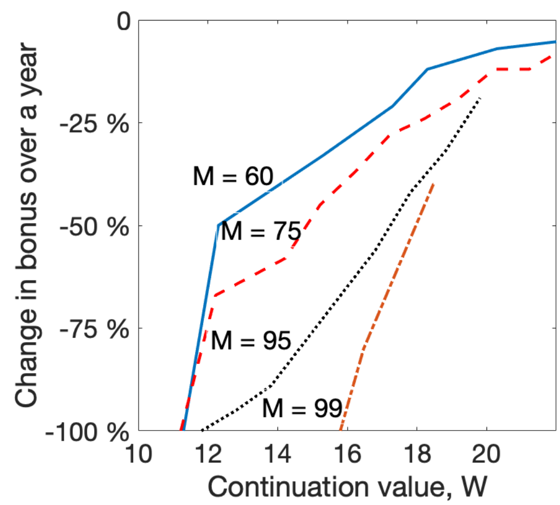

Figure 6. Change in Expected Annual Bonus following the advent of technology.

Hence, we could for instance infer the value of M from the size of the firm's I.T. department or the firm's investment in fintechs. This leads to the following cross-sectional implication :

Implication 1: *The average decline in the annual bonus following the advent of technology is higher in firms that are better adapted to the usage of automation technology.*

Threat of Termination, Implementation Lag, and the Advent of a Robot. Does the advent of a robot always put an agent at a greater risk of termination ? Are robots adopted faster in firms better adapted to their usage ?

To answer these questions, let us first investigate the expected contract duration denoted by τ_1^t. It satisfies

$$\tau^1(w) = \mathbb{E}_w \left[\inf\{t > 0 \mid W_t = 0 \text{ such that } \quad W_0 = w \quad \& \quad N_0 = 1\} \right], \tag{27}$$

where, according to Proposition (3), $W_{0^-} = w$ jumps instantaneously to $W_0 = w + \delta(w)$ because a robot has appeared ($N_0 = 1$). In the counterfactual scenario, the expected contract

M = 99

w	15.80 ($\hat{W}$)	17.47	18.47 ($\bar{W^0}$)
Φ_0^1(w)	3.19	4.17	5.11
Φ_1^1(w)	0	1.73	3.12

M = 95

w	11.44 ($\hat{W}$)	12.81	13.81	14.81	15.81	16.81	17.81	18.81	19.81 ($\bar{W^0}$)
Φ_0^1(w)	0.24	0.38	0.65	1.00	1.44	2.09	2.83	3.66	4.57
Φ_1^1(w)	0	0.02	0.07	0.22	0.47	0.92	1.63	2.5	3.7

M = 75

w	6.41 ($\hat{W}$)	8.22	9.22	10.22	11.22	12.22	13.22	14.22	15.22	16.22
Φ_0^1(w)	0	0	0	0,01	0.01	0.03	0.06	0.12	0.22	0.38
Φ_1^1(w)	0	0	0	0	0	0.01	0.02	0.05	0.12	0.24
w	17.22	18.22	19.22	20.22	21.22	22.22	23.22 ($\bar{W^0}$)			
Φ_0^1(w)	0.61	0.97	1.44	1.98	2.76	3.58	4.5			
Φ_1^1(w)	0.44	0.74	1.17	1.74	2.43	3.32	4.12			

M = 60

w	4.64 ($\hat{W}$)	6.3	7.3	8.3	9.3	10.3	11.3	12.3	13.3	14.3
Φ_0^1(w)	0	0	0	0	0	0	0.01	0.02	0.03	0.06
Φ_1^1(w)	0	0	0	0	0	0	0	0.01	0.01	0.03
w	15.3	16.3	17.3	18.3	19.3	20.3	21.3	22.3	23.3	24.3 ($\bar{W^0}$)
Φ_0^1(w)	0.12	0.24	0.39	0.59	0.98	1.46	2.01	2.8	3.64	4.59
Φ_1^1(w)	0.08	0.16	0.31	0.52	0.82	1.36	1.85	2.65	3.48	4.26

Table II
Estimates of the average bonuses paid over the year following the advent of the automation technology (Φ_1^1) and in the counterfactual scenario of the automation technology not emerging (Φ_0^1). Estimates are provided for the set $M = (60; 75; 95; 99)$ of value of the automation technology.

duration is denoted by $\tau^0(w)$ and satisfies

$$\tau^0(w) = \mathbb{E}_w \left[\inf\{t > 0 \text{ such that } W_t = 0, \text{ given } \quad W_0 = w \quad \& \quad N_0 = 0\}\right]. \tag{28}$$

Estimating the change in contract duration due to the advent of the automation technology means comparing the cross-sectional estimates $\tau^1(w)$ and $\tau^0(w)$. Our estimates are provided in Table III for $M = \{60, 75, 95, 99\}$.

As illustrated in Figure 7 for the case of $M = 95$, τ^0 dominates τ^1 so the advent of a robot makes contract termination occurs sooner. This is particularly true if the number of points in the incentive programme is small. The decline in the expected annual bonus following the advent of a robot and the greater risk of contract termination once the robot has becomes available, considered together, mean that the optimal contract effectively adjusts the provision of incentives to the availability of the robot. Thus, we posit that the early-stage period – the period prior to the advent of the automation technology – is a *golden age* of asset managers. While many authors argue that the costs of asset management, and in particular the delegation of active investment are too high (see French (2008), etc.), our paper shows that a golden-age at an early stage may not be abnormal and rather serves the purpose of increasing the efficiency of the contractual relationship in the long term when an automation technology capable of replacing the agent is foreseen to emerge.

Finally, we offer a prediction for the expected time necessary to adopt the robot after its advent. Brynjolfsson et al. (2017) refer to this as the *implementation lag* and argue that it is the main reason we have not yet observed a more significant contribution of A.I. in the economy. This lag is defined in our setting by $(\tau(w) \vee \tilde{T}) - \tilde{T}$, where $\tau(w)$ is the contract termination time, and $\tilde{T}$ represents the advent of a robot. As illustrated in 8, the implementation lag decreases with M. We note that it is not significantly sensitive to a change in the initial continuation value W, because W does not measure the future performance of the agent involved in the criteria for automating. Again, M can be interpreted

M = 99

w	15.80 ($\hat{W}$)	17.47	18.47 ($\bar{W}^0$)
τ^0(w)	3173	3175	3131
τ_1(w)	282	504	584

M = 95

w	11.44 ($\hat{W}$)	12.81	13.81	14.81	15.81	16.81	17.81	18.81	19.81 ($\bar{W}^0$)
τ^0(w)	6898	6808	6961	7238	7465	7646	7721	7759	7768
τ_1(w)	977	3233	4294	5162	5783	5918	6225	6341	6399

M = 75

w	6.41 ($\hat{W}$)	8.22	9.22	10.22	11.22	12.22	13.22	14.22	15.22	16.22
τ^0(w)	6725	7500	7966	8323	8855	9307	9625	9765	9830	10153
τ_1(w)	2074	4082	5728	6849	7615	8439	8832	9299	9532	9748
w	17.22	18.22	19.22	20.22	21.22	22.22	23.22 ($\bar{W}^0$)			
τ^0(w)	10095	10248	10435	10321	10371	10430	10333			
τ_1(w)	10020	10010	10133	10212	10178	10268	10237			

M = 60

w	4.64 ($\hat{W}$)	6.3	7.3	8.3	9.3	10.3	11.3	12.3	13.3	14.3
τ^0(w)	5506	6495	7129	7788	8275	8893	9229	9454	9713	10165
τ_1(w)	1744	3542	4843	5899	6827	7829	8634	9123	9455	9755
w	15.3	16.3	17.3	18.3	19.3	20.3	21.3	22.3	23.3	24.3 ($\bar{W}^0$)
τ^0(w)	10259	10317	10337	10433	10563	10611	10577	10594	10602	10628
τ_1(w)	9999	10141	10393	10407	10525	10598	10622	10650	10614	10651

Table III
Estimates of the average contract duration following the advent of the automation technology (τ_1^1) and in the counterfactual scenario of the automation technology not emerging (τ_0^1). Estimates are provided for the set $M = (60; 75; 95; 99)$ of value of the automation technology.

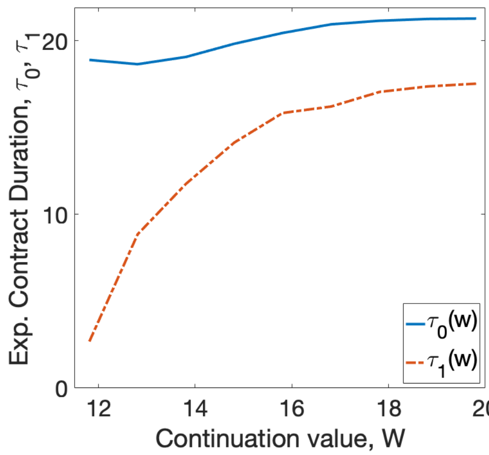

Figure 7. Expected contract duration following the advent of technology τ_1 and in the counterfactual scenario τ_0.

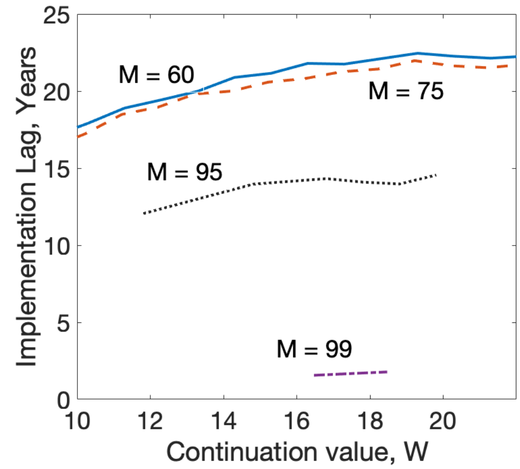

Figure 8. Implementation Lag for different values of M.

as the firm-specific efficiency of using the robot, and we have the following implication:

Implication 2: *Firms that are better adapted to the usage of a forthcoming technology will implement it faster.*

V. Extensions

A. Optimal investment in automation technology

Thus far, both parties have been presumed to take as given the characteristics and the arrival time of the robot. However, varied empirical evidence suggests that the principal may actually take action to improve the efficiency of the technology. First, Brynjolfsson et al. (2017) show that a complementary investment and structural changes are a prerequisite to an efficient usage of artificial intelligence. Second, a study by Boston Consulting Group[24] (2016) suggests that firms in the capital markets and asset management industry may also

[24] *Fintech in Capital Markets: A Land of Opportunity*, 2016.

participate directly in the innovation process through investments in fintechs. It shows that external innovation tends to be preferred over internal R&D, and that fintechs are mainly engaged through M&A and VC funding by such firms. In both of these examples, large investments are made prior to the advent of the automation technology to increase its profitability. However, the delegation to an asset manager that is subject to a financial friction before the adoption of the automation technology may impact this investment decision.

In this section, we extend the model to allow for the case of the principal making an investment at $t = 0^-$ – prior to the contracting date – to increase the profitability of the forthcoming automation technology from M to $(1+\kappa)M$ at a sunk cost $c(\kappa)M$, with $\kappa > 0$. We assume a quadratic investment cost, and $c(\kappa) = \frac{\kappa^2}{2}$, and the firm has to bear the cost of investment at $t = 0^-$ and wait up until $t = \tilde{T}$ to benefit from the implementation of the automation technology. As in the main model, the automation technology emerges following a single jump process of intensity λ and the principal offers a contract to delegate the management of the asset to an agent at $t = 0$. Our goal is to explore how the presence of the agency conflict impacts the optimal investment compared to the first-best benchmark. The necessary condition for such an investment not to be worthless is that

$$\frac{\lambda}{\lambda + r}\kappa > c(\kappa), \tag{29}$$

so it follows that the principal investment is within $[0; 2\frac{\lambda}{\lambda+r}]$.

Next, let us consider the first-best benchmark. If the value of automation after the investment – net of the cost of investment at $t = 0^-$ – remains lower than the value of continuing to delegate to the asset manager, then the principal never invests. The investment is made in the first-best benchmark if and only if[25] :

$$\Pi(\kappa)M > \frac{\lambda}{\lambda + r}\frac{\bar{a}\mu}{r}, \tag{30}$$

[25] We note that if inequality (31) holds, then the inequality (30) holds as well.

where $\Pi(\kappa) = \frac{\lambda}{\lambda+r}(1+\kappa) - c(\kappa)$ is the scaled (by M) profit of investing at level κ. Thus, the first-order condition with respect to κ shows us that the optimal level of investment in the first-best framework satisfies $c'(\kappa^{FB}) = \frac{\lambda}{\lambda+r}$. The first-best level of investment is

$$\kappa^{FB*} = \begin{cases} \frac{\lambda}{\lambda+r} \quad \text{if} \quad (1 + \frac{1}{2}\frac{\lambda}{\lambda+r})M > \frac{\bar{a}\mu}{r}; \\ 0 \quad \text{otherwise.} \end{cases} \tag{31}$$

Next, let us examine the optimal level of investment in the presence of agency friction. The principal's problem is to determine

$$\sup_{\kappa} V_0^{\kappa}(w) - c(\kappa)M \tag{32}$$

where $V_0^{\kappa}(w)$ denotes the principal's value function when investment κ is made and when the agent manages the asset under an optimal contract. It satisfies the following second-order differential equation

$$\begin{aligned} \forall w \in [0; \overline{W}_0^{\kappa}], \quad & (\lambda + r)V_0^{\kappa}(w) = \\ & \bar{a}\mu + (\gamma w - \lambda\delta)V_0^{\kappa\prime}(w) + \frac{1}{2}\underline{\beta}^2 V_0^{\kappa\prime\prime}(w) + \lambda V_1^{\kappa}(w+\delta) \end{aligned} \tag{33}$$

together with $V_0^{\kappa}(0) = \frac{\lambda}{\lambda+r}(1+\kappa)M$, $V_0^{\kappa\prime}(\bar{W}_0^{\kappa}) = -1$ and $V_0^{\kappa\prime\prime}(\bar{W}_0^{\kappa}) = 0$ for some $\bar{W}_0^{\kappa} \geq 0$ that is the payment boundary prior to the availability of the technology. It depends on $V_1^{\kappa}(w)$, the principal's value function after the advent of the automation technology that is the solution of equation (19) together with the boundary conditions $V_1^{\kappa}(0) = (1+\kappa)M$, $V_1^{\kappa\prime}(\bar{W}_1^{\kappa}) = -1$ and $V_1^{\kappa\prime\prime}(\bar{W}_1^{\kappa}) = 0$ for some $\bar{W}_1^{\kappa} \geq 0$ that plays the role of the payment boundary once the technology is available. Applying Lemma 2, we obtain that the sensitivity to the asset realization remains at $\beta = \underline{\beta}$, and the sensitivity δ to the advent of automation technology is the solution of $V_1'(W_{t^-} + \delta(W_{t^-})) = V_0'(W_{t^-})$ for all W_{t^-} such that $V_1'(W_{t^-}) \geq V_1'(0)$ and $\delta(W_{t^-}) = -(W_{t^-})$ otherwise (the limited liability condition).

For now, let us focus on the case of the automation technology being more efficient than the agent, so inequality (30) holds. We show that in this case, the contracting problem and the investment problem can be separated. The following Proposition summarizes this result:

Proposition 5: *If condition (30) is satisfied, the principal's problem becomes*

$$V_0(w) + \sup_{\kappa} \Pi(\kappa) M, \tag{34}$$

where V_0 is the value function prior to the advent of technology; this function was characterized in Proposition 3. Thus, the contracting problem and the investment problem can be separated, and the principal always optimally invests as in the first-best benchmark.

In turns out that the first-best level of investment in the automation technology is reached if inequality (30) holds. Consequently, investing at a level κ shifts the principal's value function upward by $\Pi(\kappa)M$, as illustrated in Figure 9. In particular, it means that the investment problem does not impact the contractual characteristics.

If the automation technology is not as efficient as the agent, then the principal does not necessarily substitute the robot for the agent at the technology's advent $\tilde{T}$, but does so only at the contract termination τ[26]. Consequently, while at first-best the principal never invests in the automation technology if the inequality (30) does not hold, this is not the case in the presence of agency friction. We draw in Figure 10 a typical form of the optimal investment strategy. We can interpret this result as long-termism, as the principal over-invests in the quality of the *future* automation technology (with respect to the first-best case) to mitigate the *current* agency concern.

[26] $\tau = \tilde{T}$ if the agent has not performed sufficiently well at managing the asset.

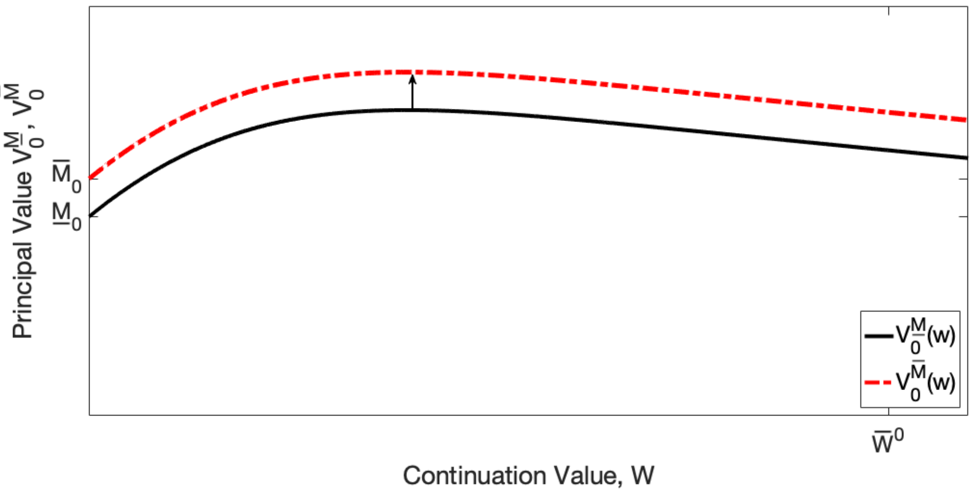

Figure 9. Change in the value function associated with an increase in $\underline{M}$ to $\bar{M}$, where $\frac{\bar{a}\mu}{r} < \underline{M} < \bar{M}$.
When the technology is more efficient than the agent, the investment problem and the contracting problem can be separated. Consequently, investing moves upward the value function, and does not impact the payment barrier $\bar{W}^0$.
Parameters are $r = 10\%, \gamma = 12\%, \lambda = 12.5\%, \bar{a} = 10, \mu = 1, B = 0.8, \underline{M} = 105,\ \bar{M} = 125$.

B. *Technology Enhancing or Replacing the Agent*

So far, our model has solely considered substituability between robots and high-skill jobs, as advocated by Acemoglu and Restrepo (2018b). Next, we depart from such a setting and assume that the robot can also work in synergy with the agent, so contract termination is no longer a prerequisite for the implementation of the robot. At the advent of the automation technology that follows a single jump process with intensity λ, suppose that the principal faces two mutually exclusive opportunities: either (1) terminate the agent's contract and replace the agent with the robot (this is the case that has been covered thus far in this paper) or (2) continue the contractual relationship and instead of replaced the agent, enhance the agent with the same technology at a sunk cost $I > 0$. Enhancing the agent permanently

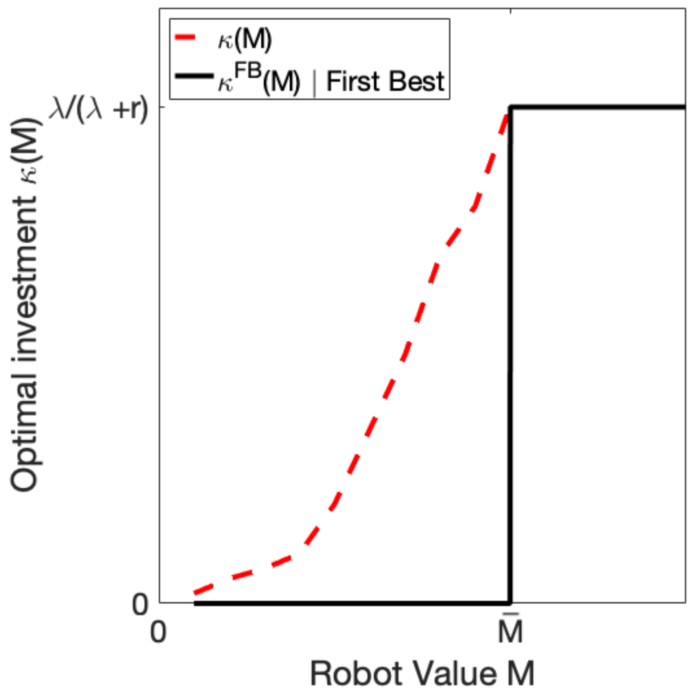

Figure 10. Optimal investment in the automation technology at first best and in presence of agency friction.
Parameters are $r = 10\%, \gamma = 12\%, \lambda = 12.5\%, \bar{a} = 10, \mu = 1, B = 0.8, M = 90$. $\bar{M}$ is such that $\Pi(\kappa)\bar{M} = \frac{\lambda}{\lambda+r}\frac{\bar{a}\mu}{r}$. For values of M over $\bar{M}$, the optimal investment is not distorted by the presence of agency friction. Under $\bar{M}$, the presence of the agency friction makes the principal over-invest in the automation technology.

augments the agent's productivity,[27] so it increases from $a\mu$ to $(1+\theta)a\mu$, where $\theta > 0$ is the parameter of *synergy* between the agent and the robot. As the changes in production are at the firm level and not agent-specific, we assume that it does not impact the agent's outside option that is normalized to 0. Then, it leads us to the following questions : What are the optimal incentives to provide to an agent when both parties foresee the advent of the automation technology that can either replace or enhance the agent? When is it optimal to enhance the agent or to replace an agent with a robot?

[27]Factor-augmenting technological change is one of the approaches to the study of automation in labour economics (See Bessen (2017), etc.).

To answer these questions, we derive heuristically an optimal contract under the assumption that the full-effort strategy is optimal. Using backward induction, we first characterize the value function after the advent of the automation technology, which is associated with the principal's decision on the optimal time to enhance or replace the agent. Following the growth option model of Décamps and Villeneuve (2007), such a value function is the solution of the stopping problem

$$V_1(W_t) = \sup_{\tau \geq t} \mathbb{E}\left[e^{-r(\{\tau\wedge\tau_0\}-t)} \left(V_s(W_{\tau\wedge\tau_0}), V_e(W_{\tau\wedge\tau_0})\right)\right], \tag{35}$$

where $\tau_0 = \inf\{s \geq 0 \mid W_s = 0\}$ as defined in Proposition (2), and V_s (respectively V_e) is the value function associated with the opportunity to replace the agent with a robot (respectively, to enhance the agent). In the rest of the section, V_1 is assumed to be concave. While V_s has been characterized already and satisfies equation (19) in Proposition 2, we still have to characterize V_e. It is the largest solution of

$$rV_e(w) = \sup_{\beta\geq\underline{\beta}} \left((1+\theta)\bar{a}\mu + \gamma w V_e'(w) + \frac{1}{2}\beta^2 V_e''(w) \right) \tag{36}$$

with boundary conditions $V_e(0) = 0$, $V_e'(\bar{W}_e) = -1$, and $V_e''(\bar{W}_e) = 0$ for some $\bar{W}_e \geq 0$ that plays the role of the payment boundary once the human-enhancing technology has been implemented. As it is optimal to lay off the agent when the agent's continuation value drops to 0, we set $V_e(0) = 0$ because the implementation of an agent-enhancing robot does not produce any benefit if no agent is managing the asset. Then, V_1 is the largest solution to

$$rV_1(w) = \sup_{\beta\geq\underline{\beta}} \left(\bar{a}\mu + \gamma w V_1'(w) + \frac{1}{2}\beta^2 V_1''(w) \right) \tag{37}$$

with boundary conditions $V_1(0) = V_s(0) = M$ $V_1(b) = V_e(b) - I$ (the value matching condition) and $V_1'(b) = V_e'(b)$ (the smooth pasting condition) for some threshold $b \geq 0$ where it

is optimal to implement the agent-enhancing robot, and where we assume $M < \frac{\bar{a}\mu}{r}$. Figure 11 shows a typical construction of the value function V_1, where the sensitivity to the asset realization is set to $\beta_t = \underline{\beta}, \quad \forall t$ as stated in Lemma 2. As before, the principal waits for $\tau_0 = \inf\{t \geq 0 \mid W_t = 0\}$ to replace the agent with the robot, and thus the equality between V_1 and V_s holds for $W = 0$. Moreover, the agent-enhancing robot is not implemented when the agent is at too great a risk of termination, and the principal defers its implementation up to a positive threshold far enough to the termination boundary. Then, the principal waits for the agent's continuation value to hit $\bar{W^1} = \bar{W}_e$ to compensate the agent. We note that the presence of a robot able to either replace or enhance the agent leads to complex non-monetary incentives.

Next, let us consider the value function V_0 prior to the advent of the automation technology. Under an optimal contract that implements the full-effort strategy, the dynamics of the agent's continuation value satisfy equation (23) in Proposition 3, and V_0 is the largest solution of

$$\begin{aligned} &rV_0(w) = \\ &\sup_{\beta \geq \underline{\beta}; \delta \geq -w} \left(\bar{a}\mu + (\gamma w - \lambda\delta)\, V_0'(w) + \frac{1}{2}\beta^2 V_0''(w) + \lambda(V_1\,(w+\delta) - V_0\,(w)) \right) \end{aligned} \tag{38}$$

with boundary conditions $V_0(0) = \frac{\lambda}{\lambda+r}M$ $V_0'(\bar{W^0}) = -1$ and $V_0''(\bar{W^0}) = 0$. V_1 is the value function after the advent of the automation technology and is characterized by the equation (37). Then, the sensitivity to the asset realization remains as before to $\beta = \underline{\beta}$, and the sensitivity δ to the advent of the automation technology is the solution of $V_1'(W_{t^-} + \delta(W_{t^-})) = V_0'(W_{t^-})$ for all W_{t^-} such that $V_1'(W_{t^-}) \geq V_1'(0)$ and $\delta(W_{t^-}) = -(W_{t^-})$ otherwise (the limited liability condition). By construction, $V_0'(\bar{W^0}) = V_1'(\bar{W^1})$, and thus $\delta(\bar{W^0}) = \bar{W^1} - \bar{W^0}$. The next proposition states that here $\bar{W^1} > \bar{W^0}$, and thus $\delta(w)$ is always positive on the upper part of the employment interval $[0, \bar{W^0}]$.

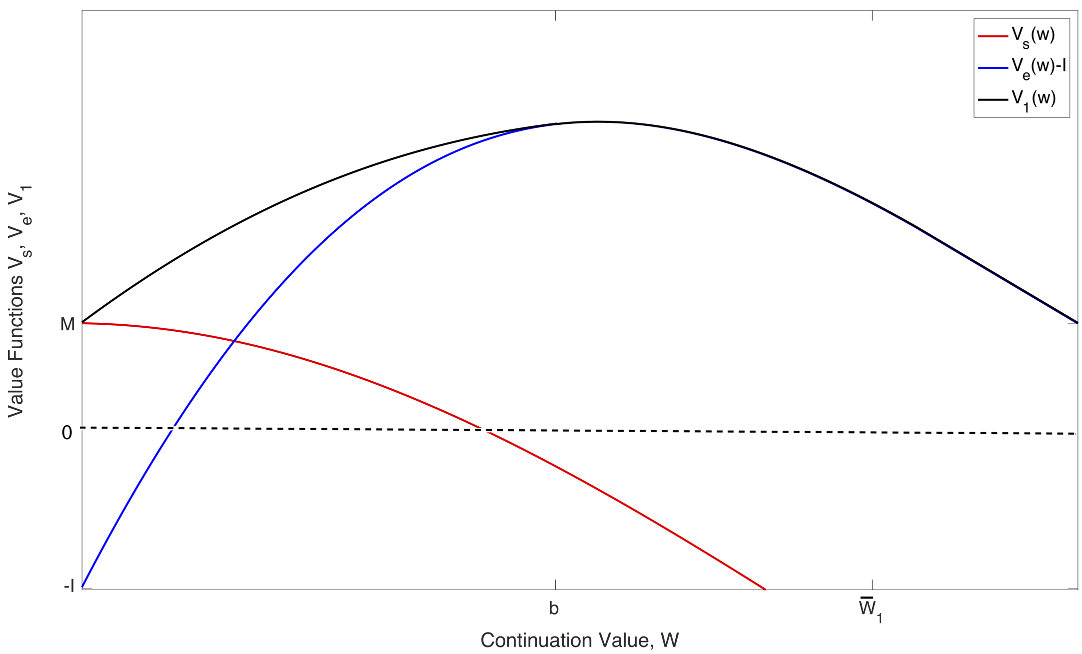

Figure 11. Principal value after the advent of the automation technology when robot can either replace or enhance the agent.
Parameters are $r = 10\%, \gamma = 12\%, \lambda = 12.5\%, \bar{a} = 10, \mu = 1, B = 0.8, M = 90, I = 10$. The principal substitutes the robot for the agent when his continuation value hits 0 for the first time, and the robot enhances the agent when his continuation value is above the threshold b for the first time. b is by $V_1(b) = V_e(b) - I$. Then, payments are provided at the threshold $\bar{W}^1$ to an agent who has been enhanced.

Proposition 6: *When the robot enhances the agent, the payment boundary moves upwards at the advent of the automation technology, i.e.* $\bar{W}^1 > \bar{W}^0$. *Furthermore,* $V_1'(0)$ *increases with* θ.

Consequently, two cases that depend on the value of the synergy parameter θ may arise. If θ is sufficiently low, the agent is pushed towards the tails of the employment interval $[0; \bar{W}^1]$, and the sensitivity δ exhibits a "hollowing-out" pattern for a middle-performing agent, as shown in Figure12.a. Indeed, δ is negative (positive) for small (respectively, large) values of w because $V_0'(0) > V_1'(0)$, so the principal prefers to put at risk of automation an

agent that has not performed sufficiently well (and, respectively, to push an agent that has performed well towards the region where the agent-enhancing robot is implemented). As in the main model of this paper, the contract instantaneously automates a bad performer. If θ is sufficiently large, the Figure12.b indicates that the enhancement effect may also dominate the "displacement" effect. Indeed, if θ is large enough then $V_0'(w) \leq V_1'(w) \quad \forall \quad w$, so $\delta \geq 0$ for all w. Consequently, the agent's continuation value is always *boosted* by the advent of technology in this case. We note that due to the full-commitment assumption, the threshold b still has to be reached for implementation of the agent-enhancing robot and the agent is still replaced at $\tau_0 = \inf\{t \geq 0 \mid W_t = 0\}$, so the contract can still be terminated after the advent of technology and before enhancing the agent. To conclude, this extension suggests that the propensity of the advent of the automation technology to polarize asset managers and more generally high-skill jobs may have a contractual foundation and be due to the ability of robots to either enhance or replace agents.

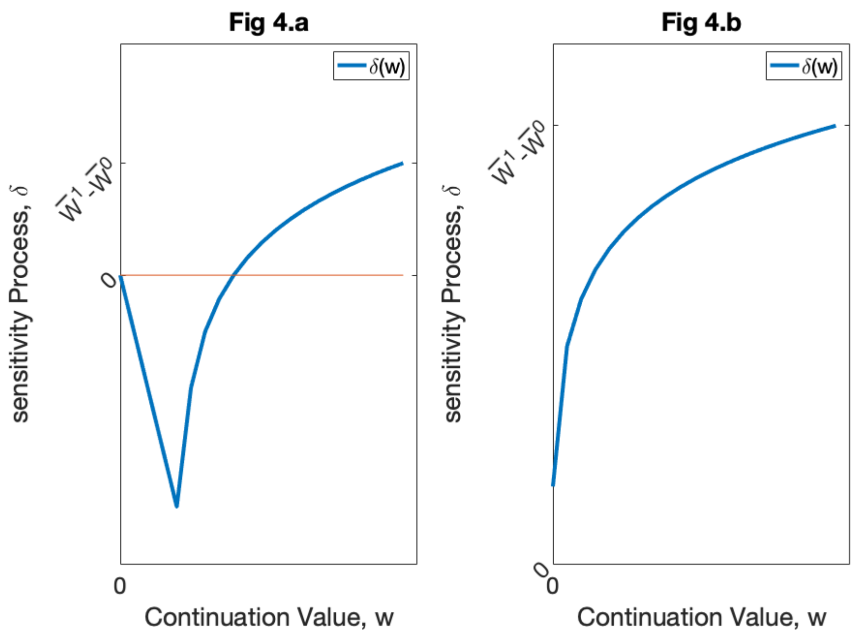

Figure 12. Process δ of sensitivity to the advent of the automation technology when the technology can either replace or enhance the agent. $\theta = \theta_{low}$ on the left panel, and $\theta = \theta_{high}$ on the right panel, with $\theta_{low} < \theta_{high}$.
Parameters are $r = 10\%, \gamma = 12\%, \lambda = 12.5\%, \bar{a} = 10, \mu = 1, B = 0.8, M = 90, \theta_{low} = 20\%, \theta_{high} = 150\%$. The payment boundary jumps from $\bar{W}^0$ before the advent of technology to $\bar{W}^1 > \bar{W}^0$. For low values of synergy which is the case on the left panel, the advent of the automation technology has a "hollowing-out" pattern for middle-performers : continuation value of poor-performers decreases and continuation value of good-performers increases. If the synergy parameter is sufficiently large as in the right panel, the enhancing effect dominates the displacement effect, and the agent's continuation value always increases instantly at the advent of the automation technology.

VI. Concluding Remarks

In this paper, we study a continuous-time principal-agent model à la DeMarzo and Sannikov (2006) with effort. We embed the advent of a robot that can compete with the agent at managing the asset. It is regarded as an irreversible substitution device that arises stochastically.

We derive an optimal long-term contract that adjusts the provision of incentives to the

availability of the robot. Indeed, the principal foresees from the contracting date that a valuable alternative to the agent will become available in the future, and thus designs an incentive-compatible contract to take advantage of this forthcoming and valuable alternative.

We show that it is optimal to automate at the advent of technology if the agent has performed poorly at managing the asset, or if the robot is more efficient than the agent. Hence, we predict a *performance-biased automation* of asset managers, as the impact of the advent of technology on the contract depends on the agent's history of performance of managing the asset.

Finally, we acknowledge that many important factors that impact the decision to automate in the presence of moral hazard are left unaccounted for. Further studies may intend to, e.g., embed ambiguity on the inherent black-box nature of the automation technology, internalize the technology's development within the firm, and include the presence of a sunk cost to automate when the principal has cash constraints.

REFERENCES

Acemoglu, Daron, and David Autor, 2011, Skills,Tasks and Technologies : Implications for Employment and Earnings, *Handbook of Labor Economics* 4, 154.

Acemoglu, Daron, and Pascual Restrepo, 2018b, Modeling Automation, *AEA Papers and Proceedings* 108, 48–53.

Acemoglu, Daron, and Pascual Restrepo, 2020, Robots and jobs: evidence from U.S. labor markets, *Journal of Political Economy* 128(6), 2188–2244.

Arntz, Melanie, Terry Gregory, and Ulrich Zierahn, 2016, The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis, *OECD Social, Employment and Migration Working Papers* 2, 47–54.

Autor, David, and David Dorn, 2013, The growth of low-skill service jobs and the polarization of the us labor market, *American Economic Review* 103, 1553–1597.

Autor, David H., David Dorn, and Gordon H. Hanson, 2013, The China Syndrome: Local labor market impacts of import Competition in the United States, *American Economic Review* 103, 2121–2168.

Autor, David H., Frank Levy, and Richard J. Murnane, 2003, The skill content of recent technological change: An empirical exploration, *The Quarterly journal of economics* 1279–1333.

Ben-David, Itzhak, Justin Birru, and Andrea. Rossi, 2020, The Performance of Hedge Fund Performance Fees, *NBER WP* .

Bertrand, Marianne, and Sendhil Mullainathan, 2001, Are CEOS Rewarded for Luck? The Ones without Principals Are, *The Quarterly Journal of Economics* 116(3), 901–932.

Bessen, James, 2017, Automation and jobs: When technology boosts employment, *Boston Univ. School of Law, Law and Economics Research Paper* .

Biais, Bruno, Thomas Mariotti, Jean-Charles Rochet, and Stéphane Villeneuve, 2010, Large risks, limited liability, and dynamic moral hazard., *Econometrica* 78, 73–118.

Brémaud, Pierre., 1981, *Point processes and queues: martingale dynamics* (New York: Springer-Verlag).

Brynjolfsson, Erik, and Andrew Mcafee, 2014, *The Second Machine Age - Work, Progress, and Prosperity in a Time of Brilliant Technologies* (W. W. Norton Company).

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson, 2017, Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics, *National Bureau of Economic Research No. w24001* .

Daley, D.J., and D. Vere-Jones, 1989, *An introduction to the Theory of Point Processes* (Springer).

Décamps, Jean Paul, and Stéphane Villeneuve, 2007, Optimal dividend policy and growth option, *Finance and Stochastics* 11, 3–27.

Demarzo, Peter M., Michael J. Fishman, Zhiguo He, and Neng Wang, 2012, Dynamic Agency and the q Theory of Investment, *Journal of Finance* 67, 2295–2340.

DeMarzo, Peter M., and Yuliy Sannikov, 2006, Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model, *Journal of finance* 61(6).

Ford, Martin, 2015, *The Rise of the Robots - Technology and the Threat of Mass Unemployment* (Oneworld Publications).

Francis, Bill, Iftekhar Hasan, Kose John, and Zenu Sharma, 2013, Asymmetric benchmarking of pay in firms, *Journal of Corporate Finance* 23, 39–53.

French, Kenneth, 2008, Presidential Address: The Cost of Active Investing, *Journal of Finance* 73, 1537–1573.

Frey, Carl Benedikt, and Michael Osborne, 2017, The Future of Employment: How Susceptible Are Jobs to Computerisation?*, *Technological Forecasting and Social Change* 114, 254–280.

Garvey, Gerald T, and Todd T Milbourn, 2006, Asymmetric benchmarking in compensation : Executives are rewarded for good luck but not penalized for bad, *Journal of Financial Economics* 82(1), 197–225.

Goos, Maarten, and Alan Manning, 2007, Lousy and lovely jobs: The rising polarization of work in britain, *The Review of Economics and Statistics* 89(1), 118–133.

Goos, Maarten, Alan Manning, and Anna Salomons, 2014, Explaining job polarization: Routine-biased technological change and offshoring, *American Economic Review* 104, 2509–2526.

Graetz, Georg, and Guy Michaels, 2018, Robots at work, *Review of Economics and Statistics* 100(5), 753–768.

Grenadier, S R, and N Wang, 2005, Investment timing, agency, and information, *Journal of Financial Economics* 75, 493–533.

Gryglewicz, Sebastian, Barney Hartman-Glaser, and Geoffery Zheng, 2019, Growth Options, Incentives, and Pay-for-Performance : Theory and Evidence, *Management Science* .

He, Zhiguo, 2009, Optimal executive compensation when firm size follows geometric brownian motion, *Review of Financial Studies* 22, 859–892.

Hoffmann, Florian, and Sebastian Pfeil, 2010, Reward for Luck in a Dynamic Agency Model, *Review of Financial Studies* 23(9), 3329–3345.

Holmstrom, Bengt, 1979, Moral Hazard and Observability, *The Bell Journal of Economics* 10, 74.

Li, Rui, 2017, Dynamic agency with persistent observable shocks, *Journal of Mathematical Economics* 71, 74–91.

Philippon, Thomas, 2018, The fintech opportunity, *NBER Paper* .

Sannikov, Yuliy, 2008, A Continuous-Time Version of the Principal-Agent, *Review of Economic Studies* 75(3), 957–984.

Spear, Stephen E, and Cheng Wang, 2005, When to fire a CEO: Optimal termination in dynamic contracts, *Journal of Economic Theory* 120, 239–256.

Appendix A. Optimality of the Full-Effort Strategy

In this section, we provide necessary and sufficient conditions for the optimality of the high-effort strategy. It is derived directly from the Proposition 8, section III of DeMarzo and Sannikov (2006) which is our baseline model.

Assume that an optimal contract lets the asset manager shirk on a small period $[t; t+dt)$ without terminating the contract. Then, the agent enjoys a private benefit Bdt while the asset is expected to earn no value during this period. The dynamics of the asset manager's continuation value on $[t, t+dt)$ satisfies

$$dW_t^{\Pi} = \gamma W_t^{\Pi} dt - Bdt + \delta_t (dN_t - \lambda dt) 1_{n_t = 0} \tag{A1}$$

Therefore, allowing the agent to shirk is never profitable if and only if

$$V_i(W_t) > e^{-rdt} V_i(W_t + dW_t) \quad \text{where } i = \{0; 1\} \tag{A2}$$

A necessary and sufficient condition for the optimality of the full-effort strategy is

$$\begin{cases} \min\limits_{w \in [0, \bar{W}^0]} \left\{ V_0(w) + \frac{\gamma}{r} \left(\frac{B}{\gamma \bar{a}} - w \right) V_0'(w) \right\} & \geq 0; \\ \min\limits_{w \in [0, \bar{W}^1]} \left\{ V_1(w) + \frac{\gamma}{r} \left(\frac{B}{\gamma \bar{a}} - w \right) V_1'(w) \right\} & \geq 0 \end{cases} \tag{A3}$$

Given $\bar{a}$, this conditions impose an upper boundary on B, so the private benefit of shirking has to be not too large.

Appendix B. Probabilistic background of the model

Here, we define formally the probability measure induced by any effort process $(a_t)_t$, coming from both the observation of the asset value process and from the observation of

the availability of the automation technology up to date t. We show its equivalence to the standard Weiner measure $\mathbb{P}^0$ on the classical Weiner Space $\Omega = C([0, +\infty), \mathbb{R})$, the set of all continuous real functions that takes their values in $[0, +\infty)$. Let (Z_t^0) be a $\mathcal{F}_t$-Brownian motion under $\mathbb{P}^0$, where $(\mathcal{F}_t)_t$ is the completion of the natural filtration generated by (Z_t^0). Under $\mathbb{P}^0$, we assume that the dynamics of the cash-flow process evolves as

$$dX_t = dZ_t^0$$

Thus, $\mathbb{P}^0$ corresponds to the probability distribution of the cash flows when no active management or effort is exerted. It can be the case when either the agent shirks or when he is laid off. Besides, one may consider $(\mathcal{H}_t)_t$, the completion of the natural filtration generated by the advent of the automation technology, such that

$$\mathcal{H}_t = \begin{cases} 0 \text{ if } t < T; \\ 1 \text{ otherwise.} \end{cases}$$

We call $(\mathcal{G}_t)_t = (\mathcal{F}_t \vee \mathcal{H}_t)_t$ the information set at date t. For any effort strategy $a = (a_t)_{t \geq 0}$, which is assumed to be $\mathcal{G}$-adapted and that takes its values in $\{0, \bar{a}\}$, we define a G-predictable process

$$\eta_t(a) = exp(-\int_0^t (a_s \mu) dZ_s^0 - \frac{1}{2} \int_0^t (a_s \mu)^2 ds)$$

$(\eta_t(a))_{t \geq 0}$ is a G_t-martingale as the effort process takes its values in a bounded interval. Its expectation equals 1 when no effort is exerted. A probability measure $\mathbb{P}^a$ on Ω can then be defined as

$$\frac{d\mathbb{P}^a}{d\mathbb{P}^0} \mid G_t = \eta_t(a) \tag{B1}$$

Assuming enough integrability conditions, the process (Z^a_t) defined as

$$Z^a_t = Z^0_t + \int_0^t a_s \mu ds \tag{B2}$$

is a Brownian motion under $\mathbb{P}^a$. Then, any effort strategy $a = (a_t)_{t \leq \tau}$ induces a probability measure $\mathbb{P}^a$ on Ω for which the dynamics of the cash flows is given by (1).

Consequently, we have that

- The asset manager's expected value of shirking forever from date t with respect to the filtration G_t is given by:

$$\mathbb{E}^0 \left[\int_t^{+\infty} e^{-\gamma(s-t)} \frac{B}{\bar{a}} ds \mid G_t \right] = \frac{B}{\bar{a}\gamma} \tag{B3}$$

- The expected value of the asset if $\bar{a}$ is enforced forever from date t is given with respect to the filtration G_t by:

$$\mathbb{E}^{\bar{a}} \left[\int_t^{+\infty} e^{-r(s-t)} (\bar{a}\mu ds + dZ^{\bar{a}}_s) \mid G_t \right] = \frac{\bar{a}\mu}{r}, \tag{B4}$$

- The expected asset value from automating irreversibly at date t is given with respect to the filtration G_t by:

$$M := \mathbb{E}^0 \left[\int_t^{+\infty} e^{-r(s-t)} (m\mu ds + dZ^0_s) \mid G_t \right] = \frac{m\mu}{r}, \tag{B5}$$

- The expected value of the forthcoming automation, seen from t before the technological

advent is given with respect to the filtration G_t by

$$\mathbb{E}\left[e^{-r(T-t)}M \mid G_t\right] = \frac{\lambda}{\lambda+r}M. \tag{B6}$$

Appendix C. Omitted Proofs

Proof. of Proposition 1

Assume that $M \leq \frac{\bar{a}\mu}{r}$, then the optimal contract in the first-best benchmark (7)-(8) is given by $\Pi^{FB} = (\tau = +\infty, U_0 = w_0)$. Hence, the principal is better off delegating to the agent forever, and he never automates. If the principal cannot get a pledgeable income larger then the value to wait for the technology to arise (i.e. if $\frac{\bar{a}\mu}{r} - w_0 < \frac{\lambda}{\lambda+r}M_0$), he does not delegate to the asset manager.

Now, assume that $M > \frac{\bar{a}\mu}{r}$, then the principal is better off automating at $\tilde{T}$ and the optimal contract in the first-best benchmark is $\Pi^{\tilde{FB}} = (\tau = \tilde{T}, U_0 = w_0)$, and so exhibits a stochastic termination at $\tilde{T}$. The delegation to the agent on $[0, \tilde{T})$ followed by the robot-driven asset generates a the total benefit that is given by

$$\left[\int_0^{+\infty} e^{-rs}\bar{a}\mu ds + e^{-r(s-\tau)}rM ds\right] = \frac{1}{\lambda+r}\bar{a}\mu + \frac{\lambda}{\lambda+r}M \tag{C1}$$

Again, the principal offers such a contract to the agent if and only if it generates a positive pledgeable income over $[0, \tilde{T})$, i.e. if and only if $\frac{1}{\lambda+r}\bar{a}\mu \geq w_0$. Otherwise, the principal waits for the advent of the automation technology and automates from $\tilde{T}$.

□

Proof. of Lemma 1

At any time t, the agent's total expected value from the incentive-compatible contract Π and an agent's continuation value W_t^{Π} – the sum of the prior earnings on $[0, t]$ and the

expected future earnings over the time interval $[t, \tau]$– is given by

$$\Upsilon_t^a = \int_0^t e^{-\gamma s}(dU_s + \frac{B}{\bar{a}}(\bar{a} - a_s)ds) + e^{-\gamma t}W_t^{\Pi} \qquad \text{for } t \leq \tau \tag{C2}$$

$$= \mathbb{E}^a \left[\int_0^{\tau^-} e^{-\gamma s}(dU_s + \frac{B}{\bar{a}}(\bar{a} - a_s)ds) + e^{-\gamma \tau}\Delta U_\tau \right] \qquad \text{for } t \leq \tau \tag{C3}$$

It is an uniformly-integrable $\mathcal{G}$-martingale under the probability measure $\mathbb{P}^a$.

Now, let us consider the process

$$(N_t - \int_0^{t\wedge\tilde{T}} \lambda ds)_t, \quad \forall t \leq \tilde{T}. \tag{C4}$$

It is a $\mathcal{H}$-martingale following the theory of point processes, where $\mathcal{H}_t$ is the information set available at date t as it has been introduced in section B of the appendix that concerns the probabilistic background of the model. Thus, we can apply the martingale representation theorem and there exists a unique $\mathcal{G}$-predictable and square-integrable pair of processes $(\beta^{\Pi}, \delta^{\Pi})$ associated with the incentive-compatible contract Π such that, for all $t \leq \tau$

$$\Upsilon_t^a = \Upsilon_0^a + \int_0^t e^{-\gamma s}\beta_s^{\Pi} dZ_s^a + \int_0^{t\wedge\tilde{T}} e^{-\gamma s}\delta_s^{\Pi}(dN_s - \lambda ds) \tag{C5}$$

Applying the Itô's formula yields to (11). □

Proof. of Lemma 2

Again, let us consider the agent's total expected value at a date t for an incentive-compatible contract Π and given an agent's continuation value W_t^{α} that satisfies the dynamics given by

(13). This former is defined as a stochastic process R_t^a that follows the dynamics

$$R_t^a = \int_0^t e^{-\gamma s}(dU_s + \frac{B}{\bar{a}}(\bar{a} - a_s)ds) + e^{-\gamma t} W_t^{\alpha} \tag{C6}$$

we must determine the function f for $(R_t^a)_t$ to be a $\mathbb{P}^a$-supermartingale $\forall\ a$.

$$\begin{aligned} dR_t^a &= e^{-\gamma t}\left(dU_t + \frac{B}{\bar{a}}(\bar{a} - a_t)dt + dW_t^{\alpha} - \gamma W_t^{\alpha} dt\right) && \text{(C7)} \\ &= e^{-\gamma t}\left(\frac{B}{\bar{a}}(\bar{a} - a_t)dt + f(\beta_t)dt + \beta_t dZ_t^{\bar{a}} + \delta_t(dN_t - \lambda dt)1_{t \le \tilde{T}}\right) && \text{(C8)} \\ &= e^{-\gamma t}\left(\frac{B}{\bar{a}}(\bar{a} - a_t)dt + f(\beta_t)dt + \beta_t(dZ_t^a - (a_t - \bar{a})dt) + \delta_t(dN_t - \lambda dt)1_{t \le \tilde{T}}\right) && \text{(C9)} \\ &= e^{-\gamma t}\left(\left((\frac{B}{\bar{a}} - \beta_t)(\bar{a} - a_t) + f(\beta_t)\right)dt + \beta_t dZ_t^a + \delta_t(dN_t - \lambda dt)1_{t \le \tilde{T}}\right) && \text{(C10)} \end{aligned}$$

Thus, $(R_t^a)_t$ is a supermartingale under $\mathbb{P}^a$ if and only if

$$\begin{aligned} f(\beta) &:= \underset{a \in \{0, \bar{a}\}}{inf}\left((\bar{a} - a)(\beta - \frac{B}{\bar{a}})\right) && \text{(C11)} \\ &= \begin{cases} \bar{a} & \text{if } \beta \ge \frac{B}{\bar{a}} \\ 0 & \text{otherwise} \end{cases} && \text{(C12)} \end{aligned}$$

Then, R_t^a is a supermartingale under $\mathbb{P}^a$ and $R_t^{a^*}$ is a martingale under $\mathbb{P}^{a^*}$. The contract $((U_t)_t, (\beta_t)_t, (\delta_t)_t)$ is incentive compatible with $a_t^*(\beta) = \bar{a} 1_{\beta_t \ge \underline{\beta}}$, where $\underline{\beta} = \frac{B}{\bar{a}}$. Therefore,

$$\begin{aligned} \mathbb{E}^a\left[\int_0^{\tau^- \wedge \tau_0} e^{-\gamma s}(dU_s + \frac{B}{\bar{a}}(\bar{a} - a_s)ds) + e^{-\gamma(\tau \wedge \tau_0)} \Delta U_{\tau \wedge \tau_0}\right] &\le R_0^a && \text{(C13)} \\ &= W_0^{\alpha} && \text{(C14)} \\ &= R_0^{a^*} && \text{(C15)} \end{aligned}$$

And as it is a $R_t^{a^*}$ is a martingale under $\mathbb{P}^{a^*}$,

$$R_0^{a^*} = \mathbb{E}^{a^*}\left[\int_0^{\tau^- \wedge \tau_0} e^{-\gamma s}(dU_s + \frac{B}{\bar{a}}(\bar{a} - a_s^*)ds) + e^{-\gamma(\tau \wedge \tau_0)}\Delta U_{\tau\wedge\tau_0}\right]. \tag{C16}$$

Therefore, if in addition $W_0^\alpha \geq w_0$, the contract $((U_t)_t, (\beta_t)_t, (\delta_t)_t)$ is a candidate solution to the principal's problem (5)-(6). □

Proof. of Proposition 2

For now, we consider that the automation technology is always available. For simplicity and without loss of generality, we set $\tilde{T} = 0$. The dynamics of the agent's continuation value follows:

$$dW_t^\beta = (\gamma W_t^\beta - -\frac{B}{\bar{a}}(\bar{a} - a_t))dt + \beta dZ_t^a - dU_t, \quad \forall \quad t \leq \tau \wedge \tau_0.$$

We start by proving that the value function V_1 satisfies the dynamic programming principle.

Lemma 3: *The value function V_1 satisfies the dynamic programming principle :*

$$V_1(w) = \sup_{\beta,\tau} \mathbb{E}_w\left[\int_0^{(\tau\wedge\tau_0)^-} e^{-rs}\left(\bar{a}1_{\beta\geq\underline{\beta}}\mu ds - dU_S\right) + e^{-r(\tau\wedge\tau_0)}\left(M - W_{\tau\wedge\tau_0}^\beta\right)\right] \tag{C17}$$

We now derive the value function V_1 as a mixed optimal control / stopping problem. Let us consider

$$\Phi(w) = \sup_{\beta,\tau} \mathbb{E}_w\left[e^{-r(\tau\wedge\tau_0)}\Psi(W_{\tau\wedge\tau_0})\right]. \tag{C18}$$

where Ψ is given by:

$$\Psi(w) = max\left(M - w, \tilde{F}_\oslash(w)\right) \tag{C19}$$

and $\tilde{F}_{\oslash}$ satisfies the Stochastic Differential Equation

$$0 = -r\tilde{F}_{\oslash}(w) + \bar{a}1_{\beta \geq \underline{\beta}}\mu + \gamma w \tilde{F}'_{\oslash}(w) + \frac{\beta^2}{2}\tilde{F}''_{\oslash}(w) \tag{C20}$$

together with the conditions at the boundaries $\tilde{F}_{\oslash}(0) = 0$, $\tilde{F}'_{\oslash}(\bar{W}_{\oslash}) = -1$, and $\tilde{F}''_{\oslash}(\bar{W}_{\oslash}) = 0$. We show that the following holds

Lemma 4: *For all* w*,* $V_1(w) = \Phi(w)$

This implies that the principal only considers two strategies: (i) ignore the automation technology and design a contract in a setting à la DeMarzo and Sannikov (2006); (ii) pays what is owed to the agent – his current continuation value– and implements the automation technology with value M.

Proof. Fix $\beta = \underline{\beta}$, then the problem reduces to an optimal stopping problem and the equality holds following the theorem 3.1 in Décamps and Villeneuve (2007). We still have to show that it is true for any value of the process β. It is done by proving that $\left(e^{-r(t\wedge\tau_0)}V_1(W_{t\wedge\tau_0})\right)_{t\geq 0}$ is the Snell envelope of $\Psi(.) = max\left(M - ., \tilde{F}_{\oslash}(.)\right)$.

First, we consider the solution to the baseline model of DeMarzo and Sannikov (2006) provided in section III Proposition 7, and noted F_M. It is concave and solves

$$0 = -rF_M(w) + \bar{a}\mu + \gamma w F'_M(w) + \frac{\beta^2}{2}F''_M(w)$$

together with $F_M(0) = M$, $F'_M(\bar{W}_M) = -1$, $F''_M(\bar{W}_M) = 0$, and it extends linearly with slope -1 after $\bar{W}_M$. Now, we show that $F_M = V_1$.

Lemma 5: *The following holds:* $V_1(w) = \begin{cases} F_M(w) & \text{if } M \leq \frac{\bar{a}\mu}{r} \\ M - w & \text{otherwise} \end{cases}$

Proof. The proof relies on the property of the Snell envelope. Fix $\beta = \underline{\beta}$ and consider a

subsolution to our problem noted $\tilde{V}_1$ and given by

$$\tilde{V}_1(w) = \sup_{\tau} \mathbb{E}_w \left[\int\limits_0^{(\tau\wedge\tau_0)^-} e^{-rs} \left(\bar{a}\mu ds - dU_S\right) + e^{-r(\tau\wedge\tau_0)} \left(M - W^{\underline{\beta}}_{\tau\wedge\tau_0}\right) \right] \tag{C21}$$

It is straightforward to see that whenever $M \geq \frac{\bar{a}\mu}{r}$, then it optimal to set $\tau = 0$. Consequently, there is no contract offered to the agent and the terminal payment is $W^{\underline{\beta}}_{0^-} = 0$. Furthermore, we have by definition and for any given τ

$$\tilde{V}_1(w) \geq \mathbb{E}_w \left[\int\limits_0^{(\tau\wedge\tau_0)^-} e^{-rs} \left(\bar{a}\mu ds - dU_S\right) + e^{-r(\tau\wedge\tau_0)} \left(M - W^{\underline{\beta}}_{\tau\wedge\tau_0}\right) \right] \tag{C22}$$

$$= \mathbb{E}_w \left[\int\limits_0^{(\tau\wedge\tau_0)^-} e^{-rs} \left(\bar{a}\mu ds - dU_S\right) + e^{-r(\tau\wedge\tau_0)} \tilde{V}_1(W^{\underline{\beta}}_{\tau\wedge\tau_0}) \right] \tag{C23}$$

The strategy with $U_s = 0$ for all s up to t and $\tau = t$ gives that

$$\tilde{V}_1(w) \geq \mathbb{E}\left(e^{-r(t\wedge\tau_0)} \tilde{V}_1(W^{\underline{\beta}}_{t\wedge\tau_0})\right) \tag{C24}$$

According to the Markov property, $(e^{-r(t\wedge\tau_0)}\tilde{V}_1(W^{\underline{\beta}}_{t\wedge\tau_0}))_{t\geq 0}$ is a supermartingale which dominates Ψ. From Lemma (4), we have $\tilde{V}_1(w) = ess\sup_{\tau} \mathbb{E}_w \left[e^{-r(\tau\wedge\tau_0)}\Psi(W^{\underline{\beta}}_{\tau\wedge\tau_0})\right]$ for all w, so $\tilde{V}_1$ is the Snell envelope of Ψ when $\beta = \underline{\beta}$.

Moreover, we have from DeMarzo and Sannikov (2006) :

- $(e^{-r(t\wedge\tau_0)}F_M(W^{\underline{\beta}}_{t\wedge\tau_0}))_{t\geq 0}$ is a supermartingale,
- F_M is concave,
- and F_M dominates Ψ, as $F_M(0) = \Psi(0)$, $F'_M(w) \geq -1$ and $F_M \geq F_{\oslash}$.

By definition of the Snell envelope, $\tilde{V}_1 = F_M$.

As according to DeMarzo and Sannikov (2006) $(e^{-r(t\wedge\tau_0)}F_M(W^{\beta}_{t\wedge\tau_0}))_{t\geq 0}$ is a supermartin-

gale for any β, we have that $F_M = \sup_{\beta,\tau} \mathbb{E}_w \left[e^{-r(\tau\wedge\tau_0)} \Psi(W^{\beta}_{\tau\wedge\tau_0}) \right]$, which leads to the desired result. □

Hence, F_M is the Snell envelope of Ψ. So it extends the result of Décamps and Villeneuve (2007) for any value of the control process β. □

□

Proof. of Proposition 3

When $M_0 \leq \frac{\bar{a}\mu}{r}$, we have derived an optimal contract where the optimal sensitivity parameter $\delta(W_{t^-})$ is not differentiable, as it is shown in Figure 2. Therefore, we cannot apply directly the proof provided in Hoffmann and Pfeil (2010) or Demarzo et al. (2012) to our Proposition, and we provide here an alternative proof that the value function $V_0(w)$ is concave. Then, we verify that it corresponds indeed to the principal's value function before the advent of the automation technology.

- **Step 1 – Concavity :**

From the boundary condition, there exists $\epsilon > 0$ such that:

$$V_0''(\bar{W}^0 - \epsilon) > V_0''(\bar{W}^0) \tag{C25}$$

If we assume that V_0 is concave close to the boundary $\bar{W}^0$, then in addition

$$0 > V_0''(\bar{W}^0 - \epsilon) \tag{C26}$$

Now, let us show that it implies that the function is concave over the whole interval $[\frac{B}{\gamma}; \bar{W}^0]$. In order to do so, let us assume that there exists $\tilde{W} := \sup_{w\in[\frac{B}{\gamma};\bar{W}^0]} \{V_0''(W) \geq 0\}$. We have by continuity that $V_0''(\tilde{W}) = 0$ while $V_0''(\tilde{W} + h) < 0$, for a small $h > 0$ taken such that $(\tilde{w} + h)V_0'(\tilde{w} + h) = \tilde{w}V_0'(\tilde{w})$. From (22) we have that $V_0'(\tilde{w}) > 0$.

We can also write the following expression for the difference quotient:

$$
\begin{aligned}
(r+\lambda)\left[\frac{V_0(\tilde{w}+h)-V_0(\tilde{w})}{h}\right] = \\
\frac{1}{h}[\gamma(\tilde{w}+h)-\lambda\delta(\tilde{w}+h))V_0'(\tilde{w}+h)-(\gamma\tilde{w}-\lambda\delta(\tilde{w})))V_0'(w) \\
+\underbrace{\frac{\beta^2}{2}V_0''(\tilde{w}+h)}_{<0} \\
+\lambda\underbrace{(V_1(w+h+\delta(w+h))-V_1(w+\delta(w)))}_{\text{roughly zero}}]
\end{aligned} \tag{C27}
$$

So,

$$
\begin{aligned}
(r+\lambda)\left[\frac{V_0(\tilde{w}+h)-V_0(\tilde{w})}{h}\right] < \\
\frac{1}{h}\left[\gamma(\tilde{w}+h)-\lambda\delta(\tilde{w}+h)\right)V_0'(\tilde{w}+h)-(\gamma\tilde{w}-\lambda\delta(\tilde{w})))V_0'(w)
\end{aligned} \tag{C28}
$$

Now, we use that $(\tilde{w}+h)V_0'(\tilde{w}+h)=\tilde{w}V_0'(\tilde{w})$, then

$$
\begin{aligned}
(r+\lambda)\left[\frac{V_0(\tilde{w}+h)-V_0(\tilde{w})}{h}\right] < \\
\frac{1}{h}[-(B+\lambda\delta(\tilde{w}+h))V_0'(\tilde{w}+h)+(B+\lambda\delta(\tilde{w}))V_0'(w)]
\end{aligned} \tag{C29}
$$

Which translates, if $\lambda\delta(w)$ is sufficiently small, to

$$
(r+\lambda)\left[\frac{V_0(\tilde{w}+h)-V_0(\tilde{w})}{h}\right] < -B\left[\frac{V_0'(\tilde{w}+h)-V_0'(\tilde{w})}{h}\right] \tag{C30}
$$

$$
\text{Therefore,}\quad \left[\frac{V_0(\tilde{w}+h)-V_0(\tilde{w})}{h}\right] < 0 \tag{C31}
$$

As it contradicts with the assumption that $V_0'(w)>0$, therefore we have that $V_0''(W)<0$ $\forall w \in [\frac{B}{\gamma};\bar{W^0}]$. We conclude that V_0 is concave over the whole employment interval as long

as it is concave close to $\bar{W}^0$. The complete proof of the concavity of the value function in this case is not provided yet.

- **Step 2 – Verification :** As usual in dynamic contracting theory, our last step is to verify that we have indeed derived an optimal contract. For any incentive-compatible contract, let us define
$\mathrm{F}_t = \int_0^t e^{-rt}(a^*\mu dt - dU_t) + e^{-rt}V_1(W_t,)1_{\{t\geq T\}} + e^{-rt}V_0(W_t)1_{\{\tilde{T}>t>\tau\}}$
$+ e^{-rt}\left(\frac{\lambda}{\lambda+r}M\right)1_{\{t\leq\tau\}}$

By Itô's Lemma, its drift is

$$e^{-rt}\left(\frac{a^*\mu}{r} - M\right) \tag{C32}$$

which is always negative as here $M \geq \frac{a^*\mu}{r}$. Therefore, F_t is a supermartingale and

$$V_0(W_0) = F_0 \geq \mathbb{E}^{a^*}[F_t \mid \mathcal{G}_t] \tag{C33}$$

with an equality for the contract derived in the Proposition 3. Then, the optimal choice of sensitivity δ satisfies:

$$V_0'(w) = V_1'(w+\delta) \tag{C34}$$

as long as it remains larger than $-w$ to fulfill the limited-liability condition.

When $M_0 > \frac{\bar{a}\mu}{r}$, then $M_0 - w$ dominates the largest solution of the HJB Equation (16) for $n = 0$. Consequently, the principal is better off solely waiting for the robot without offering a contract to the agent and implements the robot-driven asset management from $\tilde{T}$.

Proof. of Proposition 4

First, we show that $\delta(w) < 0$ in the neighbourhood of 0. We call V_1^M the principal's value

function associated to the contract à la DeMarzo and Sannikov (2006) together with a value of automation M. It is the optimal contract after the advent of technology. Because this contract does not make the agent sensitive to the advent of robots, it is a subsolution to our problem. Consequently, $V_0(0) = V_1^{M_0}(0) = M_0$ together with $V_0'(0) > V_1^{M_0}{'}(0)$. Now, we use that $\frac{dV_1(0)}{dM} \leq 0$ from Table A1 of Explicit Comparative Statics Calculations in DeMarzo and Sannikov (2006). Consequently, $V_1^{M_0}{'}(0) > V_1^{M}{'}(0)$, and we have $V_0'(0) > V_1^{M}{'}(0)$. Thus, $\delta(W) < 0$ in the neighbourhood of 0 because value functions are concave.

Now, we prove the existence of $\hat{W} = \inf\{w \mid V_1^{M}{'}(0) = V_0'(w)\}$. By construction, $V_1^{M}{'}(0) \geq -1$, so we have $V_0^{M_0}{'}(0) \geq V_1^{M}{'}(0) \geq V_0'(\bar{W^0})$, where $\bar{W^0}$ is such that $V_0^{M_0}{'}(\bar{W^0}) = -1$. Because V_0 is concave, we can apply the Intermediate Value Theorem to conclude that $\hat{W}$ exists. By limited-liability, we impose that $\delta(w) = -w$, $\forall w \leq \hat{W}$. Furthermore, there exists by construction $\bar{W^1}$ such that $V_1{'}(0) = -1$. To show that $\delta(w) \leq 0$, $\forall w$, it remains to show that $\bar{W^1} < \bar{W^0}$, which is analogous to the proof of Lemma B2 in Hoffmann and Pfeil (2010). $\square$

Proof. of Proposition 5

Assume that the value of the technology jumps from $M \geq \frac{\bar{a}\mu}{r}$ to $M + \epsilon$. Then, it leads $V_1(w)$ to increase to $V_1(w) + \epsilon$. As the value function V_0 satisfies Equation 22, V_0 increases by $\frac{\lambda}{\lambda+r}\epsilon$, and $V_0(0) = \frac{\lambda}{\lambda+r}(M + \epsilon)$. Consequently, V_0 moves upwards by $\frac{\lambda}{\lambda+r}\epsilon$ when $M \geq \frac{\bar{a}\mu}{r}$ increases by ϵ, which leads to the desired result. $\square$

Proof. of Proposition 6

The proof of $\bar{W^1} > \bar{W^0}$ relies on analogous arguments than the proof of Lemma B2 in Hoffmann and Pfeil (2010). To show that $V_1'(0)$ increases with θ, we start by showing that V_e increases with θ. Let us examine 2 value functions $V_e^{\underline{\theta}}$ and $V_e^{\bar{\theta}}$ associated to $\underline{\theta}$ and $\bar{\theta} > \underline{\theta}$. We have $V_e^{\bar{\theta}} = V_e^{\underline{\theta}} = 0$ together with $V_e^{\bar{\theta}}(W) > V_e^{\underline{\theta}}$ for $W > 0$. As (1) we have by construction $V_1(b) = V_e^{\underline{\theta}}(b)$, while (2) $V_1(0) = M$, the concavity of V_1 leads to the desired result.

$\square$

Appendix D. Algorithm

Finding an optimal contract is a free-boundary problem as both the principal's value function and the payment boundary are unknown. We solve for $(V_n, \bar{W}_n$ where $n = 0; 1$ using the shooting method.

1. First, we solve for $(V_1, \bar{W}^1)$,
2. Then, we set $\delta(w) = -w$,
3. (a) Given δ, we solve for $(V_0, \bar{W}^0)$,
   (b) For each w such that $V_0'(w) \leq V_1'(0)$ we set δ such that $V_0'(w + \delta(w)) = V_1'(w)$, and otherwise we let $\delta(w) = -w$.
4. We repeat step 3 until convergence, with the convergence criteria :

$$\mid V_0^i(w) - V_0^{i-1}(w) \mid < 10^{-3}, \tag{D1}$$

where i stands for the number of iterations already made.