FUNDAÇÃO GETULIO VARGAS
ESCOLA de PÓS-GRADUAÇÃO em ECONOMIA

Pedro de Bitencourt Melgaré

Trading off static efficiency and dynamic incentives in apprenticeship relations

Rio de Janeiro
2021

Pedro de Bitencourt Melgaré

Trading off static efficiency and dynamic incentives in apprenticeship relations

Dissertação para obtenção do grau de mestre apresentada à Escola de Pós-Grauação em Economia

Área de concentração: Teoria Econômica

Orientador: Lucas Jóver Maestri

Rio de Janeiro
2021

Dados Internacionais de Catalogação na Publicação (CIP)
Ficha catalográfica elaborada pelo Sistema de Bibliotecas/FGV

Melgaré, Pedro de Bitencourt
Trading off static efficiency and dynamic incentives in apprenticeship relations / Pedro de Bitencourt Melgaré. – 2021.
22 f.

Dissertação (mestrado) - Fundação Getulio Vargas, Escola Brasileira de Economia e Finanças.
Orientador: Lucas Jóver Maestri.
Inclui bibliografia.

1. Economia. 2. Contratos incompletos. 3. Capital Humano. 4. Transferência de aprendizagem. I. Maestri, Lucas Jóver. II. Fundação Getulio Vargas. Escola Brasileira de Economia e Finanças. III. Título.

CDD – 330.1

Elaborada por Márcia Nunes Bacha – CRB-7/4403

PEDRO DE BITENCOURT MELGARÉ

“TRADING OFF STATIC EFFICIENCY AND DYNAMIC INCENTIVES IN APPRENTICESHIP RELATIONS’’.

Dissertação apresentado(a) ao Curso de Mestrado em Economia do(a) EPGE Escola Brasileira de Economia e Finanças - FGV EPGE para obtenção do grau de Mestre(a) em Economia.

Data da defesa: 29/03/2021

ASSINATURA DOS MEMBROS DA BANCA EXAMINADORA

Presidente da Comissão Examinadora: Profº/ª **Lucas Jóver Maestri**

Lucas Jóver Maestri
Orientador

P/

Leandro Gorno
Membro

P/

Daniel Monte
Membro

Nos termos da Lei nº 13.979 de 06/02/20 - DOU nº 27 de 07/02/20 e Portaria MEC nº 544 de 16/06/20 - DOU nº 114 de 17/06/20 que dispõem sobre a suspensão temporária das atividades acadêmicas presenciais e a utilização de recursos tecnológicos face ao COVID-19, as apresentações das defesas de Tese e Dissertação, de forma excepcional, serão realizadas de forma remota e síncrona, incluindo-se nessa modalidade membros da banca e discente.

Lucas Jóver Maestri
Coordenador

Antonio de Araujo Freitas Junior
Pró-Reitor de Ensino, Pesquisa e Pós-Graduação FGV

Antonio Freitas, PhD
Pró-Reitor de Ensino, Pesquisa e Pós-Graduação
Fundação Getúlio Vargas

Instrução Normativa nº 01/19, de 09/07/19 - Pró-Reitoria FGV
Em caso de participação de Membro(s) da Banca Examinadora de forma não-presencial*, o Presidente da Comissão Examinadora assinará o documento como representante legal, delegado por esta I.N.
*Skype, Videoconferência, Apps de vídeo etc

Resumo

Neste artigo, estudamos um modelo de treinamento on-the-job em um ambiente de monopsônio com mercado de crédito imperfeito e com ausência de comprometimento. Um equilíbrio de mercado é um contrato relacional que especifica uma sequência de salários, de produtividades para o trabalhador e uma data de terminação que satisfazem as restrições de incentivos da firma e do trabalhador. Mostramos que nenhum contrato auto-sustentável pode terminar em um período em que o excedente da relação é negativo. Se o excedente da relação é sempre positivo, em equilíbrio os contratos são perpétuos e não pagam salário antes do trabalhador estar plenamente treinado. A produtividade esperada do trabalhador em um emprego alternativo cresce exponencialmente à taxa de juros. Para o caso particular de uma função de produtividade Cobb-Douglas, mostramos que o treinamento é mais devagar para maiores coeficientes de produtividade e elasticidade do produto. O efeito da taxa de desconto é ambíguo.

Palavras-chave: Teoria Econômica, Contratos Incompletos, Capital Humano, Relações Mestre-Aprendiz

Abstract

In this paper, we study a model of on-the-job training in a monopsonic environment with no enforceability and with imperfect credit markets. A market equilibrium is a relational contract which specifies a sequence of wages, a sequence of productivities for the worker and a termination date which satisfy the worker's and the firm's incentive constraint and maximize profits. We show that no self-enforceable contract can end in a period in which total surplus from employment is negative. If the surplus of employment is always positive, equilibrium contracts are perpetual and involve no wage payment before the worker is fully trained. The worker's expected productivity in an alternative employment grows exactly at the interest rate. For the particular case of a Cobb-Douglas production function for the agent, we show that training is slower for higher productivity coefficient and output elasticity. The effect of the discount rate is ambiguous.

Keywords: Economic Theory, Incomplete Contracts, Human Capital, Apprenticeships

Sumário

1 Introduction

Unlike a consumer buying a house, a worker that wants to invest in training or education cannot use the recently acquired human capital as collateral for any loan. This fact has important consequences for the literature on human capital, and is studied by economists since at least Becker's early works (Becker (1962)). Credit market imperfections as described may result not only in suboptimal investment in formal education or on-the-job training, as it also can shape the design of institutions that supply human capital. Policies such as federal college loans are a salient example.

On the other hand, on-the-job training and apprenticeship relations additionally feature another complication, regarding contracts enforcement. Even though a worker may be motivated to accept a job by the expectation of learning, it is difficult to enforce any contract that states that his employer will properly teach him. If a firm expects that, once the employment relationship is over, the worker may start his own business or join a competitor, it may be encouraged to either avoid teaching him or to make him engage in menial work and tasks less learning intensive.

This paper theoretically investigates the design of employment relations featuring on-the-job training in a simple environment in which the worker has a liquidity constraint, and contracts are hard to enforce. Job positions are scarcer than the supply of potential worker to fill them, such that in equilibrium the relationship between firm and worker is designed in order to maximize profits for the firm. Since neither part can ensure enforcement, any feasible relationship must take the form of a relational contract, which establishes the future trajectory of training, payment flows and a date of termination in such a way that neither part desires to leave it before previously agreed. The contract may either have a finite duration or it may last forever; the important feature is that this must be mutually agreed to in the initial period.

The absence of credible commitment by both parties restricts the relationship's possibilities in two important ways. For a given flow of payments, training cannot be too fast; otherwise, it might be more advantageous for the worker to leave with the knowledge already acquired than to stay in the relationship. The firm, on the other hand, cannot promise wages too high, for an analogous reason.

After termination, the worker uses his acquired skills to either work for another firm or to start his own business. At the termination date, there is a clear conflict of interest: the worker prefers to be provided with more training, since this increases his productivity outside of the relationship, while the firm prefers him to have fewer skills, since in the next period he becomes a potential competitor.

The impossibility of external enforcement creates a important constraint over the set of feasible contracts. If at the termination date the worker is more productive outside than inside the firm, when properly accounting for any possible externality his competition imposes to the firm, then it

is impossible to motivate both the firm and the worker not to leave the relationship at the begginning of the period. This result is robust to an alternative specification allowing the firm to provide an intensive training to the worker at the termination date, since it has no incentive to do so. Similarly, if the contract is perpetual, in the limit the social surplus has to be greater when the worker is inside rather than outside the firm.

These contractability results imply that any profit maximizing contract is in fact perpetual. We show that in equilibrium, contracts take the form of an apprenticeship, in a result similar to Garicano and Rayo (2017). The worker receives no wage until a pre-specified date in the future which marks the end of his training, and after the end of his training he receives a flow of wages with present value exactly equal to his outside option. There are in general multiple equilibria, since there might be multiple flows of after-training wages which are both incentive compatible and yield the same present value at the end of training. The firm trains the worker to at least the level of productivity that maximizes profits for static contracts. Also, its productivity outside the firm grows exponentially with rate equal to the common interest rate.

Given a date for the end of training, we can specify the path for the worker's productivity, and find the firm's profit associated with this particular contract. The optimal training duration must balance two conflicting effects: a faster training increases production, but also increase the present value of the flow of wages required by the worker to stay in the relationship. The firm's objective of profit maximization leads to a training duration that is greater than in the first-best solution, which involves instantaneous training.

In general, there is no closed-form solution for the optimal training duration if we do not specify the outside option function for the worker. For the case in which the worker has a Cobb-Douglas production function with diminishing returns, we explicitly solve for the optimal duration, and perform some simple comparative statics exercises. For this particular case, a higher productivity coefficient leads to a longer training duration. This is because, in this case, a higher productivity coefficient increases the worker's outside option at his final productivity level more than for intermediary ones. A higher elasticity of output of the worker leads to a higher training length, since it decreases the worker's output function for intermediary productivity levels and rises the rent extraction margin for the firm.

The subject of apprenticeship relations is important for two reasons. First, this institution is of particular interest to economic history, since this institution is believed to have had a meaningful role in the early economic growth in Europe before the Industrial Revolution (De la Croix et al. (2018), Mokyr (2019)). Historically, the main features of our model may have been particularly relevant, since going to a formal court of law in order to enforce a specific contract was expensive. Two, there is anecdotal evidence that many modern day professional relationships seem to be designed to slow the transmission of knowledge in order to incentivize workers to receive wages below

their marginal product. Some salient examples are in the professional services firms, such as law, consulting, and architecture. In these industries, the firms offer a significant general knowledge to junior employees, which work mostly on routine tasks, slowing down their rate of learning.

This project is mainly related to two strands of literature. First, it is related to the literature on principal-agent models with relational contracts Levin (2003), MacLeod and Malcomson (1989). This literature studies implicit contracts, which cannot be enforced by a third party, and therefore must satisfy self-enforceability constraints. More specifically, this project is related to a subset of this literature which study on-the-job-training in the presence of borrowing constraints by the part of the novice Garicano and Rayo (2017), Fudenberg and Rayo (2019). In Garicano and Rayo (2017) (henceforth GR), it is shown that these environments generate an agreement which takes the form of an apprenticeship relation, in which the novice works receiving less than its marginal product in exchange for training. The main differences relative to this project are highlighted in the next section.

The second related strand is the literature on dynamic knowledge or information transfers. Hörner and Skrzypacz (2016) study a setting where a informed party has information regarding the state of the world which is useful for an uninformed party. The optimal way of selling this information is shown to be giving a initial free disclosure and then sell the remainder sequentially in small amounts. Zhao et al. (2020) study a similar setting, in which an informed principal wants to incentivize an agent of exerting effort, and the only instrument it has to do that is through the disclosure of information.

The rest of this paper is structured as follows. Section 2 highlights the contributions of this paper to the literature; section 3 formally describes our model; section 4 explores the first best arrangements; section 5 presents results on contractability; section 6 characterizes the equilibrium and section 7 concludes.

2 Our contribution

Even though our model is closely related to Garicano and Rayo (2017), it's greater generality yields some important new insights. In GR, the agent produces the same output whether inside or outside the firm, and if the relationship ends, the agent's skill level does not influence the firm's profit at all. Moreover, GR assumes away the possibility of competition between the worker and the firm, which may be an object of interest in markets where there are few producers. This way, total production is always equal when the worker is employed or not by the firm, which makes both players indifferent between whether continuing or breaking the relationship, for a given productivity level of the agent.

In this paper, we show that a slight generalization of GR leads to some interesting insights in the problem of apprenticeship relations. Allowing the agent to have an arbitrary production technology

and the firm to have an arbitrary externality from the agent's productivity leads to a natural concept of static efficiency, which is shown to restrict the set of self-enforceable contracts. This leads to the result that there must always exist an equilibrium in which the relationship is perpetual, which is an insight GR is not able to provide, since total production does not change if the worker leaves or stays in the firm.

Lastly, our model also leads to some comparative statics questions regarding the agent's production function. Even though the profit maximizing contract in general has no closed form solution, we provide some simple comparative statics for the case of a Cobb-Douglas production function. We study how the length of the profit maximizing contract varies with the output elasticity, the productivity coefficient of the agent's technology, and the common discount rate.

3 The model

There is one worker and one firm, and both players are risk-neutral. The firm A possesses a fixed stock of general purpose knowledge, which it can use to train the worker without any costs and at any rate. The firm can train the worker to achieve any level of productivity in the set $\mathcal{Y} = [0,y^{\max}]$. The two players interact over a sequence of periods $t = 0,1,...$ and discount the future using the same discount factor δ. Let $r = -log(\delta)$ be the associated interest rate. Also, let y_t be the productivity of the worker at the begginning of period t if working for firm A. At each period t, both players can decide to break the relationship before production happens.

If the relationship ends at time t, the worker and the firm receive a payoff of $\frac{h^f(y_t)}{1-\delta}$ and $\frac{g^f(y_t)}{1-\delta}$, respectively. We assume throughout this paper that $h^f(0) = g^f(0) = 0$, h is a non-decreasing and g is a non-increasing function and that g^f,h^f are continuous functions. In this way, one can interpret the function $h^f(y)$ as the expected flow of income of the worker if he chooses to open his own business or to work for a competitor possessing y stock of knowledge, or as the worker's production function. Since it is non-decreasing, more knowledge cannot hurt the worker in any way in case of a separation. The function $g^f(y)$, on the other hand, can be interpreted to represent the expected flow loss of the firm when the worker leaves with y stock of knowledge. To ease notation, let $h(y) = \frac{h^f(y)}{1-\delta}$ and $g(y) = \frac{g^f(y)}{1-\delta}$ denote the present value of the perpetuities $h^f(y)$ and $g^f(y)$, respectively.

In the first period, both players agree to a self-enforcing contract $\mathcal{C} = \{(y_{t+1},w_t)_{t=0}^{\infty},T\}$. If some of the players deviates, the other one breaks the relationship and both receive payoff equal to their outside option. For our purposes, this is equivalent to studying the best subgame perfect equilibria of a repeated game in which, for each period, both players decide whether to break or not the relationship and the firm decides how much to pay and train the worker (MacLeod and Malcomson (1989)). A contract specifies a termination date $T \in \mathbb{N} \cup \{\infty\}$, a non-decreasing sequence of

productivities $(y_t)_{t=0}^{\infty}$, and a sequence of transfers $(w_t)_{t=0}^{\infty}$, such that $w_t = 0$ and $y_t = y_T$ for all $t > T$. Let

$$\mathbb{C}^0 = \{((y_t,w_t),T) \in \mathcal{Y}^{\infty} \times \mathbb{R}^{\infty} \times \mathbb{N} : w_t = 0, y_t = y_T \forall t > T\}$$

be the space of contracts.

Given a contract $\mathcal{C}$, the payoff of the firm $\Pi_0(\mathcal{C})$ and the payoff of the worker $V_0(\mathcal{C})$ can be defined as following:

$$\Pi_0(\mathcal{C}) = \sum_{t=0}^{T} \delta^t [y_t - w_t] + \delta^{T+1} g(y_T)$$

$$V_0(\mathcal{C}) = \sum_{t=0}^{T} \delta^t w_t + \delta^{T+1} h(y_T)$$

More generally, we can define the continuation payoffs Π_t and V_t as the continuation values for the firm and the worker:

$$\Pi_t(\mathcal{C}) = \sum_{\tau=t}^{T} \delta^{\tau-t} [y_\tau - w_\tau] + \delta^{T+1-t} g(y_T)$$

$$V_t(\mathcal{C}) = \sum_{\tau=t}^{T} \delta^{\tau-t} w_\tau + \delta^{T+1-t} h(y_T)$$

A self-enforcing contract must ensure that both players do not break the relationship before the set termination date T. Thus, the following conditions must hold for all $t \leq T$:

$$\begin{aligned} \Pi_t(\mathcal{C}) &\geq g(y_t) \qquad [FC_t] \\ V_t(\mathcal{C}) &\geq h(y_t) \qquad [WC_t] \end{aligned}$$

We assume that the worker has no access to credit markets and starts with 0 assets. At period t, the present value of all wages received up to period t is $\sum_{\tau=0}^{t} \delta^{\tau-t} w_\tau$. Therefore, in order to be feasible, a contract must also satisfy the following liquidity constraint for $t \leq T$:

$$\sum_{\tau=0}^{t} \delta^{\tau-t} w_\tau \geq 0 \quad [LC_t]$$

We denote by $\mathbb{C}$ the set of self-enforceable feasible contracts; that is, the set of contracts $\mathcal{C}$ which satisfy $[FC_t]$,$[WC_t]$,$[LC_t]$ for all $t \leq T$.

Following GR, we assume that the supply of workers outweighs the supply of jobs. We are thus interested in the contract which maximizes profits for the firm:

$$\mathcal{C}^* = \arg\max_{\mathcal{C}\in\mathbb{C}} \Pi_0(\mathcal{C})$$

We will call such $\mathcal{C}^*$ as both the equilibrium contract and the profit maximizing contract throughout this paper.

4 First best contracts

Since utilities are quasi-linear, the first best contract can be found by maximizing joint surplus. That is, a first best contract must solve

$$\mathcal{C}^{FB} = \arg\max_{\mathcal{C}\in\mathbb{C}^0} \Pi_0(\mathcal{C}) + V_0(\mathcal{C}) = \sum_{t=0}^{T} \delta^t y_t + \delta^{T+1}[g(y_T) + h(y_T)]$$

Let $y_o^* = \arg\max_{y\in\mathcal{Y}} g(y)+h(y)$ be the productivity level that maximizes the total surplus when the worker leaves the firm. The problem above has a straightforward solution, and can be characterized by the following proposition:

Proposition 1 *The first-best contract features $T = \infty$ and $y_t = y^{\max}$ $\forall t \in \mathbb{N}$ if $y^{\max} > g(y_o^*) + h(y_o^*)$ and $y_1 = y_o^*$ if $y^{\max} < g(y_o^*) + h(y_0^*)$.*

To see this, consider the best knowledge path for a perpetual contract. This sequence $(\tilde{y}_t)$ must solve

$$\begin{aligned} (\tilde{y}_t) \in \quad & \arg\max_{(y_t)\in\mathbb{R}^\infty} \sum_{t=0}^{\infty} \delta^t y_t \\ \text{s.t} \quad & (y_t) \text{ is non-decreasing} \end{aligned}$$

, so trivially we have that $\tilde{y}_t = y^{\max}$ for all $t \in \mathbb{N}$. Now, if there is a best knowledge path for a temporary contract (it may not exist), we must have $\frac{y_T}{1-\delta} \leq g(y_T) + h(y_T)$; otherwise, we can

construct an alternative contract with $T' = T + 1$, $y'_t = y_t$ for all $t \leq T$ and $y'_T = y_{T+1}$, which is strictly more profitable. By a similar argument, we must have $T = 1$. The best knowledge path for a temporary contract is then a one-time transfer, and must solve

$$\max_y g(y) + h(y)$$

That is, we can only compare the surplus of the best perpetual contract (which features $y_t = y^{\max}$) with the surplus of the best temporary contract, which is a one-period contract that maximizes the sum of the termination payoffs $g(y) + h(y)$. An important feature of the first-best contract is that, if it is temporary, it must last for only one period. This reflects our assumption that teaching has no cost and has no speed restriction; there is no reason for any delay.

Can the first-best contract be implemented? We show that the answer is no if the first-best contract is temporary; in fact, we show that no temporary self-enforceable contract can be (strictly) optimal. If the first-best contract is perpetual, then it can in fact be implemented, but we will show that it is not, in general, the most profitable contract for the firm.

5 Contractability results

In this section we characterize the space of feasible self-enforcing contracts. If the firm cannot credibly commit to effectively train the worker in the last period of the contract, we show that this restricts the possibility of finite self-enforcing contracts. This restriction turns out to be related to a notion of static efficiency, described below. Results are similar to MacLeod and Malcomson (1989).

Definition. The function $S(y) = \frac{1}{1-\delta}y - [g(y) + h(y)]$ is the relationship static surplus function.

The function $S(y)$ determines the expected excess production of the worker if he works for the firm. The sign of the function $S(y)$ is of particular importance, since it determines whether, for a given level of knowledge, the total production is higher with the worker working for the firm ($S(y) > 0$) or not ($S(y) < 0$).

Definition. For a given knowledge level y, the relationship is statically efficient if $S(y) \geq 0$.

In a first-best world, in which both players can costlessly commit to a contract, there are no learning dynamics, as we have seen in the last section. The firm has no reason to propose a contract where it stays in a statically inefficient relationship for more than one period, which is the minimum time length to effectively train the novice up to his optimal knowledge level.

The following proposition restricts the space of self-enforceable contracts by showing that the in the last period of the relationship, the relationship must be statically efficient:

Proposition 2 *If $\mathcal{C} \in \mathbb{C}$ features $T < \infty$, then for y_T the relationship is statically efficient (i.e,*

$S(y_T) \geq 0$).

At the beginning of the last period of the contract, when deciding whether to continuing the relationship, both players compare what they receive if they stay in the relationship ($y_T - w_T$ for the firm and w_t for the worker) to the interest on delaying by one period the payoff of their outside option ($(1-\delta)g(y_T)$ for the firm and $(1-\delta)h(y_T)$ for the worker). We conclude that the total product of continuing the relationship y_T must be larger than the interest on the product after breaking the relationship $(1-\delta)[g(y_T)+h(y_T)]$. The relationship, therefore, must be statically efficient at productivity y_T.

This result imposes a serious constraint in the space of self-enforceable contracts. This logic is also robust to allowing the firm to train the worker at the end of the last period, if we assume that any additional knowledge given by the firm to the worker hurts the firm in expectation after the relationship ends. In this case, if the firm cannot commit to such training at the very last end of the contract, it would simply not do so, in order to avoid helping its future competitor.

A similar result holds for perpetual contracts.

Proposition 3 *Assume g,h are continuous functions. Then, if $\mathcal{C}$ is self-enforcing and features $T = \infty$, then at $y^* \equiv \sup_{t'} y_{t'}$ the relationship is statically efficient.*

The intuition for this result is virtually identical to the last one: the sequence of productivity levels $(y_t)_{t=1}^{\infty}$ must converge, and if at the limit the relationship is statically inefficient, then it is also for a t large enough forwards. As a consequence, the size of the pie to be distributed is not large enough to offset the opportunity costs of the players, and therefore one of the players must strictly prefer to break the contract.

Given both propositions above, we can obtain the following corollary:

Corollary 1 *Assume g is strictly increasing. If $S(y) < 0$ for all $y > 0$, then there is no self-enforceable contract with a positive knowledge transfer.*

This corollary emphasizes the severity of the contractability difficulty under the absence of credible commitment instruments. Even if the relationship is never statically efficient, it can still provide a net positive surplus for both parties involved. This is due to a divergence between the myopic efficiency concept of $S(y) \geq 0$ and the global, dynamic concept of efficiency present in the first-best contract section.

For the context of apprenticeship relations, this means that there can be no apprenticeship relation if the apprentice always produces more by himself (discounting the competition externalities he imposes to the firm), even though that without being trained by the firm, he has no means to actually producing anything.

Lastly, observe that neither result above prohibits the firm from enforcing a contract in which for some periods the relationship is statically inefficient. The results above restricts the enforceability of contracts with static inefficiency only at the very end of the training process. Self-enforceable contracts can be statically inefficient mid training: future flows of income can motivate the worker and the firm to stay in the relationship, even though the production in this organizational structure is suboptimal.

These results are derived assuming that a firm-worker relationship is a one-time event. If we let the worker observe the history of the firm's past relationships, then reputational effects may enable contracts which end in periods in which the relationship is not efficient. Therefore, these results are particularly important in contexts unfavourable to reputational mechanisms. This is indeed the case for pre-industrial apprenticeships, since even though reputation and trust had some relevance in the apprenticeship institution (De la Croix et al. (2018)), the number of apprentices trained by a particular artisan over his whole career was small, and prospective artisans usually did not took part in more than one apprenticeship (Mokyr (2019)).

6 Characterizing the profit maximizing contract

In this section, we characterize the profit maximizing contract. The optimal contract is shown to have similar features to the one found in GR. The optimal contract is perpetual and takes the form of an apprenticeship relation, in which the worker receives no wage until a finite date T_g, and receives exactly his outside option after that.

Throughout this section, we will assume that $g(y)$,$h(y)$ are continuous functions and that $h(y)$ is strictly increasing.

By the previous section, we know that if a temporary contract $\mathcal{C}$ is self-enforceable, it features $S(y_T) \geq 0$. If $S(y_T) > 0$, then clearly $\mathcal{C}$ cannot be profit maximizing, since we could then extend it by only repeating infinitely the last period and paying the worker the interest of his outside option flow $w = (1-\delta)h(y_T)$. This would yield the worker the same payoff as the original contract, and since the relationship is strictly statically efficient, the firm would derive a strictly higher profit: $y_T - (1-\delta)h(y_T) > g(y_T)$. On the other hand, if $S(y_T) = 0$, then the alternative contract described above would yield the same payoff for both worker and firm. We conclude that without any loss we can consider only perpetual contracts if our goal is to maximize profits.

This is in itself a remarkable fact, to the extent that this means that, in equilibrium, relationships either provide no positive surplus for neither party or it lasts forever. Apprenticeship relations in this environment are designed in a way that neither the worker has any incentive to search another firm or establish his own business, nor the firm has the incentive to fire and replace the worker. This is remarkably distinct from apprenticeship relations as observed in the economy, which usually are

temporary, even in setting in which contracts are hard to enforce. As discussed in the last section, this disparity may be explained by the absence of any reputational mechanism in our model.

Some additional notation and definitions will be convenient in this section. First, let $R(y) = \frac{1}{1-\delta}y - h(y)$ be the firm rent function, and $R_a(y) = \frac{1}{1-\delta}y - g(y)$ be the worker's rent function. For any given contract $\mathcal{C} = \{(y_t,w_t)_{t=1}^{\infty},T\}$, let $y_\infty = \lim_t y_t$ be the productivity limit, and let $T_g = \min\{t \in \mathbb{N} : y_t = y_{t'} \text{ for all } t' \geq t\}$ be the worker's graduation date. Lastly, let $W_t = \sum_{\tau=1}^{t-1} \delta^{t-\tau} w_\tau$ denote the present value of all wages received until period t, evaluated at period t. Observe that the liquidity constraint of the worker can be rewritten as $w_t + W_t \geq 0$, and that for any period t we can rewrite the worker's payoff as $V_0(\mathcal{C}) = \delta^t[W_t + V_t(\mathcal{C})]$.

Now, to properly characterize the equilibrium contract, we start by the related problem of finding the optimal static contract. This is not only a warm up, as its solution also is shown to discipline the equilibrium contract.

A static contract is a sequence $(y_t,w_t)_{t=1}^{\infty}$ such that $y_t = y_s$, $w_t = w_s$ for all $t \in \mathbb{N}$. Clearly, an optimal static contract for the firm must pay the worker no more than the interest on his outside option $(1-\delta)h(y)$. Therefore, if $\mathcal{C}^s = \{(y_t^s,w_t^s)_{t=1}^{\infty},\infty\}$ is an optimal static contract with $y_t^s = y_s$ for all t, y_s must satisfy

$$
\begin{aligned}
y_s \in \quad & \arg\max_{y \in \mathcal{Y}} R(y) \\
\text{s.t} \quad & S(y) \geq 0
\end{aligned}
$$

Let $y_s^* = \max\arg\max_y\{R(y) : S(y) \geq 0\}$, and call y_s^* the optimal static productivity level. We take the maximum of the set of maximizers of the firm rent function not only to drop the possible ambiguity, as also because this exact quantity is shown to discipline the productivity limit for any optimal contract.

Now, in order to solve for the equilibrium contract, we start by the simple observation that

Lemma 1 *For any self-enforceable contract with $y_\infty > 0$, there is a period $t_0 \in \mathbb{N}$ such that the present value of wages received up to period t is strictly positive for all $t > t_0$.*

This result is a consequence of the firm's incentive constraint: the firm cannot commit of providing the worker a continuation value of more than $\frac{y_\infty}{1-\delta} - g(y_\infty)$. The continuation value for the worker, at any point in time t_0, can be separated as the total discounted flow of wages up to a period t plus the discounted continuation value at period t. Since this continuation value is bounded by the firm's own incentive constraint, for a large t the wages received from t onwards yield a very small utility for the worker, and therefore the flow of wages up to t must be big enough to offset the worker's outside option.

Lemma 1 states that for a period sufficiently far in the future, the liquidity constraint of the worker is always slack. In particular, Lemma 1 is useful to deriving the next result:

Lemma 2 *For any profit maximizing contract, the worker must be trained up to, at least, the optimal static productivity level y_s^*.*

The intuition for this result is straightforward. If in some contract the worker's continuation value is always less than his outside option at y_r^*, then for a period t distant enough, the firm's continuation payoff must be less than $R(y_\infty)$. The firm can therefore offer an alternative contract which trains the worker up to y_s^*, and this is a weak Pareto improvement. Since by Lemma 1 the liquidity constraint is slack for a period far enough, the worker's payment up to t can be slightly decreased, generating a strictly higher profit for the firm. Otherwise, if for some period t_0 the worker strictly prefers to stay in the contract rather than to leave with productivity y_s^*, the firm can increase his productivity after t_0 without altering his payment, which again generates higher profits. In other words, the only reason for the firm not to maximize its rent function in the limit is to offer a higher promised continuation value for the worker, in order to extract higher rents.

The next lemma restricts the training length; specifically, training must be completed in a finite amount of time:

Lemma 3 *For any profit maximizing contract, there is a $T_g < \infty$ such that $y_t = y_{T_g}$ for any $t \geq T_g$.*

This Lemma relates to a fundamental trade-off in the contract design. If the firm delays the worker's training, it makes quitting a relatively less attractive option, which allows it to provide a lower discounted flow of wages in the rest of the contract. On the other hand, a faster training increases the total output, allowing a increase in both the worker's payment and the firm's profit. The key insight behind this result is that, since the worker's outside option varies smoothly with his production to the firm, when the firm outstanding stock of knowledge is small enough, the attractiveness of quitting is close to the attractiveness of quitting when the training is finished. Therefore, if for some contract training takes infinitely long, it is possible to give the worker the residual knowledge the firm possesses without altering the present value of his flow of payments, which strictly increases profits.

Observe that profit maximizing contracts are not, in general, unique: observe that any alternative sequence of wages which provide a continuation value of at least $h(y_{T_g})$ for the worker and $g(y_{T_g})$ for the firm after the end of training and does not change the continuation value of the worker at period T_g yield the same payoff for both the firm and the worker. In particular, there are always multiple equilibria if $S(y_{T_g}) > 0$ [1].

[1] Too see this, notice that if $S(y_{T_g}) > 0$, for any period $t > T_g$, the incentive constraint of at least one of the agents must be slack. Assuming for simplicity that the incentive constraint of the firm is slack for periods $t' < t''$, observe that an alternative sequence of wages which increases the wage of period t' by some $\epsilon > 0$ and increases the wage of period t'' by $\delta^{t'-t''}\epsilon$ not only yield the same payoffs for both agents as also does not decrease the worker's continuation value for any period

The next lemma provides the finest characterization of the equilibrium contract for an arbitrary outside option function $h(y)$; the result has a close resemblance to the results in GR.

For tractability concerns, we impose that the relationship is statically efficient for all productivity levels $y \in \mathcal{Y}$:

Assumption 1. For all $y \in \mathcal{Y}$, $S(y) \geq 0$.

This assumption ensures that any contract which, for all periods, offers to the worker a continuation value exactly equal to his outside option, satisfies the firm's incentive constraint. In this case, the firm's incentive constraint is inactive at the solution, and the profit maximizing contract does not depend on the function $g(y)$.

Lemma 4 *Assume Assumption 1. In any profit maximizing contract, the worker receives no wage before the period immediately before the graduation date; that is $w_t = 0$ for all $t < T_g - 1$. Also, his outside option grows exponentially at rate r: $h(y_{t+1}) = \delta h(y_t)$.*

This results has an intuitive proof, which can be sketched as following. First, observe that if the worker's continuation value is equal to or higher than his outside option at the end of training $V_t(\mathcal{C}) \geq h(y_\infty)$ for a period t before the graduation date T_g, the firm can redesign the contract in order to anticipate the end of training by paying the worker wage equal to the interest on his continuation value $(1-\delta)V_t(\mathcal{C})$. This new contract must also be feasible, since it maintains the present value of wages and does not involve any negative payments, which could violate the liquidity constraint. Also, since it speeds up the training, this contract is strictly more profitable than the original one.

For intuition purposes, let's assume that time is perfectly divisible. Consider any self-enforceable contract $\mathcal{C}$ which features positive wages before T_g. From this, we can construct an alternative contract which pays nothing before some date T_0, and pays exactly the interest on the worker's continuation value at the graduation date $(1-\delta)V_{T_g}(\mathcal{C})$ after that. We can choose a period T_0 which maintains the exact same present value of the stream of wages at period 0. This makes the contract more back-loaded, increasing the continuation value for the worker for any period before the graduation, enabling the firm to increase the worker's productivity for at least some periods. This speeds up the training and increases total discounted production, which translates into higher profits since the worker's discounted wage flow remains constant.

Even though the explanation above assumes continuous time,we can pin down the pre-graduation wage to 0 without any assumption over δ. If we assume the concavity of $h(y)$, we can further pin down the value of the stream of wages after graduation.

Lemma 5 *Assume assumption 1, and that the function $h(y)$ is strictly concave. In any profit maximizing contract the firm pays no wage to the worker before the end of training, and in the graduation period the continuation value for the worker is exactly $h(y_\infty)$.*

Notice that assuming the concavity of $h(y)$ is not a decreasing returns on knowledge assumption, because y is not properly the knowledge of the worker as much as it is the total production of the worker, if working for the firm, for a given knowledge level. Therefore, assuming $h(y)$ is concave is equivalent to assume that the worker experiences more dramatic decreasing returns for his knowledge when working alone than when working for the firm. This assumption is made primarily for tractability concerns, but it is also economically meaningful. One can interpret such assumption as that the firm must first give general knowledge to the worker, applicable to other firm's in the same industry, and that after such initial training takes place, the additional productivity gains come from specific knowledge acquirement.

In this case, in the equilibrium, the firm pays no wage to the worker until his graduation date, and pays exactly the interest on his outside option after the training ends.

As a consequence of Lemma 5, the equilibrium contract can be found by solving the following reduced form problem:

$$
\begin{aligned}
[PF] \quad & \max_{T_g, y_\infty} \quad \sum_{t=1}^{T_g} \delta^t h^{-1}(\delta^{T_g - t} h(y_\infty)) + \delta^{T_g} R(y_\infty) \\
& \text{s. t} \quad\quad y_\infty \in [y_s^*, y^{\max}] \\
& \quad\quad\quad\quad T_g \in \mathbb{N}
\end{aligned}
$$

Since wages are zero up to the graduation date T_g, are exactly equal to the interest on the worker's outside option at y_∞ after that, and the rate of training is determined by the graduation date and the worker's continuation value after that, we can reduce the problem of finding the profit maximizing contract to the problem of finding the optimal graduation date T_g and productivity level after training y_∞. Without any more structure on $h(y)$, however, there is no way of deriving a closed form solution. In the next section, we solve explicitly the problem above for a isoelastic outside option function $h(y)$.

6.1 Particular case of a Cobb-Douglas production function

In this section, we solve for the equilibrium contract for a particular class of functions $h(y)$. A natural class of functions which provides a nice analytical tractability is the Cobb-Douglas production function, in which the outside option function takes the following form $h^f(y) = ay^\alpha$, for $\alpha \in (0,1), a > 0$. For convenience, normalize $y^{\max} = 1$, so that we have $\mathcal{Y} = [0,1]$. The parameter α is of particular interest, since it provides a measure of how steeper are the returns from knowledge for the worker when using his alternative technology instead of the firm's. We will call α the worker's output elasticity (with respect to his production when working for the firm).

For convenience, let $A = \frac{1}{1-\delta}$, and observe that $h(y) = Ay^{\alpha}$. The parameter a denotes the fraction of the relationship production that the worker can produce outside of the firm. We will call a the productivity coefficient of the agent.

Observe that this class of functions satisfies all the assumptions from Lemma 5. Therefore, we know that, so long as it obeys $S(y) \geq 0$, the function $g(y)$ does not influence in any way the solution, so we omit from specifying it. Also, notice that $y_s^* = 1$, so we know that in any profit maximizing contract, the worker is fully trained (ie, $y_\infty = 1$).

Let $\sigma = 1/\alpha \in (1{,}\infty)$, and observe $h^{-1}(v) = A^{-\sigma} v^{\sigma}$. To ease notation, $R(1) = R = \frac{1}{1-\delta} - A$. We can rewrite problem $[PF]$ as

$$[P1] \quad \max_{T \in \mathbb{N}} \Pi(T) = \delta^T \left[\delta^{(\sigma-1)T} \sum_{t=1}^{T-1} \delta^{-(\sigma-1)t} y_\infty + R \right]$$

The expression inside the brackets denote exactly the firm's profit evaluated at period T monetary units. Observe that we can simplify the objective function to

$$\Pi(T) = \delta^T \left[\frac{\delta^{-(\sigma-1)}}{\delta^{-(\sigma-1)} - 1} (1 - \delta^{(\sigma-1)T}) + R \right]$$

Let $c = \delta^{-(\sigma-1)} > 1$. Then,

$$\Pi(T) = \frac{1}{c-1} \delta^T \left[(1 - c^{-(T-1)}) + (c-1)R \right]$$

Taking the derivative with respect to T,

$$\frac{\partial \Pi(T)}{\partial T} = \log(\delta) \left[\Pi(T) - \delta^T (\sigma - 1) \frac{1}{c-1} c^{-(T-1)} \right]$$

Rearranging,

$$\frac{\partial \Pi(T)}{\partial T} = \log(\delta) \frac{1}{c-1} \delta^T \left[(1 - \sigma \delta^{(\sigma-1)(T-1)}) + (c-1)R \right]$$

Observe that the function $\Pi(T)$ has the nice feature that there is only one T^* such that $\Pi'(T^*) = 0$, and furthermore, we have $\Pi'(T) > 0$ for $T < T^*$ and $\Pi'(T) < 0$ for $T < T^*$. Therefore, the solution for problem $[P1]$ must be either ceil(T^*) or floor(T^*). We can therefore state the exact

solution for this particular case:

Proposition 4 *Assume* $h(y) = \frac{a}{1-\delta} y^{\alpha}$*, with* $\alpha \in (0,1)$*, and* $\mathcal{Y} = [0,1]$*. If* $g(y)$ *satisfies* $S(y) \geq 0$ *for all* $y \in [0,1]$*, in any profit maximizing contract, the worker receives no wage before the graduation date* T_g*, and at* T_g *offers a continuation value of* $h(1)$ *for the worker. Lastly, if* $\mathcal{T} \geq 1$*,* T_g *satisfies* $T_g = ceil(\mathcal{T})$ *or* $T_g = floor(\mathcal{T})$*, with* $\mathcal{T}$ *defined in the expression below.*

$$\mathcal{T} = \frac{1}{\log(\delta)} \frac{\alpha}{1-\alpha} \left[\log\left(1 + \frac{1-a}{1-\delta}\left(\delta^{-\frac{(1-\alpha)}{\alpha}} - 1\right)\right) + \log(\alpha) \right] + 1$$

This expression for T_g enables some comparative statics analysis, assuming $\mathcal{T} \geq 1$. First, a change in discount rate affects the graduation date by two different channels. An increase in the discount rate raises the magnitude of $R(y_\infty)$, leading to higher profits it $R(y_\infty) > 0$ and higher losses if $R(y_\infty) < 0$. This imply a early graduation is more and less attractive to the firm, respectively. However, since the worker's outside option grows at rate $r = -log(\delta)$, a higher δ means a lower growth rate. The overall effect is ambiguous in general; for an elasticity α close enough to 1, nevertheless, the first effect dominates, and higher δ leads to earlier and later graduation dates for $a < 1$ and $a > 1$, respectively [2].

On the other hand, if $R > 0$, $\frac{\partial T_g}{\partial \alpha} > 0$. The intuition for this result is that, if $\alpha' < \alpha$, we have $y^{\alpha'} > y^{\alpha}$ for all $y \in [0,1]$. The worker's production function, for lower values of α, features not only steeper decreasing returns, as also increases the worker's outside option for any level of productivity. This reduces the rent extraction margin for the firm, and makes longer contracts less attractive. Lastly, $\frac{\partial T^*}{\partial a} > 0$, since a higher productivity coefficient diminishes profits at the end of training productivity more than for intermediary productivity levels.

7 Conclusion

In this paper, we have shown that allowing the worker and the firm to have an arbitrary monotonous outside option function in the environment presented in Garicano and Rayo (2017) yields novel insights on the nature of apprenticeships relations. In this environment, organizational efficiency naturally restricts the possibility of contracts, and generates a puzzle regarding temporary apprenticeships relations: in the last period before termination, the relationship must be statically efficient, implying therefore that there is no incentive to break it.

We also showed that the main results from Garicano and Rayo (2017) that the structure of apprenticeships can be rationalized as a monopsonic firm designing a relational contract in order

[2] This is because $\frac{\partial}{\partial \delta}\left(\frac{\delta^{-\frac{(1-\alpha)}{\alpha}}-1}{1-\delta}\right) = \frac{1}{1-\delta}\left[\left(\delta^{-\frac{(1-\alpha)}{\alpha}} - 1\right)\frac{1}{1-\delta} - \frac{1-\alpha}{\alpha}\delta^{-\frac{(1-\alpha)}{\alpha}-1}\right]$. By a simple application of L'Hôpital's rule, we have that the second term dominates the first as $\alpha \to 1$.

to maximize profits in an environment without enforceability and with credit market imperfections hold in a more general setting. In equilibrium, the training dynamics is such that the outside option of the worker grows exponentially at the discount rate. This is a simple, but important insight: the relevant productivity for determining the speed of training is of the one associated with the outside option of the worker, instead of the productivity of the worker when working for the firm. This means that for more specific the knowledge is to the firm itself, the faster the training is.

In the case of a Cobb-Douglas production function for the worker, we were able to solve explicitly for the equilibrium contract. Under the assumption that when fully trained the worker is more productive at the firm than at an alternative employment, we show that the training duration is longer when i. the elasticity of substitution of the worker's production function is higher, since this decreases the worker's outside option for intermediary productivity levels and ii. the higher is the productivity coefficient, since this decreases the firm's profits after the end of training. The effect of the discount rate is ambiguous, since a higher δ changes both the magnitude of the profits or losses after the worker is fully trained and changes the growth rate of the worker's productivity.

Regarding future work, this paper raises some natural questions which may lead to new insights. First, relaxing the assumption that at the maximum productivity level the rent function is positive necessarily leads to an equilibrium qualitatively different from the one studied in this paper. And second, the relationship between the function $h(y)$, the productivity of the worker inside the firm and the composition of the training between general and firm specific skills could be further explored.

8 References

Gary S Becker. Investment in human capital: A theoretical analysis. *Journal of political economy*, 70(5, Part 2):9–49, 1962.

Luis Garicano and Luis Rayo. Relational knowledge transfers. *American Economic Review*, 107 (9):2695–2730, 2017.

David De la Croix, Matthias Doepke, and Joel Mokyr. Clans, guilds, and markets: Apprenticeship institutions and growth in the preindustrial economy. *The Quarterly Journal of Economics*, 133 (1):1–70, 2018.

Joel Mokyr. The economics of apprenticeship. *Apprenticeship in Early Modern Europe*, pages 20–43, 2019.

Jonathan Levin. Relational incentive contracts. *American Economic Review*, 93(3):835–857, 2003.

W Bentley MacLeod and James M Malcomson. Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica: Journal of the Econometric Society*, pages 447–480, 1989.

Drew Fudenberg and Luis Rayo. Training and effort dynamics in apprenticeship. *American Economic Review*, 109(11):3780–3812, 2019.

Johannes Hörner and Andrzej Skrzypacz. Selling information. *Journal of Political Economy*, 124 (6):1515–1562, 2016.

Wei Zhao, Claudio Mezzetti, Ludovic Renou, and Tristan Tomala. Contracting over persistent information. *arXiv preprint arXiv:2007.05983*, 2020.

A Mathematical Appendix

Proof of Proposition 1.

This results follows simply by summing both players' incentive condition at the last period of the contract

$$y_T + \delta[g(y_T) + h(y_T)] = V_T(\mathcal{C}) + \Pi_T(\mathcal{C}) \geq g(y_T) + h(y_T)$$

, which implies $S(y_T) \geq 0$.

Proof of Proposition 2.

Since (y_t) is increasing and bounded, it converges to y^*. Now, summing $[FIC_{t-1}]$ and $[WIC_t]$, we have

$$\sum_{\tau=t}^{\infty} \delta^{\tau-t} y_\tau \geq g(y_t) + h(y_{t-1})$$

Taking the limit as $t \to \infty$ and using the fact that g,h are continuous, we get $\frac{1}{1-\delta} y^* \geq g(y^*) + h(y^*)$, which implies $S(y^*) \geq 0$.

Proof of Lemma 1.

If $\mathcal{C}$ features $y_\infty > 0$, there is $t_0 \in \mathbb{N}$ such that $y_{t_0} > 0$. Let $W_t^{t_0} = \sum_{\tau=t_0}^{t-1} \delta^{\tau-t} w_\tau$, and observe that $W_t \geq W_t^{t_0}$. For all $t > t_0$, the agent's incentive constraint at t_0 can be rewritten $V_{t_0}(\mathcal{C}) = \delta^{t-t_0}[W_t^{t_0} + V_t(\mathcal{C})] \geq h(y_{t_0})$. By the firm's incentive constraint, $V_t(\mathcal{C}) \leq \frac{y_\infty}{1-\delta} - g(y_\infty)$. Rearranging and using the liquidity constraint for period t_0,

$$\begin{aligned} W_t^{t_0} &\geq \delta^{t_0} \left[\delta^{-t} h(y_{t_0}) - \delta^{t-t_0} V_{t+1}(\mathcal{C})\right] \\ &\geq \delta^{t_0} \left[\delta^{-t} h(y_{t_0}) - \delta^{t-t_0} \left(\frac{y_\infty}{1-\delta} - g(y_\infty)\right)\right] \end{aligned}$$

Observe that the right hand side goes to ∞ as $t \to \infty$, so we conclude $\lim_{t\to\infty} W_t = +\infty$.

Proof of Lemma 2.

Suppose $\mathcal{C}$ is a profit maximizing contract with $0 < y_\infty < y_s^*$. We will divide this proof into two cases.

Case 1. If there is t_0 such that $V_{t_0}(\mathcal{C}) \geq h(y_s^*)$, consider the alternative contract $\mathcal{C}'$ identical to $\mathcal{C}$ up to t_0 and with $y_t' = y_s^*$ and $w_{t_0}' = (1-\delta)V_t(\mathcal{C})$ for all $t \geq t_0$. Observe that $\mathcal{C}'$ maintains the same continuation value for the agent at period t_0, and since it is identical to $\mathcal{C}$ up to t_0, the agent's incentive constraint is satisfied for all $t \leq t_0$. Also, for $t > t_0$, observe that the contract $\mathcal{C}'$ yields the agent a continuation value strictly higher than its outside option: $V_t(\mathcal{C}') = V_{t_0}(\mathcal{C}) \geq h(y_s^*) = h(y_t')$.

The liquidity constraint is satisfied for $t \geq t_0$ since $w'_t > 0$ for all $t \geq t_0$, and $W'_t = W_t \geq 0$ for $t < t_0$.

This contract yields a strictly higher continuation value for the firm at t_0, since $\Pi_{t_0}(\mathcal{C}) = \frac{1}{1-\delta} y_s^* - V_{t_0}(\mathcal{C}') > \frac{1}{1-\delta} y_\infty - V_{t_0}(\mathcal{C}) \geq \Pi_{t_0}(\mathcal{C})$. This implies that, for any $t < t_0$,

$$\begin{aligned} \Pi_t(\mathcal{C}') &= \sum_{\tau=t}^{t_0-1} \delta^{\tau-t} \left[y'_\tau - w'_\tau \right] + \delta^{t-t_0} \Pi_{t_0}(\mathcal{C}') \\ &> \sum_{\tau=t}^{t_0-1} \delta^{\tau-t} \left[y_\tau - w_\tau \right] + \delta^{t-t_0} \Pi_{t_0}(\mathcal{C}) \\ &= \Pi_t(\mathcal{C}) \end{aligned}$$

, so not only the firm's incentive constraint is satisfied for all $t < t_0$ as the contract $\mathcal{C}'$ is strictly more profitable than $\mathcal{C}$. Lastly, for $t \geq t_0$, $\Pi_t(\mathcal{C}) = \frac{1}{1-\delta} y_s^* - h(y_s^*) \geq g(y_s^*)$ using the fact that $(y_s^*) \geq 0$.

Case 2. If for all t the agent's continuation value is smaller than his outside option at y_s^* $V_t(\mathcal{C}) < h(y_s^*)$, then for any $\epsilon > 0$, there is t_0 big enough such that $\frac{y_\infty}{1-\delta} - V_t(\mathcal{C}) - \epsilon < R(y_s^*)$ for all $t > t_0$. In particular, this is true for $\epsilon = \delta(h(y_s^*) - h(y_\infty))$. Without loss of generality, assume that $W_{t_0} + w_{t_0} > \epsilon$. Consider an alternative contract $\mathcal{C}'$ identical to $\mathcal{C}$ up to $t_0 - 1$, and with $w'_{t_0} = w_{t_0} - \epsilon$, $y'_{t_0} = y_{t_0}$, and $w'_t = (1-\delta) h(y_s^*)$, $y'_t = y_s^*$ for all $t \geq t_0$. By construction, this contract satisfies the liquidity constrain. Also, it yields a higher continuation value at t_0 for the agent:

$$\begin{aligned} V_{t_0}(\mathcal{C}') - V_{t_0}(\mathcal{C}) &= \delta \left[h(y_s^*) - V_{t_0+1}(\mathcal{C}) \right] - \epsilon \\ &\geq \delta \left[h(y_s^*) - h(y_\infty) \right] - \epsilon = 0 \end{aligned}$$

, and therefore the agent's incentive constraints are satisfied up to t_0. For $t > t_0$, they are also satisfied since the wage is exactly the interest on the agent's outside option.

Observe that this contract yields a strictly higher continuation value for the firm at period t_0:

$$\begin{aligned} \Pi_{t_0}(\mathcal{C}') - \Pi_{t_0}(\mathcal{C}) &= \epsilon + \delta \left[R(y_s^*) - \Pi_{t_0+1}(\mathcal{C}) \right] \\ &\geq \epsilon + \delta \left[R(y_s^*) - R(y_\infty) \right] \\ &= \frac{\delta}{1-\delta} \left[y_s^* - y_\infty \right] > 0 \end{aligned}$$

, which implies $\Pi_t(\mathcal{C}') - \Pi_t(\mathcal{C}) = \delta^{t_0-t} [\Pi_{t_0}(\mathcal{C}') - \Pi_{t_0}(\mathcal{C})] > 0$ for any $t < t_0$; in particular, evaluating at $t = 0$ we conclude that contract $\mathcal{C}'$ is strictly more profitable. Lastly, the firm's incentive constraint is satisfied for $t > t_0$ since $\Pi_t(\mathcal{C}') = R(y_s^*) \geq g(y_s^*)$ by the definition of y_s^*.

Proof of Lemma 3.

Assume that $\mathcal{C}$ is a profit maximizing contract with $T_g = \infty$. Since $y_t \to y_\infty$, by the continuity of $h(.)$ we have $h(y_t) - h(y_\infty) \to 0$. This implies we can choose k large enough such that $h(y_k) \geq \delta h(y_\infty)$. Consider the alternative $\mathcal{C}'$ identical to $\mathcal{C}$ up to period $k-1$ and with $w'_k = V_k(\mathcal{C}) - \delta h(y_\infty)$, $y'_k = y_k$, $y_t = y_\infty$, $w_t = h(y_\infty)$ for all $t > k$. The wage w'_k is necessarily non-negative, since that by $V_k(\mathcal{C}) \geq h(y_k)$, we have $w'_k = V_k(\mathcal{C}) - \delta h(y_\infty) \geq h(y_k) - \delta h(y_\infty) \geq 0$.

We claim that $\mathcal{C}'$ is a feasible contract:

For period k, the wage w'_k does not change the agent's continuation payoff: $V_k(\mathcal{C}') = w'_k + \delta h(y_\infty) = V_k(\mathcal{C}) \geq h(y_k) = h(y'_k)$. Since $V_k(\mathcal{C}') = V_k(\mathcal{C})$ and $\mathcal{C}'$ is identical to $\mathcal{C}$ up to period k, we have $V_t(\mathcal{C}') = V_t(\mathcal{C}) \geq h(y_t) = h(y'_t)$, so the agent incentive constraint is satisfied up to period k. Lastly, for $t > k$, we have $V(\mathcal{C}') = h(y_\infty) = h(y'_t)$. Similarly, the firm's profit in period k is higher for $\mathcal{C}'$, since

$$\begin{aligned}
\Pi_k(\mathcal{C}') = \;& y_k - w_k + \delta\left[\frac{1}{1-\delta} y_\infty - h(y_\infty)\right] \\
&= y_k + \delta \frac{1}{1-\delta} y_\infty - V_k(\mathcal{C}) \\
&> \sum_{\tau=k}^{\infty} \delta^{\tau-k} y_\tau - V_k(\mathcal{C}) \\
&= \Pi_k(\mathcal{C})
\end{aligned}$$

For any $t < k$, $\Pi_t(\mathcal{C}') - \Pi_t(\mathcal{C}) = \delta^{k-t}\left[\Pi_k(\mathcal{C}') - \Pi_k(\mathcal{C})\right] > 0$; in particular, using $t = 0$ we conclude that $\mathcal{C}'$ is strictly more profitable than $\mathcal{C}$. Since $y'_t = y_t$ for $t \leq k$, the firm's incentive constraint is satisfied up to k. And lastly, for $t > k$, we have $\Pi_t(\mathcal{C}') = \frac{1}{1-\delta} y_\infty - h(y_\infty) \geq g(y_\infty)$ using Proposition 2.

Proof of Lemma 4.

We say that $\mathcal{C}$ is a quasi-delayed rewards contract if $w_t = 0$ for all $t < T_g - 1$. Let $\mathcal{Q}$ be the set of all the contracts with this property.

Step 1. For any profit maximizing contract, $V_t < h(y_\infty)$ for all $t < T_g$, and $V_{T_g} < \delta^{-1} h(y_\infty)$.

Proof. If $\mathcal{C}$ features $V_{t_0} \geq h(y_\infty)$ for some $t_0 < T_g$, consider an alternative contract $\mathcal{C}'$ identical to $\mathcal{C}$ up to period t_0 and with $w'_t = (1-\delta)V_{t_0}$, $y'_t = y_\infty$ for all $t \geq t_0$. Observe that $\mathcal{C}'$ offers the same continuation value than $\mathcal{C}$ at t_0, and clearly satisfies the agent's incentive constraint. Since $\mathcal{C}'$ is identical to $\mathcal{C}$ up to t_0, the agent's total utility is unchanged $V_0(\mathcal{C}') = V_{t_0}(\mathcal{C})$. On the other hand, since $t_0 < T_g$, $y'_{t_0} > y_{t_0}$, and therefore $\Pi_{t_0}(\mathcal{C}') > \Pi_{t_0}(\mathcal{C})$. Since the contract is the same up to t_0, the firm's profit is strictly higher for $\mathcal{C}'$, which completes our proof.

Step 2. If $\mathcal{C}$ is an optimal contract, $\mathcal{C} \in \mathcal{Q}$.

Suppose $\mathcal{C} \notin \mathcal{Q}$ is a profit maximizing contract. By Lemma 3, the graduation date is finite:

$T_g < \infty$. Define $\mathcal{C}'$ as following:

Let S be the smallest period t such that $\delta^t V_{T_g} \leq V_0$ (observe that $S \leq T_g$). Let $w'_t = 0$ for all $t < S$, $w'_t = (1-\delta)V_{T_g}$ for $t \geq S$ and $w'_{S-1} = \delta^{-(S-1)}V_0 - \delta V_{T_g}$. Observe that, since S is defined such that $\delta^S V_{T_g} \leq V_0$, we have $w'_S \geq 0$; and since $\delta^{S-1}V_{T_g} > V_0$, the wage at period $S-1$ can be no higher than $(1-\delta)V_{T_g}$; otherwise, we would have

$$(1-\delta)V_{T_g} \leq \delta^{-(S-1)}V_0 - \delta V_{T_g}$$

, which implies $\delta^{S-1}V_{T_g} \leq V_0$. This is impossible since S is the smallest period with this property.

This contract delivers the same utility for the agent at time $t = 0$:

$$V_0(\mathcal{C}') = \delta^{S-1}(w'_{S-1} + \delta V_S) = V_0 + \delta^S(V_{T_g} - V_{T_g}) = V_0(\mathcal{C})$$

Lastly, for all $t \in \mathbb{N}$ define y'_t such that $\min\{V'_t, h(y_\infty)\} = h(y'_t)$ this is possible since V'_t is increasing in t. Observe that by construction the agent's incentive constraint is satisfied for all t.

Since the present value of wages received until period t $\sum_{\tau=0}^{t-1} \delta^\tau w'_\tau$ is 0 for all $t \leq S-1$, we have

$$V_t(\mathcal{C}') = \sum_{\tau=t}^{\infty} \delta^{\tau-t} w'_t = \sum_{\tau=0}^{t-1} \delta^{\tau-t} w_t + \sum_{\tau=t}^{\infty} \delta^{\tau-t} w_t \geq V_t(\mathcal{C})$$

for $t < S-1$; and for $t \geq S$, we have

$$V_t(\mathcal{C}') = V_{T_g}(\mathcal{C}) \geq h(y_\infty) = h(y_\infty) \tag{1}$$

Since the utility of the agent under the alternative contract is higher than for the original $V_t(\mathcal{C}') \geq V_t(\mathcal{C})$ for all periods $t < T_g$ and using the fact that $h(y'_t) = V'_t$ if $V'_t < h(y_\infty)$, we have $y'_t \geq y_t$ for $t < T_g$.

If $\mathcal{C} \notin \mathcal{Q}$,then the inequality $V'_t \geq V_t$ is strict for some $t < T_g$ by (1). By Step 1, $V_t < h(y_\infty)$, and therefore $h(y'_t) = \min\{V'_t, h(y_\infty)\} > V_t \geq h(y_t)$, which implies $h(y'_t) > h(y_t)$ and therefore $y'_t > y_t$. We conclude that $\mathcal{C}'$ is strictly more profitable, since, using $V_0(\mathcal{C}) = V_0(\mathcal{C}')$,

$$\Pi_0(\mathcal{C}') - \Pi_0(\mathcal{C}) = \sum_{t=0}^{\infty} \delta^t [y'_t - y_t] > 0$$

We still have to show that $\mathcal{C}'$ satisfies the firm's incentive constraint. For this, observe that, for all t such that $V'_t < h(y'_t)$, we have $V'_t = h(y'_t)$ and therefore, using assumption 3, we know that

$S(y'_t) \geq 0$, and therefore

$$\Pi_t(\mathcal{C}') = \sum_{\tau=t}^{\infty} \delta^{\tau-t} y'_t - V'_t \geq \frac{1}{1-\delta} y'_t - h(y'_t) \geq g(y'_t)$$

, and if $V'_t \geq h(y_\infty)$,

$$\Pi_t(\mathcal{C}') = \frac{y_\infty}{1-\delta} - V'_t \geq \frac{y_\infty}{1-\delta} - V_{T_g} \geq g(y_\infty)$$

, which ends our proof.

Proof of Lemma 5.

By Lemma 4, any profit maximizing contract is in $\mathcal{Q}$. Therefore, any profit maximizing contract can be found by solving the following problem:

$$\begin{aligned}
\max_{\mathcal{C}\in\mathcal{Q}} \Pi_0(\mathcal{C}) = \quad & \max_{T_g, V, y_\infty} & \sum_{t=1}^{T_g-1} \delta^t h^{-1}(\delta^{T_g-1-t} V) + \delta^{T_g-1}\left(\delta \frac{1}{1-\delta} y_\infty - V\right) \\
& \text{s.t} & V \in [\delta h(y_\infty), h(y_\infty)) \\
& & V \leq R_a(y_\infty) \\
& & y_\infty \in [y_s^*, y^{\max}] \\
& & S(y_\infty) \geq 0 \\
& & T_g \in \mathbb{N}
\end{aligned}$$

Since the function $h(y)$ is strictly increasing, if $h(y)$ is strictly concave, its inverse $h^{-1}(v)$ is strictly convex. The objective function above, therefore, is strictly convex in V. This reduced formed firm's problem has an objective function which is strictly convex in V, and therefore the solution must be a corner solution. Observe that the restriction $V \leq R_a(y_\infty)$ is redundant, since using $S(y_\infty) \geq 0$ we have $R_a(y_\infty) \geq h(y_\infty)$, and therefore $V < h(y_\infty)$ implies $V \leq R_a(y_\infty)$. Therefore, since V must have a corner solution, we must have $V = \delta h(y_\infty)$, which implies $V_{T_g-1}(\mathcal{C}) = \delta h(y_\infty)$. This yields $w_{T_g-1} = 0$ by the liquidity constraint, and we conclude that $V_{T_g}(\mathcal{C}*) = h(y_\infty)$.