Optimal Chief Executive Officer Pay in Winner-take-all Settings*

David F. Larcker
Stanford University

Venky Nagar
University of Michigan

Abstract

Many modern firms compete in a "winner-take-all" setting where all firms appear to be equals at first but then a few firms quickly race ahead and dominate the space at the expense of everyone else. The CEOs of such firms also become extremely wealthy, raising the question of whether they are overpaid. We build a Bayes rational dynamic winner-take-all model where the CEO pay structure is ex ante optimal but ex post rewards some of the surviving winners excessively. Our results suggest that excess pay arguments based only on ex post winners suffer from survivorship bias; one should also consider CEOs who failed and destroyed value.

Keywords: Adverse selection; Chief executive officer compensation

*We thank Sugato Bhattacharyya, Rick Lambert, Kevin Smith, and seminar participants at Michigan and Purdue for their comments, and ECGP, the Hoover Institution, and the Rock Center for their support.

Optimal Chief Executive Officer Pay in Winner-take-all Settings

Abstract

Many modern firms compete in a "winner-take-all" setting where all firms appear to be equals at first but then a few firms quickly race ahead and dominate the space at the expense of everyone else. The CEOs of such firms also become extremely wealthy, raising the question of whether they are overpaid. We build a Bayes rational dynamic winner-take-all model where the CEO pay structure is ex ante optimal but ex post rewards some of the surviving winners excessively. Our results suggest that excess pay arguments based only on ex post winners suffer from survivorship bias; one should also consider CEOs who failed and destroyed value.

1 Introduction

Most of the financial wealth in modern capital markets is created by a few firms. Bessembinder et al. (2023) find that the top-performing 2.4 percent of firms account for all of the $US 75.7 trillion in net global stock market wealth creation from 1990 to December 2020. One potential reason is that many modern firms operate in a winner-take-all setting where all firms' products initially compete on a relatively even footing but then suddenly a confluence of external events and internal pivots—e.g., technological developments, production line innovations, and product redesign—triggers a race that only a few firms win. These firms then dominate their product space and their CEOs become extremely wealthy.[1]

The extreme CEO pay for winners has generated the criticism that these CEOs are overpaid (Bebchuk and Fried, 2003). In response, Gabaix and Landier (2008) build a rational model of an economy of firms with a few large winners and many small losers, all of whom search for CEOs from a pool of available managers whose talent is known to everyone. Top firms select top talent and pay them far more. The extreme skewness in firm size thus carries over to extreme skewness in pay for CEOs relative to their talent. But this pattern of extreme pay for some CEOs is deserved because it arises from a competitive equilibrium where every manager is paid her marginal product at the firm she matches. However, the pay in the model is lump-sum because each manager's talent is known to all. As a result, the model cannot directly address the argument that extreme CEO pay is deserved primarily because it is not a fixed lump-sum but instead performance-based where the CEO only gets paid when the firm wins (Larcker and Tayan, 2023). Accordingly, the SEC pay versus performance (PVP)

[1]Our Appendix A shows some examples. Tesla and BYD dominate the EV space, while Nvidia GPUs dominate the AI space. The key events leading to this domination were likely both internal pivots and external developments, i.e., a combination of supply chain and factory floor advances enabled Tesla to suddenly unlock its high-volume production bottlenecks, or Google's discovery of the transformer neural net architecture and its amenability to Nvidia's CUDA technology suddenly made LLM AI models feasible and immediately increased the demand for Nvidia chips. These events were also unpredictable as our Appendix A shows—the meteoric rise of these stocks over a short period relative to the losers implies that most investors had no ex ante expectations of future success. Of course, Appendix A is not the end of the story for all these firms which is why we also consider dynamic multiperiod races.

rule enacted in 2022 requires the disclosure of both the ex ante compensation granted and the ex post compensation actually paid.[2] The present study therefore takes a middle ground and builds a model of performance-based pay that is ex ante optimal but ex post excessive for the winners.

Gabaix and Landier (2008) model the hiring situation after firms have won or lost the winner-take-all race. By contrast, we argue that boards search for CEOs before the race begins, and that these CEOs are crucial to the race outcome of a few winners and many losers. The occurrence of few big and many small firms is then not just happenstance but instead two sides of the same coin: the very race that creates the winners also creates the losers. In our model, these races play over multiple times and the board has to hire and compensate a CEO who can maximize the firm's residual payoffs in these races. The performance-based pay schemes for both the winning CEO and the losing CEO are then also simultaneously determined before the race, not after.[3]

We first model the winner-take-all nature of the race by assuming that the cash flow gap between winning and losing is so large and the chances of winning so uncertain that the board cannot cross-insure across these two states and so must pay out of its actual cash flows.[4] We thus effectively disallow the conglomerate setting where the corporate parent routinely cross-insures its subsidiaries with substantive cash transfers. Instead, our setting is more akin to a tech venture capitalist portfolio where each investee CEO's fortunes depend on her firm's own performance, and the venture capital (VC) firm makes most of its income off a few extreme winners in its portfolio. We also study a Bayes rational dynamic model of a sequence of two identical races.

[2] `https://www.sec.gov/files/rules/final/2022/34-95607.pdf`

[3] For example, Axon CEO Rick Smith recently said, "One of the things I love about the design is I can make zero—I feel better about making a lot if I can also make zero," `https://www.wsj.com/business/rick-smith-axon-ceo-pay-package-2024-6e864a64?mod=mhp`

[4] For example, in Appendix A, it may not be ex ante feasible to insure across winners versus losers like Tesla versus GM or Nvidia versus Intel. One reason could be that the board and the insurer may not have common priors on the chances of winning and so may not be able to settle on a price. Although the common priors assumption is ubiquitous in the literature, Kreps (1990, p.111) calls this assumption a theoretical economists' "dogma" that he cannot defend. See Harrison and Kreps (1978) for a model without the common priors assumption.

We next model the pre-race CEO hiring and compensation situation as follows. The board selects a CEO finalist whose type it cannot assess perfectly either during the initial interview. The CEO type reflects the ability of the CEO to quickly pivot the firm to success once the race begins.[5] We assume that the finalist CEO candidate has sufficient introspection and knowledge of the firm to know her type, but cannot credibly communicate it to the board. We also assume that the board's prior beliefs about the CEO type is common knowledge. For simplicity, we assume two possible CEO types, high or low. The board makes a one-time take-it-or-leave-it offer that pins down the pay if the firm loses the race and the pay if it wins.[6]

We next discuss our results for a single race and then extend the intuition to two consecutive races. The key factor driving all the results is that there is not enough money to pay the CEO his reservation wage when the firm loses the winner-take-all race; the firm must therefore make it up for the CEO when he wins. In the single race equilibrium, the low type CEO knows that he has a lower chance of winning the race than the high type and so needs more compensation if he wins to meet his reservation wage. The board may find it optimal to pay him this amount because having a low type CEO may still be better than forgoing the project and making zero. The high type CEO can then free ride off the low type's demand for more compensation upon winning and earn excess pay. Since the high type CEO is also more likely to win, the adverse selection inefficiency empirically manifests itself as survivorship bias: pay appears excessive if one only looks at ex post winners, but rational if one looks at all ex ante participants.

Turning to the dynamic scenario of two consecutive races, we gain several additional insights.[7] The key reason the CEO is overpaid for winning in the one-period setting is because the firm cannot insure across the winning and losing outcomes and so cannot pay the

[5]Indeed, a recently published academic survey shows that finding the right type of CEO is directors' top priority that easily outranks other issues such as designing an appropriate compensation package to motivate CEO efforts (Edmans et al., 2023, Table 2).

[6]The one-time nature of the offer is supposed to capture the idea that the board cannot draw out the search process. See Section 3.

[7]We use rationality and Bayes rule to update everyone's beliefs after the first race.

CEO enough if he loses (see footnote 4). This is not true in a two-period model if the board can commit to two-period contracts. The firm can promise the CEO the reservation wage no matter the CEO type or the outcome in the first period (thus removing any premiums) and then, if the firm loses the first race, it can promise to pay him any unpaid arrears (appropriately amortized) if he wins the second race. This promise meets all finalists' ex ante expectations of earning reservation wages and so it is accepted. Thus, there is no premium paid to the CEO in the first race, but there is substantive premium paid in the second race for winning.

If the above two-period contracts cannot be written, the optimal strategy for the board is to pay the CEO the usual premium as in the one-period model for the first race, but then squeeze any premium out if he loses the first race and possibly also let him go and hire a new CEO. On the other hand, if the incumbent wins the first race, the board continues the premium in the second race.

In sum, what the two-period models show is that the ex ante optimal but ex post excessive premium or rent endures: what changes is the timing of this rent. The idea that the high type agent earns rent and that this rent evolves over time is not new to theorists, who have nonetheless built a vast number of models on this topic (Bolton and Dewatripont, 2005, Ch.9). The reason is that the nature of these rents is highly sensitive to institutional details and the specific intuition in one setting rarely carries over to another.

Turning to our specific winner-take-all setting, the rents we identify are of empirical interest because they manifest themselves as survivorship bias, i.e., ex ante rationally deserved pay may appear ex post excessive if one only looks at surviving firms and not at all the firms who entered the race. The main contribution of our model is then perhaps empirical. We show how empirical analyses could benefit from paying simultaneous and equal attention to

both winners and losers.[8]

Interestingly, the excess pay for the winning CEO obtains when both the high type and the low type CEO in our model receive identical contracts (pooling equilibrium).[9] This finding is thus contrary to the skimming argument that many CEOs who earn excess pay do so through pay-for-performance contracts that unfairly reward poor performance (Bebchuk and Fried, 2003). Instead, what our model suggests is that excess pay is a rational consequence of optimal governance.

The remainder of the paper is as follows. Section 2 provides a brief review of adverse selection models. Section 3 describes our model. Sections 4 and 5 analyze the one-period and two-period equilibria. Section 6 considers a few extensions. Concluding remarks are provided in Section 7.

2 Literature Review

Edmans et al. (2017) provide an insightful taxonomy of theories of CEO pay, namely behavioral and rational. The behavioral theory is straightforward and argues that CEOs systematically have the power to exploit captive or inattentive boards and skim off shareholder wealth as pay (Bebchuk and Fried, 2003).

Turning to rational economic explanations, Gabaix and Landier (2008) examine a setting where the CEOs who are publicly known to be of a high type endogenously match in equi-

[8]As an illustration, compare the two Wall Street Journal articles: 1) Elon Musk's roughly $48.5 billion pay:"Tesla Shareholders Vote to Uphold Musk's Pay Plan" `https://www.wsj.com/business/tesla-musk-pay-package-approved-826f71ccWSJ`; and 2) Intel CEO Pat Gelsinger's abrupt firing:"He Was Going to Save Intel. He Destroyed $150 Billion of Value Instead." `https://www.wsj.com/tech/intel-ceo-pat-gelsinger-chips-bd4d61f9`. The Intel CEO appears to have been more expensive to shareholders than a potentially overpaid Elon Musk. A less extreme example is DoorDash, whose CEO Tony Xu has not yet achieved the stock-price targets necessary earn his 2020 equity award, which was granted before the company's initial public offering and valued at about $414 million at the time. The CEO's 2023 pay, the most recent reported, totaled $316,431. "Tony's equity package is entirely performance-based and maintains ambitious goals for the business," a DoorDash spokesman said. "Until those goals are met, his take-home pay is limited primarily to his salary of $300,000." `https://www.wsj.com/business/ceo-pay-2024-data-bc6ecb11`

[9]In fact, this lack of variation is what causes excess pay for the high type CEO. We explain in Section 4.3 why we do not obtain a separating equilibrium.

librium to high value firms. The distribution of firm value in these models is exogenously heavy-tailed (i.e., a few large firms dominate a multitude of small ones) and high type CEOs' marginal product is greater for larger firms. These CEOs then end up matching to large firms in equilibrium and earning far more than the other CEOs. Empirical calibration of the model matches observed variations in CEO pay level (and firm size).[10]

These models raise the question of how winner-take-all industries with a heavy-tailed distribution of firm values emerge. One possibility that Gabaix and Landier (2008) speculate is technological network effects and economies of scale where a firm's initial advantage over other firms exponentially compounds period after period. In a later study, Gabaix (2016) models this process using the random growth model where each firm grows each period at a random growth rate. This results in a winner-take-all economy where few firms grow exponentially and become very large and dominate everyone else.[11]

The Gabaix and Landier (2008) model focuses only on the level of CEO pay and not its composition because the CEO type is public knowledge and the CEO is thus paid a flat fee that is equal to his marginal product. Yet, the entire practitioner argument for why CEO pay is deserved is precisely because the pay is not fixed – the CEO's performance-based contract makes money only when the firm does well (Larcker and Tayan, 2023). In our view, Gabaix and Landier (2008) model the hiring problem after firms have won or lost the winner-take-all race. By contrast, we argue that boards search for CEOs before the race begins, and that these CEOs are crucial to the race outcomes. The occurrence of few big and many small firms is then not just happenstance but instead two sides of the same coin:

[10]Two other reasons Gabaix and Landier (2008, Section I) offer for high CEO pay are that 1) increasing volatility in the business landscape has necessitated performance-based pay for CEOs, who being risk averse, demand an *ex ante* wage premium; and 2) the CEO job has become more complex. Our model sidesteps both these issues because 1) our CEOs are risk-neutral; and 2) the reservation wages are the same for both the high type and the low type CEOs.

[11]Our Appendix A shows that a "winner-take-all" characterization is apt for the EV industry—a few firms such as Tesla and BYD dominate the EV space and the rest have small market shares. The Wall Street Journal likewise reports that in China, "one EV company makes money, and 100 lose money." The chief financial officer of Mercedes-Benz noted that EVs are "a pretty brutal space" and "I can hardly imagine the current status quo is fully sustainable for everybody." See `https://www.wsj.com/economy/trade/a-china-u-s-decoupling-you-aint-seen-nothing-yet-12c0828e` and `https://www.wsj.com/articles/lucid-electric-vehicle-losses-51035f63`.

the very race that creates the winners also creates the losers. The performance-based pay schemes for both the winning CEO and the losing CEO are also similarly intertwined, being simultaneously determined before the race, not after.

Other models endogenously induce performance-based pay as a solution to the matching problem of CEOs to firms when the CEO type is private information (adverse selection), and conditional on the CEO having matched to the firm, use performance-based pay as a means to induce CEO effort and mitigate moral hazard (e.g., Marinovic et al., 2024). However, a recently published academic survey of directors shows that the adverse selection problem of finding the right type of CEO is directors' top priority that easily outranks other priorities such as the moral hazard problem of designing an appropriate compensation package to motivate CEO efforts (Edmans et al., 2023, Table 2). We therefore focus on the adverse selection problem.

A comprehensive reference on contracting in single-period and multi-period adverse selection settings is Bolton and Dewatripont (2005, Ch.2 & 9). It shows how the informed party can earn equilibrium rents from the uninformed party and how these rents evolve as the parties interact repeatedly. However, these rents are sensitive to a myriad of institutional features, which has resulted in a large collection of models, each emphasizing a different institutional feature. We summarize a few salient models next.

The adverse selection problem has its roots in labor economics where the firm is interested in hiring a productive worker. In Harris and Holmstrom (1982), both the risk-averse worker and the risk-neutral firm are initially uninformed about the worker's type. The firm therefore insures the worker with a lifetime constant wage. As time passes, some workers are discovered to be more productive and are hired away by other firms for higher wages. The remaining workers continue to be paid as before.[12] Taylor (2013) uses this model to explain CEOs' downward rigid pay.

Another relevant study is Garrett and Pavan (2012) that builds a dynamic adverse selec-

[12]The firms thus earn the bulk of their lifetime profits in the early time periods when average wages are low.

tion and moral hazard model where each period the manager privately learns the current-period shock to his type and exerts effort. The board on the other hand decides every period how much to compensate the current manager and whether to replace him with a new manager of unknown type, based on the current manager's disclosures and outcomes. The optimal policy is excessive firing earlier and excessive retention later.

Finally, Beaudry (1994) models a setting where the principal comes in various types and is privately informed about her type whereas the agents are all identical. However, that study acknowledges that its model is better suited for lower-level roles where the job finalists, all of whom can be considered identical, are unaware of subtle office politics and culture. By contrast, the CEO, who is the agent in our model, makes and drives this culture.

The key insight from Beaudry (1994) is that equilibrium separation of types happens when the firm can pay a high base salary irrespective of performance. This salary structure is not feasible in our setting where the losing firm has no money to cover the CEO's reservation wage. We elaborate on this point in Section 4.3.

All the models discussed thus far assume information asymmetry about the CEO type. Oyer (2004) models a two-period where the employee type changes over time but is visible to the board. Specifically, the employee's publicly observable reservation utility or the market wage in the second period moves in the same direction as the firm's performance in the first period. In contrast to our setting, the wage premium in Oyer (2004) accrues to the low type CEO. This is because, unlike our model, the firm faces no financing constraint in paying the (risk-averse) employee and so optimal risk sharing entails a somewhat lower performance bonus and a somewhat high base wage which benefits the low type employee.

3 Model Setup

3.1 The One-Race Model

We follow the extreme value theory in Gabaix and Landier (2008) and represent the possible gross profit outcomes of the winner-take-all race with two values: the winner makes 1 and the loser ϵ, with $1 > \epsilon > 0$. We assume that the firm's investment to enter the race is normalized to zero.[13]

Investors are risk-neutral and appoint a rational value-maximizing board that searches for a manager. The key relevant attribute of the manager is her skill at delivering the high vs. the low outcome in the race. The board comes up with one finalist whose skill at winning the race the board is somewhat unsure of. All other candidates are thus presumed to have failed the board's assessment; the board believes that these remaining contenders have no chance of winning the race.

The finalist manager comes in two possible types, high $0 < t_H < 1$ or low $0 < t_L < t_H < 1$. The high (low) type manager can win the race with probability $t_H(t_L)$. For exposition, the high type manager is a "she" and the low type a "he." The manager knows her type but is unable to communicate it credibly to the board—all the board has is a prior on her type, specifically the board knows is that the probability of the low type t_L manager is $0 < p_L < 1$. (The probability of the high type t_H manager is thus $1 - p_L$.)[14]

To sidestep risk-sharing considerations in the pay package, the finalist manager is also assumed to be risk-neutral. Also recall that the key to adverse selection games is single-crossing of isoquant curves and not risk-aversion. With risk-neutrality, the isoquant curves become single-crossing lines which simplifies the analysis without removing the key intuition

[13] A note on zeros: To economize on notation, we normalize several additive constants in our model to zero. If these additive zeroes are slightly perturbed, our main inequalities will still hold due to continuity. We are careful not to use zeroes for multiplicative constants because zero is multiplicatively unique in the sense that all fractions of zero are zero.

[14] In Section 6, we show that our insights hold when we consider more than two CEO types.

(Kreps, 1990, Ch.17).[15]

Each finalist manager irrespective of her type earns $\epsilon < u < 1$ in some other career.[16] The fact that u is same for both types of managers means that other career thus reveals nothing about the finalist manager's skill in running the present firm.[17] One interpretation of type, therefore, is the manager's intangible "fit" with firm. A high type manager's leadership style is more likely to lead the present firm to success. A low type manager's intangible leadership skills are a better fit for some other firm, and therefore this manager commands the same reservation wage despite being a low type from the present firm's perspective.[18] Nonetheless, our main inequalities still hold if the high type CEO has a slightly higher reservation wage than the low type.

The board credibly offers a one-time take-it-or-leave-it menu of offers to the finalist. As is standard, we assume that the finalist will always choose our firm over other opportunities if offered u in expectation. There is no renegotiation if the finalist declines (Kreps, 1990, Ch. 18). One potential justification for this assumption is that the window of opportunity is small and the firm cannot engage in protracted negotiations. Instead, the board goes back to a semifinalist, who the board has assessed as distinctly inferior to the finalist and unable to win the race and thus makes the firm zero.[19]

We assume, as described in the Introduction and footnote 4, that a firm competing in the winner-take-all race cannot borrow or hedge the high outcome 1 against the low outcome ϵ, and neither can the manager. The payoff gap is too large and the winning probability too subjective for the board and the insurer to agree on pricing terms. This assumption is quite

[15]The finalist manager's risk neutrality also implies that the board's payment menu to its CEO can either be ex ante contracted amounts or ex post discretionary performance and retention rewards over which the risk-neutral CEO has ex ante rational expectations.

[16]If $u < \epsilon < 1$ the first best strategy is to pay the manager u in all states of the world (the firm earns more than u in every state).

[17]Garrett and Pavan (2012) also make the assumption that managers of all types have the same reservation utility elsewhere.

[18]As an example of a firm that replaced its CEO due to lack-of-fit concerns, see `https://www.wsj.com/business/nestle-ceo-abruptly-ousted-as-board-seeks-better-cultural-fit-a19a63ff`.

[19]Alternatively, the semifinalist could be modeled as someone who makes the firm slightly more than zero and the main inequalities would still hold.

standard (Edmans et al., 2017), but this means that the board has to pay the manager out of what she makes for the firm. The board also cannot impose negative pay on the manager—she cannot disgorge money to the firm due to her limited wealth. We further assume the standard monotonocity condition where both the manager and the investors earn more when the firm does better (Innes, 1990).[20]

The above discussion suggests that the manager's contract is a menu of pairs of $\{\epsilon_p, 1_p\}$ where ϵ_p in any pair is the payment to the manager when she delivers ϵ and 1_p is the payment when she delivers 1. Investors keep the residuals $\epsilon - \epsilon_p$ and $1 - 1_p$.[21]

Every contract payment pair $\{\epsilon_p, 1_p\}$ obeys the following feasibility and monotonicity constraints that ensure that there is always enough money to positively pay the manager and both the manager and the investors earn more under the high outcome 1:

$$\begin{aligned} 0 \leq \epsilon_p \leq \epsilon \\ 0 \leq 1_p \leq 1 \\ \epsilon_p \leq 1_p \\ \epsilon - \epsilon_p \leq 1 - 1_p \end{aligned} \tag{1}$$

We have to add an additional feasibility condition because we assume that it is not possible to finance the manager's pay using borrowing. We therefore also have to assume that the firm's expected gross profits under a manager are larger than u, the manager's reservation wage.

$$t_L 1 + (1 - t_L)\epsilon \geq u \tag{2}$$

It is enough to assume the condition above for the t_L manager because it will then

[20] In subsequent sections, we show that not all of these restrictions are strictly binding in equilibrium. Nonetheless, these assumptions greatly constrain the nature of the feasible pay structure and prevent a separating equilibrium as we show in Section 4.3.

[21] Oyer (2004, fn. 4) notes that these contracts can be viewed either as profit sharing or as employee ownership because all proceeds are distributed to the manager and investors.

automatically hold for the t_H manager.[22] The feasibility condition (2) along with $u > \epsilon$ implies that there a payoff $\epsilon < x < 1$ such that $t_L x + (1 - t_L)\epsilon = u$, so it meets the reservation wage for the low type manager. Since $t_H > t_L, 1 > \epsilon$, $t_H x + x(1 - t_H)\epsilon > u$ and the high type manager also accepts this ownership share. It is straightforward to see that this payment scheme also meets all the no-borrowing and monotonicity requirements in (1). Given the outcome set $\{\epsilon, 1\}$, the manager earns $\{\epsilon, x = \frac{u-(1-t_L)\epsilon}{t_L}\}$ and investors earn $\{0, 1 - x\}$, and both are weakly positive and increasing. It will turn out that this is the pooling equilibrium contract.

3.2 The Two-Race Model

We assume that the first race is followed by a new identical race. There is no discounting. Following Oyer (2004, p. 1622), any winnings from the first race are returned to shareholders immediately via dividends. As a result, all the payoff and financing constraints in the second race are identical to the first race. The only things that carryover from the first race to the second are the board's Bayesian posterior beliefs on the incumbent CEO type based on the first race's outcome.

The board has more hiring and compensation choices in the second race than the first because it has an incumbent about whose type it has acquired more information from the first race's outcome. The board can either fire the incumbent or make him a new offer that he can accept or reject. If he rejects, the board searches for a new CEO—the separated incumbent can always ask to be considered in this search.[23] As before, we assume that the incumbent or the new finalist will choose the current firm over outside opportunties if she

[22]This is because $0 < t_L < t_H < 1$ and $\epsilon < 1$, so $t_H 1 + (1 - t_H)\epsilon > t_L 1 + (1 - t_L)\epsilon$. Applying this restriction to both types of managers is limiting in that if the board chooses to offer a contract that it knows only the high type manager will accept, it is fine if equation (2) is violated by a low type manager as long as it is not violated by a high type manager. We also note that the feasibility assumption (2) is not entirely innocuous. If we wish to create an extreme gap in the probability of the high outcome across the two CEO types, i.e., $t_L << t_H$, t_L needs to be very close to 0, which in turn forces u to be very close to ϵ from above. We therefore cannot choose u freely.

[23]Garrett and Pavan (2012) make the similar assumption that the board can search for a new manager only after the current manager quits.

expects to earn u.

We consider two contracting possibilities: first, the board can credibly commit to a two-period contract that allows the incumbent to make less than u in expectation in the first race but then recover any arrears in expectation from the winnings of the second race. By contrast, we prohibit bonding of management: the incumbent can quit at will after the first race. Intuitively, this prohibition prevents overpayment to the incumbent in the first race and underpayment in the second because then the incumbent will collect the first-race overpayment and then quit to earn more elsewhere in the second race.

We also require that such contracts be renegotiation-proof, i.e., the second-period contract can be altered in the second period if both parties agree. However, Oyer (2004, Section IV.A) argues that the board may not be able to credibly commit to two-period contracts, especially those that underpay in the first period and thus require the firm to overpay in the second; so we also consider the setting with only period-by-period contracts that pay at least u each period for the incumbent or the accepting finalist.

4 Solving the One-Race Model

The first observation we make is the single-crossing property of the two types' indifference curves:

$$t_L 1_p + (1 - t_L)\epsilon_p = u$$
$$t_H 1_p + (1 - t_H)\epsilon_p = u$$

Figure 1 shows these curves in the $(\epsilon_p, 1_p)$ space. Both curves are downward sloping, with the t_H curve being flatter than the t_L curve (because $0 < t_L < t_H < 1$), and meet at (u, u). Since $u > \epsilon$, this point is above and to the right of (ϵ, ϵ).

The shaded rhombus in Figure 1 is key to the solution. It represents the feasible payout region to the manager given the monotonocity and feasibility conditions (1). Note that if

t_L is too small, the low type indifference curve will be too steep and not hit the rhombus: the low type CEO is so unlikely to generate the 1 payoff and so likely to generate the $\epsilon < u$ payoff that there is no way to pay him u in expectation from the firm's proceeds. This is the reason we assert the necessary feasibility conditions (1) and (2) and show the existence of a feasible solution in the previous section.

Figure 1 shows the nature of the solution as well. If the board offers only point B as a take-it-or-leave-it menu, only the high type CEO will accept (she earns u in expectation). The low type CEO's indifference curve of u is above point B, so she will not accept. Investors make a profit of $(1 - p_L)(t_H 1 + (1 - t_H)\epsilon - u)$.

Suppose the board offers both points A and B as payoff menus. Point A is the (ϵ, x) offer introduced earlier in this section where x solves $t_L x + (1 - t_L)\epsilon = u$. Both CEO types will accept A because the low type earns u and the high type earns $t_H x + (1 - t_H)\epsilon > u$.[24] Investors make a profit of $(1 - p_L)t_H(1 - x) + p_L t_L(1 - x)$. Indeed, $t_H(1 - x) > t_L(1 - x)$ means investors earn more under the high type than they do under the low type despite paying the high type a premium.

Note that if the board offers point A, A', and B on the menu, the low type is indifferent between A and A', but the high type earns more at A' compared to A; so she will choose A'. This reduces the payoff to investors. The board is thus best off offering either point A on the right border of the feasible rhombus (the cheapest solution to both types accepting). Another possibility that we discuss for completeness is the trivially separating equilibrium of point B that only the high type finalist accepts and receives no premium; if the finalist turns out to be the low type, he will walk and the firm makes zero.

In sum, the board picks one finalist whose type it does not know. The board can give this finalist the performance-based contract A which the finalist will accept irrespective of her type. The benefit here is that the firm operates irrespective of the finalist's type and the cost is that the high type finalist earns a rent or a premium. Alternatively, the board can

[24] This is because because $x > \epsilon, t_H > t_L$.

offer the finalist the performance-based contract B that only the high type finalist accepts. The benefit here is that the board does not pay the high type finalist a premium and the cost is that the firm shuts down if the finalist is a low type (note that the low type finalist is not a liability to the firm—he just makes less for the firm than the high type). The board thus has to solve the two different programs A and B and pick the most profitable one. We summarize the outcomes of the one-race equilibrium as follows:

There are only two possible equilibrium outcomes: a) the board offers a pooling contract that both types of finalists accept; or b) the board offers a contract that only the high type finalist accepts. One can view case b) as a trivially separating equilibrium. We discuss the possibility of a meaningful separating equilibrium in Section 4.3.

The payoff to the firm from writing a pooling contract that both the low and high type finalists accept is (recall that $\epsilon < x = \frac{u-(1-t_L)\epsilon}{t_L} < 1$*):*

$$LH(p_L) \equiv \qquad (1-p_L)t_H(1-x) + p_L t_L(1-x)$$

Note that $LH(p_L)$ *is decreasing in* p_L *because* $t_L < t_H, x < 1$*. The more likely the finalist is a low type, the less the firm earns. On the other hand, the payoff to the firm from writing a contract that only the high type finalist accepts is:*

$$H(p_L) \equiv \qquad p_L 0 + (1-p_L)t_H(1 - \frac{u-(1-t_H)\epsilon}{t_H})$$

The one-race equilibrium strategy is the maximum: $\max[LH(p_L), H(p_L)]$.

We will use the terms LH, H to denote both the strategies and the firm payoffs.[25]

[25]It is also clear that the bigger of LH and H will continue to remain so if the point (u, u) is slightly perturbed for the high type in Figure 1 or if the board responds to the low type finalist's rejection by hiring a semifinalist who makes a small δ instead of 0.

4.1 An Analysis of the Equilibrium

In the pooling equilibrium, the ex ante probability that a firm is run by a high type CEO is $1 - p_L$. The ex post chance of a successful 1-outcome firm run by a high type CEO is $\frac{(1-p_L)t_H}{(1-p_L)t_H+p_L t_L} > \frac{(1-p_L)t_H}{(1-p_L)t_H+p_L t_H} = 1 - p_L$ (recall that $t_L < t_H$). A winning firm is thus ex post more likely to be run by the high type CEO relative to baseline probabilities. This CEO is also overpaid relative to u, her opportunity cost, an outcome that is not the result of a captive board, but instead is a manifestation of survivorship bias with a rational value-maximizing board, i.e., ex ante rationally deserved pay may appear ex post excessive if one only looks at surviving firms (in the case both CEO types accept).

We also note that the high type CEO receives excess pay when both the high type and the low type CEO receive identical contracts. In fact, this lack of variation is what drives excess pay for the high type CEO. This finding is thus contrary to the skimming argument that many CEOs who earn excess pay do so through pay-for-performance contracts that unfairly reward poor performance (Bebchuk and Fried, 2003). In fact, the high type CEO is more likely to generate the high outcome 1, and so she is more profitable to investors despite her premium. Investors earn $t_H(1 - x) > t_L(1 - x)$ under her than they do under the low type.

4.2 A Numerical Example

To help clarify the solution discussion, consider the following numerical example that is consistent with Figure 1. Suppose that $\epsilon = 0.1$, which means that the high outcome of 1 is $10\times$ the low outcome, and $u = 0.11 > \epsilon$. Let the probability of the low type CEO $p_L = 0.7$, and the high and the low type talent be $t_H = 0.9$ and $t_L = 0.2$. That is, the high type CEO is 4.5 times more likely to generate the high outcome than the low type CEO.

Let point $A = (\epsilon, 0.15)$ and point $B = (\epsilon, 0.111)$, below point A but above the point (ϵ, ϵ). Both A and B are below $(\epsilon, 1)$. Picking point B (where only the high type CEO accepts) gets

investors an expected profit of 0.24. Picking an inefficient point A' $= (0.05, 0.35)$, which is to the left and above A but still in the feasible rhombus gives investors a larger profit of 0.296. Finally, picking point A (where either type of CEO accepts) gives investors the maximum expected profit of 0.3485. The low type CEO earns $u = 0.11$ and the high type CEO earns a larger expected wage of $0.145 > u$. The ex post probability of a winning CEO being a high type is 0.658, almost double the ex ante probability $1 - p_L = 0.3$. Survivorship bias is thus the main reason why we see winners earning rents.

We next examine the the severity of the winner-take-all scenario, which we can model as the winning outcome of 1 becoming bigger or the losing outcome of ϵ becoming smaller (see fn. 11). Figure 2 shows the impact of ϵ shrinking to $\epsilon' < \epsilon$. The feasible rhombus becomes vertically narrower and the optimal possible points now become B (only the high type CEO accepts) or A' (both types accept)—A is no longer feasible. Since A' is to the left and above A, the high type makes more and investors earn less. Returning to the numerical example above, if $\epsilon' = 0.05$, where the high outcome 1 is $20\times$ the low outcome, and A' $= (\epsilon', 0.35)$, the low type CEO still makes $u = 0.11$, the high type CEO makes 0.32 (which is bigger than her previous expected wage of 0.145). Although 0.2665 is less than investors' previous expected profit of 0.296, it is still more than the 0.2385 they would make by offering the wage menu B' $= (\epsilon', 0.1166)$ that only the high CEO type would accept because only she can earn the minimum reservation wage of 0.11.

Intuitively, as the low outcome ϵ shrinks, but everything else remains unchanged, the low type CEO needs to be paid more of the high outcome to get to her reservation wage of u. This in turn further increases the payoffs to the high type CEO. By contrast, there is no change in the CEO payoff structure when the high outcome increases from 1 to say 2 (see Figure 3). This is because the old wage structures continue to be both feasible and the cheapest compensation plan. Investors thus reap all the benefits of increased high

outcomes.[26] In sum, the high type CEO benefits more when the winner-take-all situation reduces the low outcome, and investors benefit when the winner-take-all situation increases the high outcome.

4.3 What Prevents a Separating Equilibrium?

To see why there is no separating equilibrium other than the trivial case $H(p_L)$ where the high type does not receive a premium, assume that the board can offer the contract $\{u, u\}$ to the finalist. Both types of finalists accept and receive no premium.[27] The only problem is that the payoff u is not feasible when the firm loses the race and earns $\epsilon < u$. This observation leads to the natural two-period extension where the board promises to pay the arrears $u - \epsilon$ (appropriately amortized) if the firm wins the second race. The future firm thus acts as an insurer to the current firm. We will consider this case in Section 5.

As noted above, a trivial type of separation is $H(p_L)$ where it is optimal to have only the high type finalist accept. A feature of the model is that the premium the high type gets when both types accept (relative to when only the high type accepts) is independent of p_L, the ex ante probability of the low type. The pooling pay scheme $LH(p_L)$ is $\{\epsilon, x = \frac{u-(1-t_L)\epsilon}{t_L}\}$ whereas the high-type only pay scheme $H(p_L)$ is $\{\epsilon, \frac{u-(1-t_H)\epsilon}{t_H}\}$. If one makes p_L low enough, the quantity of expected profit to the firm from the low type CEO stays unchanged at $t_L(1-x) > 0$ but its occurrence probability p_L will go down to zero, whereas the expected wage premium to high type CEO stays unchanged at $t_H(\frac{u-(1-t_L)\epsilon}{t_L} - \frac{u-(1-t_H)\epsilon}{t_H}) > 0$ but its occurrence probability $1-p_L$ will go up to one. For p_L low enough $H(p_L) > LH(p_L)$. In the two-period model therefore, we assume that the probability of the low type in both periods is large enough to preclude this trivial separation.[28]

[26] Likewise, the investor montonocity constraint is also not strongly binding because in the low outcome ϵ situation, the CEO gets everything and investors make the minimum zero. Their fortunes can thus only improve in the high outcome case.

[27] In Figure 1, separation with no premium can be obtained by offering to the finalist CEO any point on the $A' - A$ line below and to the right of (u, u), and any point on the B line above and to the left on (u, u).

[28] To continue the numerical example from the previous subsection, assume that the probability of the low type p_L falls from 0.7 to 0.07. A pooling contract that both types accept earns the firm 0.723 whereas a contract that only the high type accepts earns the firm 0.744.

5 Solving the Two-Race Model

We first assume that the board can writing binding two-period contracts. Figure 4 shows the details of what follows. As noted in Section 4.3, we assume that p_L, the ex ante probability that the finalist in the first race is the low type, is large enough that the board prefers to hire both types of finalists instead of the trivial separation case of just the high type . We then follow the idea in Oyer (2004, Section IV.A), and attempt to hold the first period wages for both types to u irrespective of the outcome. This wage is feasible when the outcome is $1 > u$, but not when the outcome is $\epsilon < u$. We will soon show that is best to pay the full ϵ out in wages and leave the smallest arrears $u - \epsilon$ possible.

In the second race following the loss in the first race, the incumbent plays the game $LH(p_L^\epsilon)$ where $p_L^\epsilon > p_L$ is the posterior Bayes probability of the incumbent being the low type given that he lost the first race. As in the one period game, it is optimal to pay the incumbent ϵ for losing and $x = \frac{u-(1-t_L)\epsilon}{t_L}$ for winning. Note that x is independent of p_L.

However, this pay is not enough because we still owe the arrears $u - \epsilon$ in expectation to the incumbent. The high type incumbent, being risk-neutral, will accept a further pay of $\frac{u-\epsilon}{t_H}$ if she wins the second race, whereas the low type incumbent will need $\frac{u-\epsilon}{t_L} > \frac{u-\epsilon}{t_H}$. To get both types of incumbents to accept the contract before the first race, the board must promise $x + \frac{u-\epsilon}{t_L}$ for losing the first race but winning the second.[29]

The calculations above and in Figure 4 also show why it is important to have the smallest amount of unpaid arrears possible because the amortization factor is > 1. The arrears $u - \epsilon$ become $\frac{u-\epsilon}{t_L}$ in the second race that have to be paid out with the probability $p_L^\epsilon t_L + (1 - p_L^\epsilon)t_H > t_L$ (because $t_H > t_L$) and so the overall expected amortized arrears are greater than $u - \epsilon$. The value maximizing board wants to reduce this extra wage premium as much as possible which is why it pays out the entire winnings ϵ to the incumbent if he loses the first race.

[29]Here we must assume that $1 > x + \frac{u-\epsilon}{t_L}$, which we will show is feasible using a numerical example.

Things are more straightforward if the incumbent wins the first race. We pay her u and begin the second race anew (we discuss why it is optimal to pay u and no more or no less in the next paragraph). As Figure 4 shows, the board's posterior belief that the incumbent is the low type is $p_L^1 < p_L$, i.e., the board is more confident it has landed a high type incumbent and so it is best to retain her and not start a new search. Assuming $LH(p_L^1) > H(p_L^1)$, i.e., we can rule out the trivially separating case, it is best to pay the one-period wages of $\{\epsilon, x = \frac{u-(1-t_L)\epsilon}{t_L}\}$ in the second and final race. In sum, when the board can commit to a two-race contract, it can eliminate any wage premium in the first race and pay the standard premium and an appropriately amortized wage arrears in the second race.[30]

The preceding paragraph assumes that if the firm wins the first race, the board pays the incumbent the full u and incurs no unpaid arrears. The reasoning is the same as for the case when the firm loses the first race: the second-race amortization factor of unpaid arrears is > 1. Likewise, the board does not pay more than u in the first race and less than u in the second: the incumbent of either type will initially accept this offer, but then quit in the second period and re-enlist when the board starts a new search.[31] Because the board knows that the incumbent who has won the first race is more likely to be the high type than the new finalist (and thus more profitable), it will rehire the incumbent with full premium. Anticipating this, the board will pay the full u in the first period.

Figure 5 shows what happens when the board cannot write two-period contracts. It pays the one-period optimum $\{\epsilon, x = \frac{u-(1-t_L)\epsilon}{t_L}\}$ to both types of incumbent in the first period (the high type incumbent earns rents). The interesting outcome is when first race is lost. The

[30] In the event that $H(p_L^1) > LH(p_L^1)$, the board would offer a payment that holds the high type incumbent to u; the low type then quits. But then the high type would also quit and join the new CEO search where she would earn a premium if selected. Then the low type incumbent would mimic this behavior as well because he prefers the current firm when paid u. But once both types of incumbents quit and express interest in the new search, the board is better off hiring them because it knows that the incumbent is more likely the high type than the other finalists ($p_L^1 < p_L$). The board would then offer them $H(p_L^1)$ because $H(p_L^1) > LH(p_L^1)$. This creates a circularity that we wish to avoid. We can also circumvent this circularity by additionally asserting that the board offers $H(p_L^1)$ which the incumbent accepts if she is the high type and rejects if she is the low type. We acknowledge that this contingency will become more salient as the number of races increases and the incumbent wins all of them. The board will then be very sure that the incumbent is the high type.

[31] The high type will quit and re-enlist and the low type will mimic her.

board now knows that the incumbent is more likely the low type with probability $p_L^\epsilon > p_L$. The board can squeeze this incumbent by offering him the $H(p_L^\epsilon)$ pay of $\{\epsilon, \frac{u-(1-t_H)\epsilon}{t_H}\}$. The high type incumbent makes u and so will accept whereas the low type incumbent makes $< u$ and so will walk. At this point the board can start a new search and land a high type finalist with probability $p_L > p_L^\epsilon$ (recall that the value-maximizing board prefers the high type finalist despite the wage premium because the firm makes more).

Given this board strategy, why wouldn't the high type incumbent reject the $H(p_L^\epsilon)$ offer that pays no premium and enlist herself for the new CEO role where if she succeeds she will get a premium? She will not because the low type incumbent will also mimic her in enlisting for the new search and then the board assesses the incumbent to be of low type with probability $p_L^\epsilon > p_L$ and so will not consider her as a finalist (recall both types of incumbents would prefer to work for the current firm if they can make at least u in expectation). Anticipating this sequence of events, the high incumbent accepts the offer with no premium and the low type quits, enabling the board to find a new CEO for the second race. We believe that this is a meaningful separating equilibrium.

Things are more straightforward if the incumbent wins the first race. The board pays her x and begins the second race anew. As Figure 5 shows, the board's posterior belief that the incumbent is the low type is $p_L^1 < p_L$, i.e., the board is more confident it has landed a high type incumbent and so it best to retain her and not start a new search. Assuming $LH(p_L^1) > H(p_L^1)$, i.e., we can rule out the trivially separating case (see footnote 30), it is best to pay the one-period wages of $\{\epsilon, x = \frac{u-(1-t_L)\epsilon}{t_L}\}$ in the second and final race. In sum, when the board cannot commit to a two-race contract, it pays the wage premium in the first race and squeezes the incumbent should he lose but continues the premium should he win.

Comparing the two games in Figures 4 and 5, the pay squeeze in Figure 6 happens in the second race unlike the two-period contract setting in Figure 4 where the pay squeeze happens in the first period and board pays a premium plus amortized upaid arrears for winning the second race. Furthermore, note that the outcomes in both the two-period

games are renegotiation-proof because either party, the board or the incumbent CEO, will not willingly make a second-period counteroffer other than the current equilibrium that the other party will willingly accept over the second-period status quo.

A numerical calculation shows that neither the two-period contract nor the period-by-period contract setting always dominates the other. If $u = 0.11, \epsilon = 0.1, p_L = 0.7, t_H = 0.9, t_L = 0.2$, the period-by-period game in Figure 5 yields greater ex ante profits than the two-period contracting model in Figure 4 (0.798 vs. 0.706). If we change to $u = 0.4, t_L = 0.85$, the two-period contracting model dominates ex ante (0.951 vs. 0.948).[32] But the basic idea still endures, namely that the premium for winning is intricately tied to the firm's inability to pay out enough to the CEO after losing. The timing of this premium, however, depends on the specifics of the two-period model. Winning and losing are thus two sides of the same coin and must be considered jointly, not separately (see footnote 8).[33]

6 Variants of the Single-Race Model

6.1 An Endogenous Winner-take-all Contest With Two Firms Playing Simultaneously

In this subsection, we embed the one-race one-firm model in a contest between two identical firms simultaneously playing the winner-take-all race if it ever begins, and facing i.i.d uncertainty on their finalist CEO type. Each firm simultaneously decides to make a take-it-or-leave-it offer to its finalist CEO, i.e., we ignore the possibility of repeated offers and other hiring games between the two firms.

We model the winner-take-all outcome by positing that if both firms' CEOs turn to be the same type, one firm randomly and independently wins the contest with probability half and the other loses. If the CEO types are unequal, the firm with the high type CEO wins

[32] We check that all the feasibility conditions and inequalities including $LH(p_L^1) > H(p_L^1)$ are satisfied.
[33] We leave the case of infinite-period races to future research.

the 1 outcome for sure and the other firm loses for sure and gets ϵ.

A consequence of the above setup is that for any hiring strategy employed by the other firm, a high type CEO in the focal firm has a higher chance of winning than a low type. As a result, any contract acceptable to the low type is also acceptable to the high type, just as in Section 4. We therefore only consider an offer that only the high type CEO accepts (the trivial separating equilibrium) or an offer that both types accept. We then show a numerical example where it is a dominant strategy for each firm to hire both types of CEOs and as a result overpay the high type relative to her reservation utility u. Our main intuition thus continues to hold.

We now list the payoffs to firm 1 for various choices of firm 2. Note that risk neutrality is critical here because it enables us to linearly aggregate independent probabilities and payoffs across both pure and mixed strategies:[34]

1. **Both firms offer contracts that both type CEOs accept.** Recall that the probability of a low type CEO is p_L and that these probabilities are independently drawn for the two firms. Furthermore, we are endogenizing t_H, the probability that the high type CEO wins, and similarly t_L.

   If firm 1 draws a high type CEO, with independent probability $(1 - p_L)$ firm 2 also draws a high type CEO and firm 1's probability of winning is half. On the other hand, with independent probability p_L firm 2 draws a low type CEO and now firm's 1 probability of winning is 1. On net, a high type CEO at firm 1 will win with probability $t_H = \frac{1-p_L}{2} + p_L \leq 1$.

   Similarly, if firm 1 draws a low type CEO, its only hope is that firm 2 also draws a low type CEO (which will happen with independent probability p_L), in which case each firm has probability half of winning. Therefore, $t_L = \frac{p_L}{2} < t_H$.

2. **Firm 1 offers a contract that both type CEOs accept and firm 2 offers a**

[34] We will not need mixed strategies in our numerical example because we will show that a pure strategy is dominant for both firms.

contract that only the high type CEO accepts. Once again for firm 1, $t_H = \frac{1-p_L}{2} + p_L$. But t_L improves to p_L because firm 2 will shut down when it draws a low type CEO because she will not accept the contract. Note again that $t_H > t_L$.

3. **Firm 1 offers a contract that only the high type accepts.** In this case, it doesn't matter which contract firm 2 offers. For firm 1, $t_H = \frac{1-p_L}{2} + p_L$.

We recompute our previous numerical example in Section 4 where $\epsilon = 0.1$, i.e., the high outcome of 1 is $10\times$ the low outcome, and the reservation wage $u = 0.11 > \epsilon$. The probability of the low type CEO $p_L = 0.7$.

We find that offering a contract that both type CEOs accept is a dominant strategy that pays off 0.4357 to each firm in expectation (the low type CEO earns in expectation the reservation utility of 0.11 and the high type a premium wage of 0.1243). If the other firm offers a contract that only a high CEO accepts (she earns the reservation utility 0.11), that firm makes only 0.2265 whereas the focal firm makes even more at 0.6598.[35] The advantage of offering a contract that only the high type CEO accepts is that the she does not need to be paid a wage premium, but the loss to the firm from having the low type CEO walk exceeds that benefit (recall that the low type CEO still generates profits—she is just not as effective as the high type CEO). In sum, as long as the parameters of the game are such that both CEO types accept, the premium to the high type CEO endures even after endogenizing the winner-take-all game.

6.2 Multiple CEO Types and Multiple Performance Outcomes

Figure 6 considers the single-race single-firm case when there are three types of CEOs: $0 < t_L < t_M < t_H < 1$. The single-crossing conditions greatly simplify the problem. The board has to choose between offering C (only t_H accepts), or B (only t_H, t_M accept), or A

[35]This is because the low type CEO at the focal firm has a higher chance of winning (due to the low type CEO rejecting the other firm) which reduces the wage premium to the high type. Both these effects increase expected profits.

(all CEOs accept irrespective of their type).[36] The board thus trades off overpaying more talented CEOs versus earning zero when the CEO happens to be less talented and rejects the contract because it earns him less than u in expectation (the trivial separating equilibrium). More finalist types thus simply imply that there are CEOs of intermediate talent whose wage gains over u when the high outcome obtains are also of intermediate magnitude.

Reverting back to two CEO types, we next consider the setting of two low outcomes ϵ_1 and ϵ_2 and one high outcome 1: $0 < \epsilon_1 < \epsilon_2 < u < 1$ (recall that u is the reservation wage). As before, this ordering implies that reservation wage cannot be financed fully from the low outcomes. We assume that the low type CEO generates the three outcomes $\epsilon_1, \epsilon_2, 1$ with probabilities $0 < p_1, p_2, p_3 < 1$ that sum up to one. We use the notion of first order stochastic dominance and assume that the high type's CEO probabilities are $p_1 - \Delta_2 - \Delta_3, p_2 + \Delta_2, p_3 + \Delta_3$, where the Δ's are feasible strictly positive numbers. The high type CEO thus is more likely to generate higher outcomes than the low type CEO.

We next get the straightforward inequality that if $x \leq y \leq z$ with at least one strict inequality, then:

$$\begin{aligned} &(p_1 - \Delta_2 - \Delta_3)x + (p_2 + \Delta_2)y + (p_3 + \Delta_3)z \\ &= p_1 x + p_2 y + p_3 z + \Delta_2(y - x) + \Delta_3(z - x) \\ &> p_1 x + p_2 y + p_3 z \end{aligned} \tag{3}$$

The simple economic implication is that any increasing positive wage scheme over the three outcomes $\epsilon_1 < \epsilon_2 < 1$ pays the high type more in expectation. Since an equal pay scheme across the three outcomes is not possible (given that $u > \epsilon_2$), any feasible pay scheme that earns the low type CEO u will earn the high type CEO strictly more than u.

However, we first need to assume as before that there is a monotonic self-financing com-

[36] Also note that A is preferred to B and B is preferred to C by all types, so offering multiple menu items in the compensation contract is redundant.

pensation plan that is feasible for the low type CEO, i.e., we assume:

$$p_1\epsilon_1 + p_2\epsilon_2 + p_3 1 \geq u$$

The above equation implies that since $1 > u > \epsilon_2 > \epsilon_1$, there is a payment x where $1 \geq x > \epsilon_2$ that earns the low type CEO $p_1\epsilon_1 + p_2\epsilon_2 + p_3 x = u$. This plan is self-financing and both the manager and the investors earn weakly positive and monotonic amounts for the three outcomes $\{\epsilon_1, \epsilon_2, 1\}$: namely $\{\epsilon_1, \epsilon_2, x\}$ for the manager and $\{0, 0, 1-x\}$ for the investors respectively. Finally, as discussed above, the first order stochastic dominance inequality (3) implies that the above payment scheme delivers more than u to the high type CEO. Thus, there does exist a candidate solution and we can turn our attention to the optimal one.

Recall from the previous section that the board picks the best of the two optimization programs: it offers an increasing pay scheme that pays u to the high type CEO and by (3), less than u to the low type CEO who will reject it, or it will offer a pay scheme that both type CEOs accept. The second pay scheme will be more profitable to the board in many instances.[37] Furthermore, any feasible positive monotonic pay scheme that delivers u to the low type CEO will deliver strictly more than u to the high type CEO, due to (3) as discussed above. There will therefore be a strictly positive probability of excess CEO pay in the second pay scheme's equilibrium.[38]

[37]Recall that the the probability of the high type CEO $1 - p_L$ is a free variable that does not affect the high type CEO's wage premium. If $1 - p_L$ is vanishingly small, the ex ante chances that the board will have to pay the wage premium to the high type CEO in the second pay scheme will be small compared to the *ex ante* expected lost profits from the low type CEO (who occurs with very high probability p_L) rejecting the first pay scheme.

[38]It is easy to see the optimal corner pay solution in the two-dimensional linear program case in Figure 1. The same cannot be said for three dimensions because we have too many free variables. If $\Delta_2 > \Delta_3$, the optimal pay solution to the linear program may not be the corner solution $\{\epsilon_1, \epsilon_2, x\}$. It may so happen that the high type CEO is marginally more likely to generate ϵ_2 over 1, in which case it may cheaper ex ante to pay slightly less than ϵ_2 when the outcome is ϵ_2 and slightly more than x when the outcome is 1. This will yield another corner pay solution. Nonetheless, the high type CEO will continue to earn a premium.

7 Conclusion

A salient feature of modern financial markets is that most of their wealth creation is due to a few firms (Bessembinder et al., 2023). The CEOs of these firms then receive very high pay, raising questions of pay fairness and governance (Bebchuk and Fried, 2003). In response, studies argue that firms compete in winner-take-all contests that result in a few very large winners and many small losers. After winning the race, the winners hire top talent as CEOs and pay them extremely high amounts. However, because these amounts are determined through a rational CEO search process, they are deemed as deserved pay (Gabaix and Landier, 2008).

This study on the other hand argues that boards search for CEOs before the winner-take-all race begins, not after, and that these CEOs are crucial to the race outcomes. The occurrence of few big and many small firms is then not just happenstance but instead two sides of the same coin: the very race that creates the winners also creates the losers. The performance-based pay schemes for both the winning CEO and the losing CEO are also similarly intertwined, being simultaneously determined before the race, not after.

We model the possible gross profit outcomes of the winner-take-all race as high or low and then add an adverse selection component to the CEO type. This addition is motivated by Edmans et al. (2023)'s survey finding that directors' primary concern is finding the right type of CEO. The winner-take-all setting further constrains the way a CEO can be paid: there is no possibility for the firm to insure across the two highly disparate high versus low outcomes and so the CEO must be paid little in the low outcome and more in the high outcome. This constraint is more severe for the low type CEO because he cannot make enough to meet his reservation wage in the low state that he is more likely to generate, and so must be compensated a larger fraction of the high state that he is less likely to generate. This naturally advantages the high type CEO who is more likely to generate the high state. As a result, winning CEOs, who are more likely to be the high type relative to their baseline

probability, can receive pay that is both ex ante rational (the board is value-maximizing not captive) and ex post excessive (relative to the CEO's reservation opportunities). We extend the one-period model to two periods and show that the ex ante optimal but ex post excessive premium or rent endures: what changes is the timing of this rent (depending on the model).

The idea that the high type agent earns rent and that this rent evolves over time is not new to theorists, who have nonetheless built a vast number of models on this topic (Bolton and Dewatripont, 2005, Ch.9). The reason is that the nature of these rents is highly sensitive to the institutional details and the specific intuition in one setting rarely carries over to another. In our winner-take-all setting, these rents are of empirical interest because they manifest themselves as survivorship bias, i.e., pay that is rational ex ante when one looks at all the race entrants may appear excessive ex post if one only looks at the surviving winners. Indeed, as our footnote 8 shows, losing firm CEOs have been equally or even more expensive to shareholders than winning firms. The main message of our model is then an empirical one, namely that empirical analyses could benefit from paying equal attention to both winners and losers simultaneously.

References

Beaudry, P., 1994. Why an Informed Principal May Leave Rents to an Agent. International Economic Review 35, 821–832.

Bebchuk, L. A., Fried, J. M., 2003. Executive compensation as an agency problem. Journal of Economic Perspectives 17, 71–92.

Bessembinder, H., Chen, T.-F., Choi, G., and, K. C. J. W., 2023. Long-term shareholder returns: Evidence from 64,000 global stocks. Financial Analysts Journal 79, 33–63.

Bolton, P., Dewatripont, M., 2005. Contract Theory. The MIT Press.

Edmans, A., Gabaix, X., Jenter, D., 2017. Executive compensation: a survey of theory and evidence. ECGI Working Paper .

Edmans, A., Gosling, T., Jenter, D., 2023. Ceo compensation: Evidence from the field. Journal of Financial Economics 150, 103718.

Gabaix, X., 2016. Power laws in economics: An introduction. Journal of Economic Perspectives 30, 185–206.

Gabaix, X., Landier, A., 2008. Why has ceo pay increased so much? Quarterly Journal of Economics 123, 49–100.

Garrett, D. F., Pavan, A., 2012. Managerial Turnover in a Changing World. Journal of Political Economy 120, 879–925.

Harris, M., Holmstrom, B., 1982. A theory of wage dynamics. Review of Economic Studies 49, 315–333.

Harrison, J. M., Kreps, D. M., 1978. Speculative investor behavior in a stock market with heterogeneous expectations*. The Quarterly Journal of Economics 92, 323–336.

Innes, R. D., 1990. Limited liability and incentive contracting with ex-ante action choices. Journal of Economic Theory 52, 45–67.

Kreps, D., 1990. A Course in Microeconomic theory. Princeton University Press.

Larcker, D., Tayan, B., 2023. Mega grants: Why would a board approve nine-figure ceo pay. Stanford CGI Closer Look Series .

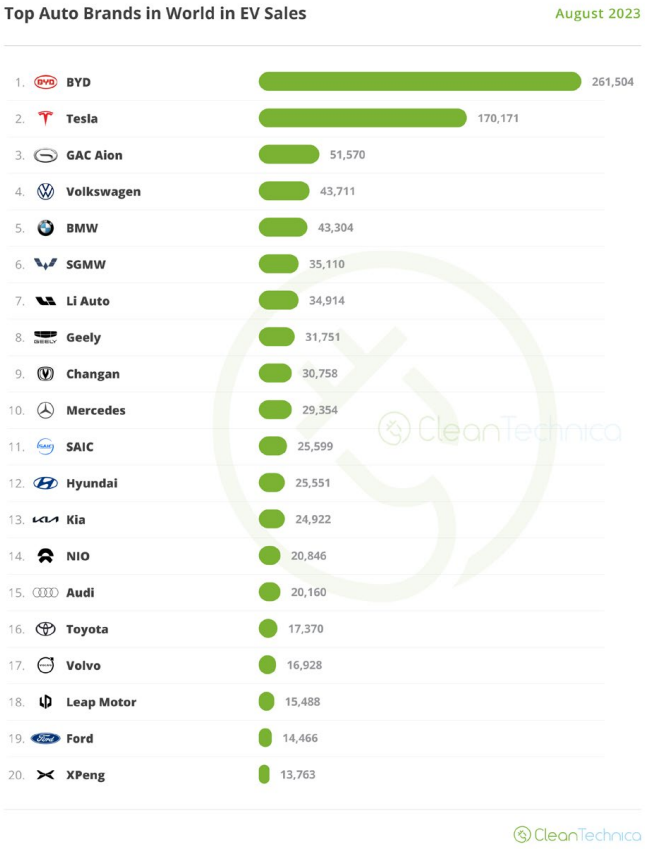
Marinovic, I., Bertomeu, J., Milone, M., 2024. Explaining the Compensation of Superstar CEOs: The Case of Elon Musk. Working paper .

Oyer, P., 2004. Why do firms use incentives that have no incentive effects? The Journal of Finance 59, 1619–1650.

Taylor, L. A., 2013. CEO wage dynamics: Estimates from a learning model. Journal of Financial Economics 108, 79–98.

Appendix A: Examples of winner-take-all markets

1) Market shares of EV brands in August 2023; 2) Stock returns of EV and GPU firms



http://cleantechnica.com/2023/10/07/world-ev-sales-now-equal-18-of-world-auto-sales/

https://finance.google.com





Figure 1: The adverse selection problem with two types $0 < t_L < t_H < 1$ and outcomes $\{\varepsilon, 1\}$ and $0 < \varepsilon < u < 1$

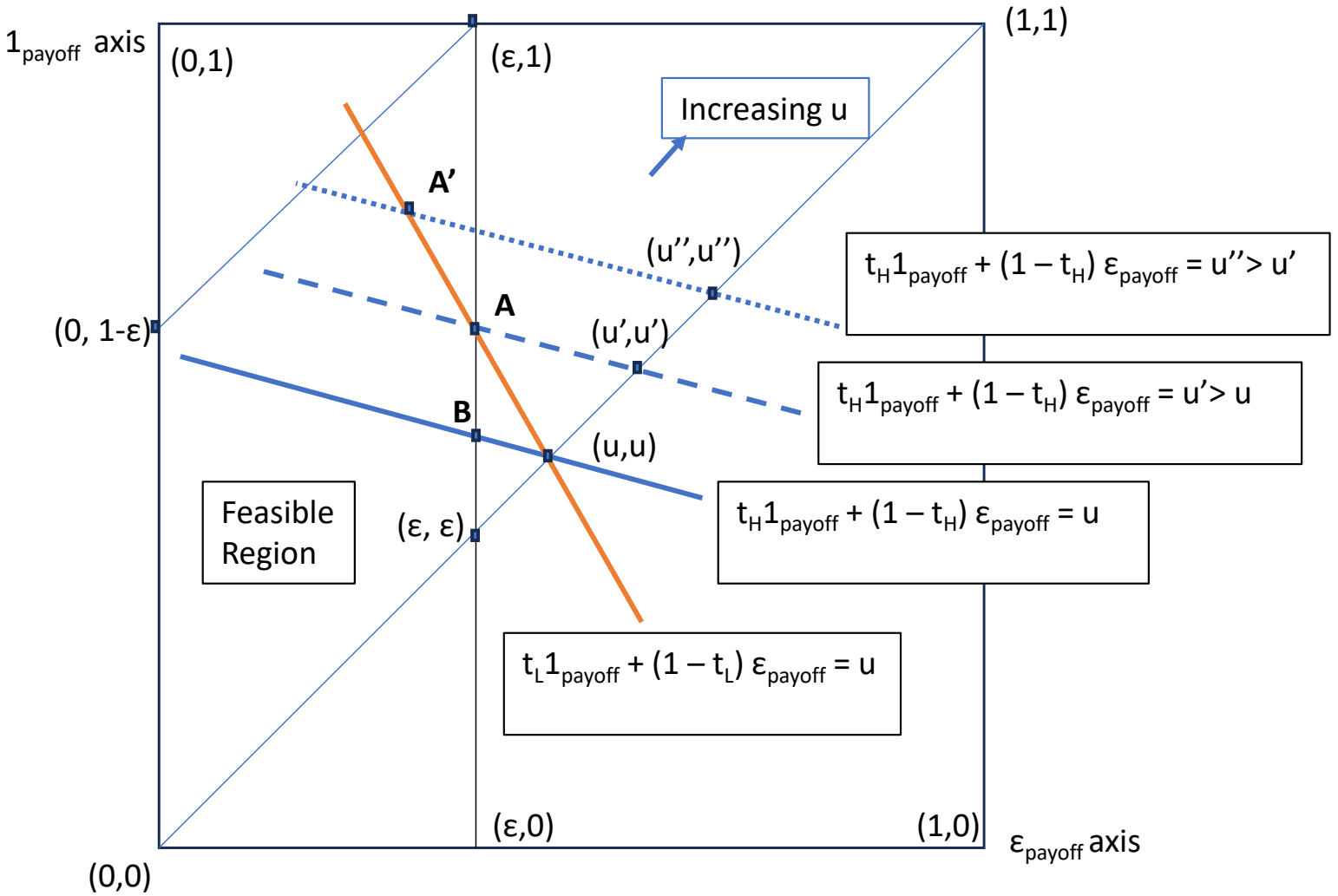


Figure 2: The adverse selection problem with two types $0 < t_L < t_H < 1$ and outcomes $\{\varepsilon', 1\}$ and $0 < \varepsilon' < \varepsilon < u < 1$

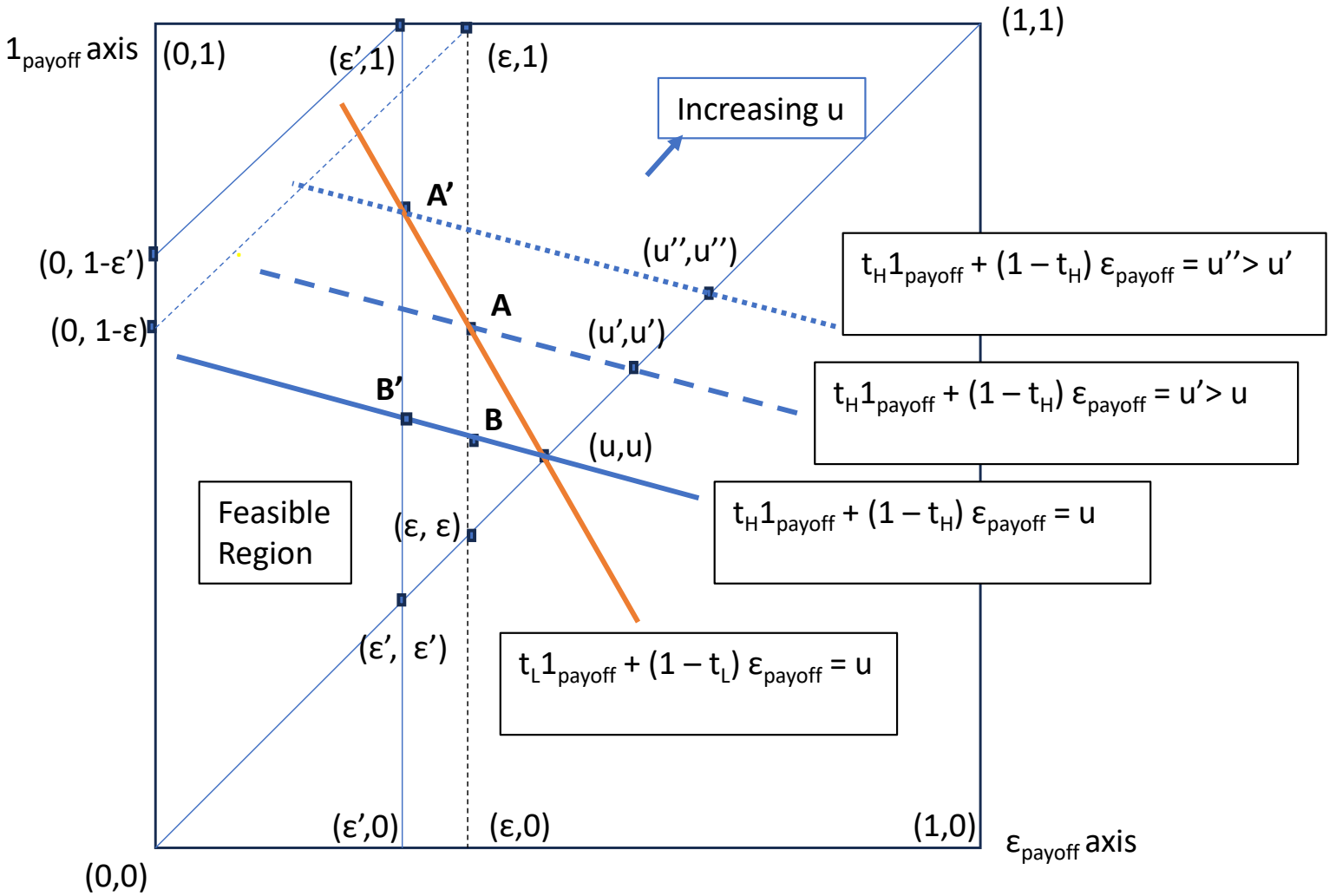


Figure 3: The adverse selection problem with two types $0 < t_L < t_H < 1$ and outcomes $\{\varepsilon, 2\}$ and $0 < \varepsilon < u < 2$

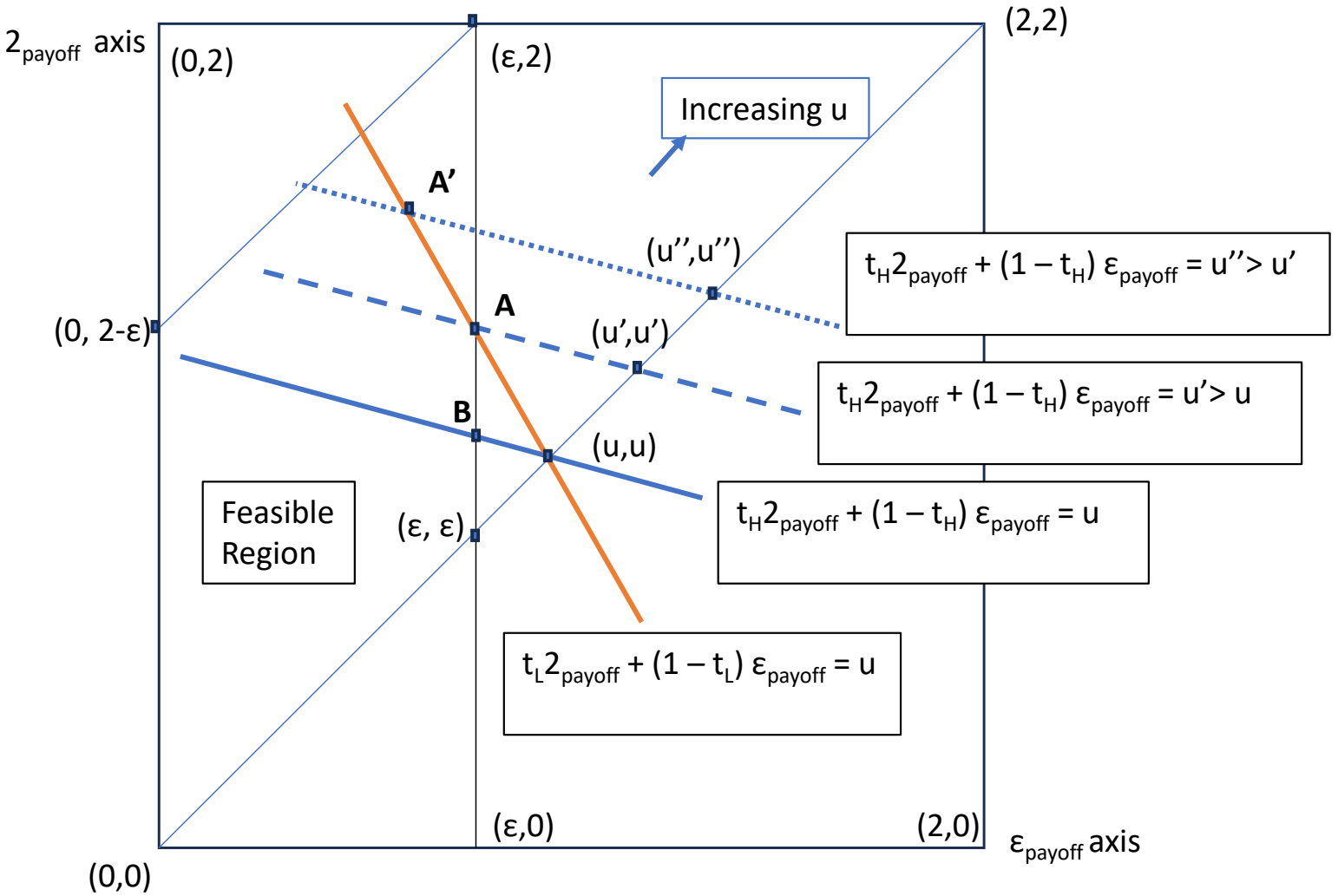


Figure 4: Expected [Gross Payoffs – Wages] in a Two-period Model with Two-period Enforceable Contracts

p_L is the ex-ante probability of a low type CEO and t_L, t_H are the low type and high type CEO's probabilities of winning 1 over ε. u is the reservation wage for both type CEOs and $x = \frac{u - \varepsilon(1-t_L)}{t_L}$. Probabilities are updated for the second period using Bayes rule.

$$p_\varepsilon^0 = p_L(1-t_L) + (1-p_L)(1-t_H)$$

$$p_\varepsilon^\varepsilon = p_L^\varepsilon(1-t_L) + (1-p_L^\varepsilon)(1-t_H)$$

where

$$p_L^\varepsilon = \frac{p_L(1-t_L)}{p_L(1-t_L) + (1-p_L)(1-t_H)}$$

Note: $p_L^\varepsilon > p_L, p_\varepsilon^\varepsilon > p_\varepsilon^0$
Because $t_H > t_L$

$$p_\varepsilon^1 = p_L^1(1-t_L) + (1-p_L^1)(1-t_H)$$

where

$$p_L^1 = \frac{p_L t_L}{p_L t_L + (1-pL)t_H}$$

Note: $p_L^1 < p_L, p_\varepsilon^1 < p_\varepsilon^0$
Because $t_H > t_L$

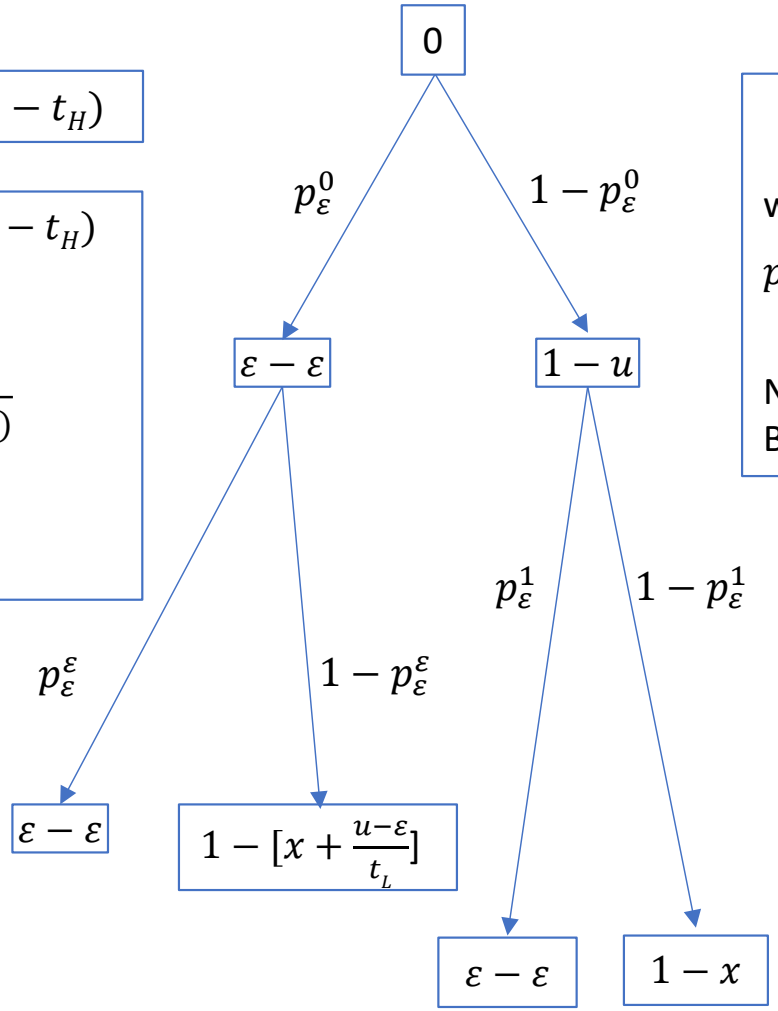


Figure 5: Expected [Gross Payoffs – Wages] in a Two-period Model with Period-by-period Contracts

p_L is the ex-ante probability of a low type CEO and t_L, t_H are the low type and high type CEO's probabilities of winning 1 over ε. u is the reservation wage for both type CEOs and $x = \frac{u - \varepsilon(1-t_L)}{t_L}$. Probabilities are updated for the second period using Bayes rule.

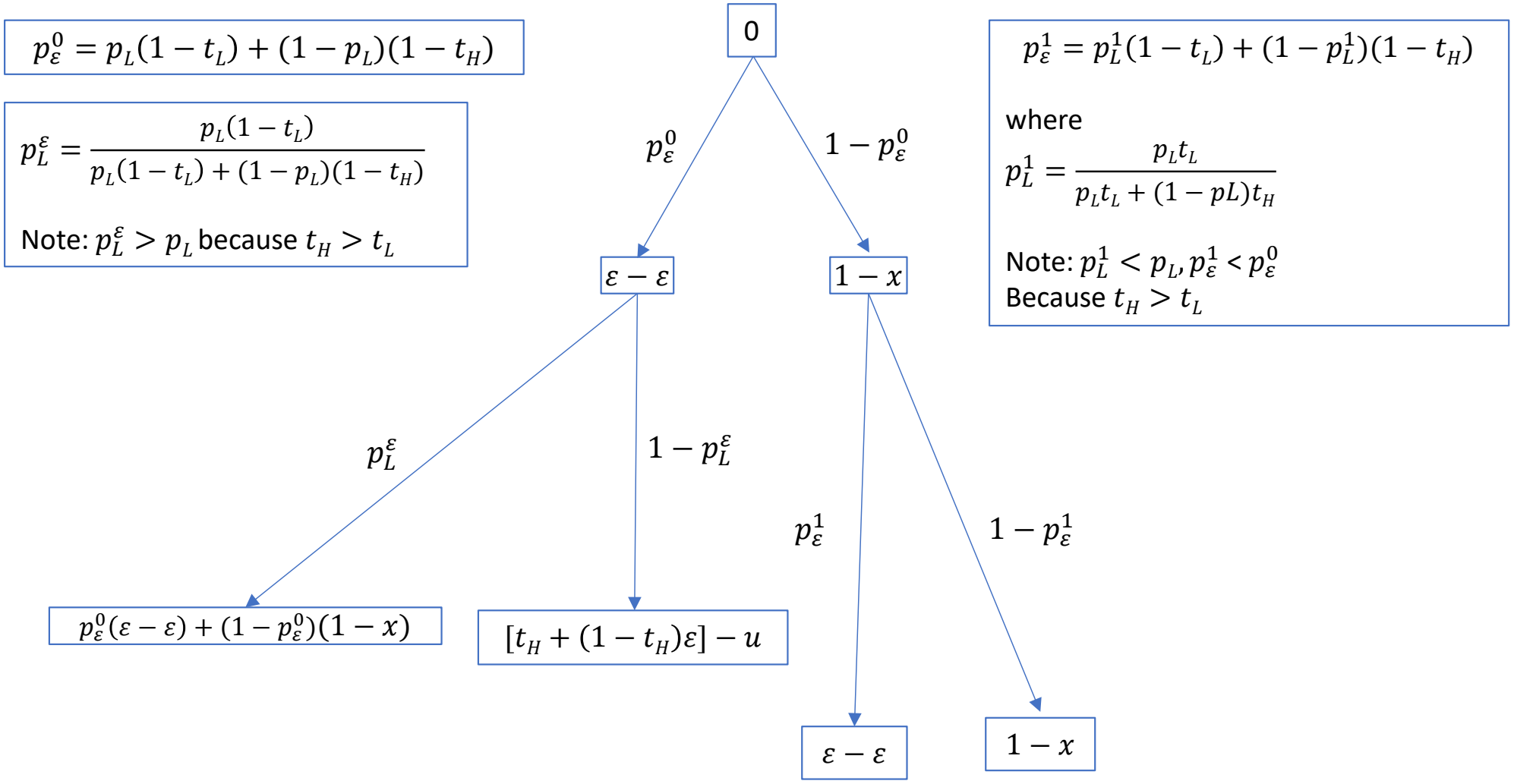


Figure 6: The adverse selection problem with three types $0 < t_L < t_M < t_H < 1$ and outcomes $\{\varepsilon, 1\}$ and $0 < \varepsilon < u < 1$

