7-1-2014

Games with Money and Status: How Best to Incentivize Work

Pradeep Dubey

John Geanakoplos

GAMES WITH MONEY AND STATUS: HOW BEST TO INCENTIVIZE WORK

By

Pradeep Dubey and John Geanakoplos

July 2014

COWLES FOUNDATION DISCUSSION PAPER NO. 1954

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Games with Money and Status: How Best to Incentivize Work*

Pradeep Dubey[†] and John Geanakoplos[‡]

9 July 2014

Abstract

Status is greatly valued in the real world, yet it has not received much attention from economic theorists. We examine how the owner of a firm can best combine money and status together to get his employees to work hard for the least total cost. We find that he should motivate workers of low skill mostly by status and high skill mostly by money. Moreover, he should do so by using a small number of titles and wage levels. This often results in star wages to the elite performers.

Keywords: Status, Incentives, Wages

JEL Classification: C70, I20, I30, I33

1 Introduction

Man is moved by the desire for status. Kings wage war for glory, soldiers give their lives for honor, and gangsters take lives for respect. Donors give more when their contributions are publicly recognized, and professors write more when they think it will bring them prestige. Children strive for excellence to win praise, and students study harder to get better grades. Athletes train longer to win medals and fame, corporate executives work harder to get promotions, and games are played competitively for the thrill of victory.

In many instances status brings money, suggesting that status is just instrumental to getting money. But as the last paragraph shows, status is also sought for its own sake. J.P. Morgan went so far as to say that money is just a way of keeping score, suggesting that money is often acquired in order to get status. Achilles became enraged when he was deprived of his booty, less because of its consumption value but more because of the signal it sent about his rank. Karna[1] viewed honor and status as

*This is a revision, with a slightly altered title, of the second half of Dubey–Geanakoplos (2005). The authors have recently adopted the convention of alternating the order of their names.

[†]Center for Game Theory in Economics, SUNY, Stony Brook and Cowles Foundation, Yale University

[‡]Cowles Foundation, Yale University, and Santa Fe Institute

[1]the tragic hero of the epic Mahabharata.

paramount, donating much of his wealth and refusing to switch to the winning side in battle, all for the glory of his reputation. Marlon Brando in "On the Waterfront" laments that in his youthful boxing days he took a dive for the money when he could instead have fought to win glory: "I coulda been a contender, I coulda been somebody".

Most societies have used status to motivate their citizens. The ancient Greeks allocated honorific prizes to the best playwright, the best painter, and the best of the Achaeans. The French bury their heroes in the Pantheon. The English bestow knighthood.

The flip side of honor, indeed its negative, is shame or "losing face". The thrill of victory and the agony of defeat both create incentive for performance. Admiral Nelson's exhortation "England expects every many to do his duty" is at once a promise of honor to those who fight and shame to those who run away.

The question is, what is the most efficacious way to allocate status? And how should status incentives and monetary incentives be combined? Should a planner who wishes to motivate everyone to work with the least expenditure of money deploy titles to lower all the wages, or is it better for him to actually raise some wages while lowering others? Should he provide status incentives to performers at the top or at the bottom? Should he focus on honor or on shame?

We take the essence of status to be the ranking of people created by titles.[2] It is immediately evident that the planner should use status as much as possible, since titles cost him nothing to bestow, and he should use money only for the residual incentive. Thus status is prior to money. It undoubtedly played a critical social role before money was invented, and even in modern times, money payment need not be contemplated until status incentives are exhausted.

Yet we shall argue that in spite of their costlessness, it is optimal for the planner who wants to minimize money expenditures to award very few titles. Even though there may be numerous levels of performance, it would be best for him not to discriminate among them too finely. He should partition them into a few broad categories and award both titles and wages based on those categories alone. The optimal wage schedule is a step function, not the steadily increasing curve that would be generated by a piece rate. The inevitable consequence is that at the cut points of the categories, a small increase in performance will lead to a higher title and a big jump in wage. We also show that the partition of performance levels into cells corresponding to different titles or wages is stable, depending only on productivity, but not the disutility of effort or the number of employees. Titles should not change over the business cycle, even if wage levels do.

One reason for broad categories is that often times higher performance is achieved by luck or ability rather than effort; steadily increasing rewards for increasing outputs can reduce effort.[3] A second reason is that broad categories generate higher status

[2] We shall abstract away from the organizational function of titles, and focus only on the status incentives created by titles.

[3] It would be naive to assume that higher output is always a sign of higher skill and better luck and more effort. This is particularly true when effort must be devoted to mastering multiple tasks. The low effort employee only gets to the first task; the high effort employee gets to many tasks,

incentives for employees with lower ability by enabling those who work hard to come equal in title with workers of higher ability, a little like handicaps do. (The difference is that handicaps discriminate between people, giving a boost say to minority candidates, whereas broad categories are anonymous).

We argue further that the planner should incentivize high performers mostly by money and low performers mostly by status. Indeed we describe circumstances under which a planner would lower his total wage bill by exorbitantly raising the wages of a tiny elite of top performers while using titles to motivate the others.

This sounds counterintuitive, but the reason is that a wage increase for middling performance forces a higher wage for all superior performance, because nobody will exert effort to strive for superior performance if he can already get a comparable wage for middling performance. Paying an astronomical wage for elite performance can thus be less costly than paying a high wage for middling performance.

Finally, we argue that when wages can be kept secret, the planner should further reduce the number of titles and further concentrate the status incentives on the lowest performers. We describe circumstances with secret wages under which it is optimal to award just two titles (member vs non-member) and in which all the status incentive goes to the very lowest ability group.

In order to make the arguments precise, we build a mathematical approach to status involving the owner of a firm and his employees, who may have differential abilities. Employees can either work or shirk. Their output is increasing *on average* with effort and ability, but the actual output depends also on chance. The owner cannot tell in general whether a particular output was caused by effort, ability, or luck. The owner observes the output of each employee, and on that basis alone awards him both a title and a wage.[4] We suppose that the rewards are non-discriminatory: employees who produce the same output get the same title and wage. Furthermore, rewards are merit based in that a unilateral increase in the ouput of any employee cannot lower his reward: his wage cannot go down, and if his title earlier outranked another's it continues to do so.

Status utility has been discussed by many authors from various points of view.[5]

---

but then is slightly less good at the first task. Imagine a student who studies just one problem and can get it right 80% of the time, while the high effort student studies both semesters and can get each of two problems right 60% of the time. The shirker has probabilities (.2,.8,0) of getting 0,1, or 2 correct answers, while the high effort student has probabilities (.16.,48,.36). (In the language we shall shortly introduce, the worker stochastically dominates the shirker, but does not uniformly stochastically dominate him.) The optimal title partition or wage schedule will turn out to be {{0,1},{2}} in which the reward comes only for getting both answers right.

[4] The owner could have considered other more general reward schedules, where the wage and title of one worker depended on the output of the others. The most common would be a relative reward schedule in which the top r_1 outputs are given the highest award, the next r_2 the second highest and so on. We prove in Section 4 that our absolute scheme is better for the owner than any relative scheme. But there are still more general schemes we do not investigate.

[5] There is a large literature on status, starting with Veblen (1899) who famously introduced conspicuous consumption, i.e., the idea that people strive to consume more than others partly for the sake of higher status. See Auriol-Renault (2008) for a recent survey of the literature, including different approaches to the modeling of status

A major distinction turns on cardinal versus ordinal utility.[6] Our model is in the ordinal tradition. We suppose that every employee gains σ utiles for each person he outranks and loses σ utiles for each person who outranks him.[7] There is some recent evidence that such utilities are quite prevalent in reality.[8]

To sum up, if the N employees are paid wages $w = (w_1, ..., w_N)$, and awarded titles $t = (t_1, ..., t_N)$, each employee n obtains utility

$$u^n(w, t) = u_n(w_n) + \sigma_n[\#\{j : t_n > t_j\} - \#\{j : t_j > t_n\}]$$

We suppose that utility u for money is increasing and concave. Our results are most powerful in the risk neutral case when u is linear, for when there is diminishing marginal utility, it is very difficult to further motivate highly paid employees with still more money.

Wages could be public or secret. When wages are public, as in state universities and government offices in the United States and many other countries, they enable or force people to compare themselves with each other. Public wages in effect create titles along with monetary compensation. In this case an output that gets a higher wage necessarily gets a higher title, but we do allow titles to go up across outputs that are given the same wage.[9]

Sometimes wages can be secret, as happens in many private institutions, such as Yale University and the Santa Fe Institute, where there is an unspoken rule that employees do not discuss their salaries. This gives the owner more flexibility, since he can pick wages and titles completely independently of each other.

To ease the analysis we completely solve two polar extremes. In Section 2 we consider employees of homogeneous (i.e. ex ante identical) abilities, and in Section

---

[6] One strand of the literaure adopts a cardinal approach which makes utility depend on the difference between an individual's wage/consumption and others' consumption (see, e.g., Duesenberry (1949), Pollak (1976), Fehr–Schmidt (1999), and Dubey-Geanakoplos-Haimanko (2013)). The ordinal approach makes utility depend on the individual's rank in the distribution of consumption (see, e.g., Frank (1985), Robson (1992), Direr (2001), and Hopkins–Kornienko (2004)). Our model of status is in the ordinal tradition. This should be contrasted with the purely instrumental role status might play, for instance when higher consumption signals higher wealth and hence eligibility as a marriage partner (see e.g., Cole–Mailath–Postlewaite (1992, 1995, 1998) and Corneo–Jeanne (1998)).

[7] We introduced this utility function for status in Dubey-Geanakoplos (2005, 2010). In an interesting paper on "Contests for Status" with incomplete information, Moldevanu, Sela, and Shi (2007) also used the same utility function.

An alternative would be to suppose that status utility comes from being the top dog, or more generally, is increasing and stricly convex in the number of people an employee outranks. Strict convexity might arise if higher rank gives higher visibility to the outside world; for example, only the CEO might enjoy media attention outside the firm. In this paper we are ruling this out, imagining a closed world in which status utility is derived from the acknowledgment of superior rank by the other employees; the owner alone is the public face of the company. However our framework can accommodate such non-linear status utilities, and in particular the existence of optimal wage-title schedules is not compromised, as we shall show in a sequel paper. The planner will adapt his rewards depending on how his employees perceive status.

[8] See for example the article "Does Wage Rank Affect Employees Well-being?" by Brown, Gardner, Oswald, and Qian (2008).

[9] We easily accomodate the case where wages alone signify title. If two titles are given the same wage, we could pay the higher title infinitesimally more if we wanted to identify titles with wages.

3 we consider employees with disparate abilities.[10] In both cases we solve for the optimal title schedule when wages cannot be paid, the optimal wage schedule when titles cannot be awarded, and finally the optimal wage-title schedule. For homogeneous workers it turns out that the optimal title schedule remains the same after it becomes possible to pay public wages as well (and thus there is nothing to be gained by keeping wages secret), but for disparate workers, the optimal title schedule must be adjusted once wages are introduced (and thus the owner can exploit secret wages). In both the homogeneous and disparate cases, increasing performance near the low end brings more status but not much more wages, while at the high end, increasing performance can bring huge additional money bonuses. In both cases the star like quality of the wage schedule is increased as status becomes more important.

The general case of employees with overlapping abilities — which lies in between our extremes — is no doubt important, and our framework makes it clear that optimal wage-title schedules exist in this case as well, though their precise structure is not investigated here.

Furthermore, for the most part we assume complete information, i.e. each employee knows not only his own ability, but also the population distribution of abilities of his rivals. In Section 5 we show that our results remain essentially intact with incomplete information where each employee has a probability distribution on the abilities of each of his rivals but does not know their actual realizations.

2 Homogeneous Employees

We first consider the case of N homogeneous, i.e. *ex ante* identical, employees. For simplicity suppose that the possible outputs lie in a finite[11] set $Q \subset \mathbb{R}_+$, with maximum $x_{\max}$ and minimum $x_{\min}$. If any employee works, his output is a random variable X with density f on Q, and if he shirks it is a random variable Y with density g on Q. For any subset $A \subset Q$, let $f(A) = \sum_{x \in A} f(x)$ and $g(A) = \sum_{x \in A} g(x)$. We suppose that the output of each employee is statistically independent of the others' outputs regardless of their effort levels.[12] The disutility for switching from shirk to work is denoted by $d_n > 0$ for all employees n. Since employees are ex ante identical, we take $\sigma_n = \sigma, u_n = u, d_n = d$ for all employees n.

We make the productivity assumption that the worker on average produces more

[10] These were also the center of attention in Dubey-Geanakoplos (2010), in the scenario where wages were not present, and students were rewarded solely by titles/grades based on their exam scores.

[11] If Q is a compact interval, we can approximate it by a fine finite grid and then use a limiting argument to derive the analogous result for a continuum of outputs.

[12] We assumed independence for ease of exposition. Our analysis goes through with a weaker hypothesis consisting of two parts. (a) If N-1 employees work and one shirks, then the shirker's performance g is independent of the workers' performance, each of which is given by f (which need not be independent from each other). (b) If they all work, their outputs (which can be distributed according to $h \neq f$) are ex ante symmetric in the following precise sense: consider an elementary event in which every person is assigned an output, and another elementary event obtained by permuting the names; then the two elementary events should have the same probability.

output than the shirker:

$$\sum_{x \in Q} f(x)x > \sum_{x \in Q} g(x)x$$

Without loss of generality, for $x \in Q$, either $f(x) > 0$ or $g(x) > 0$.

We begin by studying absolute reward schedules $(\mathcal{P}, w)$, where $\mathcal{P}$ is the titles partition and w is the wage schedule. Later, in Section 5, we shall show that relative reward schedules are inferior.

In the absolute schedule, there is a partition $\mathcal{P}$ of Q into consecutive cells (intervals) corresponding to increasing titles; $\mathcal{P}(x)$ denotes the cell of $\mathcal{P}$ in which x lies. There is also a *wage schedule* given by a weakly monotonic function $w : Q \to [w_{\min}, \infty)$, mapping outputs to wages above some stipulated minimum $w_{\min} \geq 0$. If wages are public, then w must be measurable with respect to $\mathcal{P}$, that is constant on each cell of $\mathcal{P}$, meaning that outputs which get the same title cannot get different wages. The collection of wage schedules that are measurable with respect to $\mathcal{P}$ is denoted $\mathcal{W}(\mathcal{P})$. Let Π be the (finite) set of partitions of Q into consecutive cells, denoting all possible ways $\mathcal{P}$ of allocating titles. Let $\mathcal{W} \equiv \cup\{\mathcal{W}(\mathcal{P}) : P \in \mathcal{P}\}$ denote the collection of all wage schedules, i.e. the set of all weakly monotonic functions $w : Q \to [w_{\min}, \infty)$.

Let $I(\mathcal{P})$ denote the status incentive generated by $\mathcal{P}$ when $\sigma_n = \sigma = 1$, i.e. the increase in payoff of an employee when he switches from shirk to work, assuming that all others are working, ignoring money altogether and considering only titles. (The status incentive for arbitrary $\sigma \geq 0$ is then $\sigma I(\mathcal{P})$.) Clearly his status payoff is 0 when he works, since he comes ahead of his ex ante identical competitors as often as he comes behind. Therefore, recalling that performances are independent, his status incentive is simply $N-1$ times the negative of his status payoff when he shirks and faces exactly one competitor who works.

$$I(\mathcal{P}) = (N-1)[\sum_{\{x \in Q, y \in Q : \mathcal{P}(y) < \mathcal{P}(x)\}} f(x)g(y) - \sum_{\{x \in Q, y \in Q : \mathcal{P}(y) > \mathcal{P}(x)\}} f(x)g(y)]$$

Let I^S denote the maximum incentive to work that can be generated by status alone, i.e.

$$I^S = \max_{\mathcal{P} \in \Pi} I(\mathcal{P})$$

As we shall see in Section 3.1, it follows from our productivity assumption that $I^S > 0$.

Similarly, given a wage schedule $w \in \mathcal{W}$, we can define the wage incentive

$$I(w) = \sum_{x \in Q} [f(x) - g(x)]u(w(x))$$

and the maximum incentive to work that can be created by money alone

$$I^M = \sup_{w \in \mathcal{W}} I(w)$$

A popular but naive wage schedule is the piece rate in which $w(x) = \lambda x$ for some fixed scalar $\lambda > 0$. If u is linear, the piece rate creates a positive incentive to work (on account of the productivity assumption). By increasing λ, the incentive can be increased to any level desired.

The piece rate is the first wage schedule that comes to mind, because we are so used to competitive markets. If the worker could sell pieces of his ouput to different competing firms, then a market price would be established for his output, corresponding to the piece rate. But this logic does not apply to our setting. The worker can choose among different firms (modeled by his participation constraint, which we introduce at the end of Section 2) but having made the choice he becomes an employee and must give his entire output to the owner. As long as they remain with the firm, the employeess are paid according to the policy set by the owner. As we shall see, the owner will not want to set a piece rate.

More generally, if u is concave, we can define a wage schedule $\tilde{w}(x)$ so that $u(\tilde{w}(x))$ is linear in x. Indeed, fix $0 < \lambda < (\sup_{w \in \mathbb{R}} u(w) - u(w_{\min}))/(x_{\max} - x_{\min})$. Let $\tilde{w}(x) = u^{-1}(u(w_{\min}) + \lambda(x - x_{\min}))$ for all $x \in Q$.

From the productivity assumption, this wage schedule gives positive incentive to work. Hence $I^M > 0$. Furthermore, if $u(w) \to \infty$ as $w \to \infty$, then we can take λ arbitrarily large and the incentive to work becomes arbitrarily large, hence $I^M = \infty$.

The piece rate schedule and its adapted version for concave u is simple but not economical. We shall shortly derive a much less costly wage schedule that gives the same incentive to work.

Consider the general problem of selecting the optimal reward (i.e. wage-title) schedule, taking into account both the status incentives of titles and the consumption incentives of wages.[13]

$$\min_{w,\mathcal{P}} \sum_{x \in Q} f(x)w(x)$$

$$\text{s.t.} \begin{cases} \sigma I(\mathcal{P}) + I(w) \geq d \\ \mathcal{P} \in \Pi, w \in \mathcal{W} \\ w \in \mathcal{W}(\mathcal{P}) \text{ if wages are public} \end{cases}$$

The optimal absolute reward schedule turns out to be quite simple for ex ante identical employees. To find it we break the analysis into two parts. Throughout we keep f, g fixed, and examine the solution as d varies.

In Section 3.1 we solve the pure titles problem. We ask how the owner could best use titles to motivate his employees to work, without handing out any money at all, i.e., we characterize all partitions $\mathcal{P}$ such that $I(\mathcal{P}) = I^S$. It turns out that these partitions can be identified by the easily checked "inside and outside" conditions. Furthermore, we show that they form a (complete) sublattice of the lattice of all partitions, with maximal element $\mathcal{P}^*$ and minimal element $\mathcal{P}_*$. It is evident that this sublattice is the set of all solutions to the minimization problem above, for each

[13] We implicitly assume that the output of the worker is so valuable to the employer relative to the wages he needs to pay in order to get them to work, that he deems it optimal to incentivize everyone to work.

$d \leq \sigma I^S$; and that there are no solutions without wages when $d > \sigma I^S$. Finally, for generic f, g the sublattice is a singleton.

A typical property of optimal partitions is that they are coarse, clumping many outputs into the same cell: there are far fewer titles than outputs. In fact, only in the very special scenario where f uniformly stochastically dominates g do we get as many titles as outputs in an optimal solution.

In Section 3.2 we solve the pure wage problem. We ask how the owner should best choose a wage schedule when his employees derive no status utility from titles, i.e., which $w_d \in \mathcal{W}$ solve the minimization problem above when $\sigma = 0$? This is the classical pure wage problem. It has a solution for all $d \leq I^M$. We show that the solution to the pure wage problem is connected to the pure titles problem. Every solution w_d of the pure wage problem is measurable with respect to the finest partition $\mathcal{P}^*$ that solves the pure titles problem.

Thus the optimal wage schedule gives even fewer wage levels than there are titles. This is a far cry from a piece rate schedule where wages strictly increase with each output. The optimal wage schedule is a step function with broad steps. At the jump points, a small increase in output is rewarded with a huge increase in wage.

This is even more starkly true when employees are risk neutral. In this case, w_d can be taken to be a "trigger wage" or "star wage": outputs below a threshold q^* are paid $w_{\min}$ and those above q^* get a bonus $w_{\min} + B$, i.e. we get just two wage levels no matter how many outputs. Moreover, for generic f and g, the optimal wage must be of the trigger form. In the special case where f uniformly stochastically dominates g, the bonus is given only to the very top element of Q and the trigger wage is really a "star's" wage.

When titles and wages are combined, the optimal wage-title schedule is achieved simply by the superposition of the pure titles solution and the pure wage solution just described, and this is so whether wages are public or secret.[14] Since the owner is trying to minimize the wage bill, he will first try to see how far he can go via titles alone before putting up money to motivate his workers. We find that for all $0 < d \leq \sigma I^S$, it is optimal to choose any partition $\mathcal{P}$ that solved the pure titles problem. No wages are necessary. For $\sigma I^S < d \leq \sigma I^S + I^M$, the same titles partition $\mathcal{P}$ can be accompanied by the wage schedule $w_{d-\sigma I^S}$ that solved the pure wage problem for disutility $d - \sigma I^S$. It makes no difference whether wages are public or secret. For $d > \sigma I^S + I^M$, no solution is possible.

When u is concave and satisfies increasing relative risk aversion, we show that as σ increases the optimal wage schedule becomes more and more trigger like. Thus as society becomes more status conscious, wages become more unequal. We prove these results over the next three sections.

2.1 Titles Alone

We examine the incentive to work created by titles alone, and ask which $\mathcal{P}$ maximizes $I(\mathcal{P})$. Such a $\mathcal{P}$ is optimal in the sense that if any other title scheme $\mathcal{P}' \in \Pi$ gets

[14] This is surprising because the pure titles partition implements work as a Nash equilibrium in the N-person game, whereas the optimal pure wage schedule implements work in a one-person problem.

employees to work via status incentive alone, so will $\mathcal{P}$. To characterize optimal $\mathcal{P}$, we need to recall two notions of stochastic dominance.

2.1.1 Stochastic Dominance

Definition: Let X and Y be independent random variables which take on values in a finite totally ordered set Z.[15] We say that X *(stochastically) dominates* Y on the interval $[a,b] \subset Z$ if $\Pr(X \in [a,b]) \Pr(Y \in [a,b]) = 0$, or

$$\Pr(X \in [\theta,b] | X \in [a,b]) - \Pr(Y \in [\theta,b] | Y \in [a,b]) \geq 0,$$

for all $\theta \in (a,b]$. In this case we write

$$X \succsim Y \text{ on } [a,b].$$

In words, this means that no matter at what point θ we cut the interval $[a,b]$, conditional on both X and Y lying in $[a,b]$, X is at least as likely to lie in the upper segment as Y. If every cut on $[a,b]$ gives a strict inequality, then we say that X *strictly dominates* Y and we denote it by $X \succ Y$ on $[a,b]$.

With this definition in hand, we can show that $I^S > 0$. Let $X \sim f$ and $Y \sim g$ denote the stochastic outputs of the worker and the shirker. First note that Y cannot stochastically dominate X on Q, otherwise, by the "dominance increases expectation lemma" in the appendix, $\sum_{x\in Q} f(x)x \leq \sum_{x\in Q} g(x)x$, contradicting the productivity assumption. Therefore there exists a $\theta \in Q$ such that $\Pr(X \geq \theta) - \Pr(Y \geq \theta) > 0$. Partition Q into two titles: let all outputs less than θ be accorded the low title, and all outputs θ and above be given the high title. This clearly generates positive status incentive, hence $I^S > 0$.

A moment's thought will convince the reader that $X \succsim Y$ on $[a,b]$ if and only if whenever $[a,b] = L \cup R$ is divided into two disjoint intervals, the left interval L lying below the right interval R, then

$$\frac{P(X \in L)}{P(Y \in L)} \leq \frac{P(X \in R)}{P(Y \in R)}$$

This is obviously equivalent to the "**betweenness**" property

$$\frac{P(X \in L)}{P(Y \in L)} \leq \frac{P(X \in L \cup R)}{P(Y \in L \cup R)} \leq \frac{P(X \in R)}{P(Y \in R)}$$

In the case of strict domination, these inequalities will be strict.

It will be useful to consider a strengthened form of domination.

[15] As before, we assume without loss of generality that for each $a \in Z$, either $\Pr(X = a) > 0$ or $\Pr(Y = a) > 0$.

Definition: We say that X *uniformly dominates* Y on the interval $[c,d] \subset Z$ if X dominates Y on every subinterval $[a,b] \subset [c,d]$. In this case we write $X \succsim_U Y$ on $[c,d]$.

Uniform domination $X \succsim_U Y$ on $[c,d]$ can easily be seen to be equivalent to the condition that whenever $a < b$ are consecutive elements of $[c,d]$ then

$$\frac{P(X=a)}{P(Y=a)} \leq \frac{P(X=b)}{P(Y=b)}$$

In case $[c,d]$ consists of two elements, domination and uniform domination are the same. But with three elements or more, uniform domination is a strictly stronger requirement. *Strict uniform domination,* denoted $X \succ_U Y$, is defined just like uniform domination, but with strict inequalities throughout.

2.1.2 The Optimal Titles Partition

To create the best incentives for work, we need to lower the shirker's payoff as much as possible. Thus it stands to reason that we should mask performance in regions of ouput where the shirker is better than the worker, by awarding the same title throughout; and award titles for superior performance across regions where the worker is likely to do better. These are reflected in the inside and outside domination conditions of our first Theorem.

Theorem 1 (Inside-Outside Condition): *Let X, Y denote the random output of the employee when he works, shirks. Then $I(\mathcal{P})$ is maximized over Π at $\bar{P}$ if and only if*

(i) *Inside Domination*: $Y \succeq X$ *on each cell of* $\bar{P}$

(ii) *Outside Uniform Domination*: $X \succeq_U Y$ *across the cells of the ordered set* $\bar{P}$

(The proof of Theorem 1 and all other omitted proofs can be found in the Appendix.)

Given that X and Y both lie somewhere in two consecutive partition cells, condition (ii) says that X is more likely than Y to lie in the upper cell. But given that X and Y lie in the same cell, condition (i) says that it is more likely that Y is to the right of any cut.

One might have thought that since titles create status incentive and are free to bestow, the owner should hand out as many titles as he can. However, an optimal partition often involves masked cells. Indeed we have

Lemma 1 (Coarse Partition): *Suppose that x, y are consecutive outputs and that*

$$\frac{f(x)}{g(x)} > \frac{f(y)}{g(y)}$$

Then x and y must be in the same cell of any optimal titles partition. In particular, an optimal titles partition can be perfectly fine (and hand out as many titles as there are outputs) only if $X \succeq_U Y$.

In the special case where the worker uniformly dominates the shirker, we do get the opposite.

Lemma 2 (Fine Partition): *Suppose $X \succeq_U Y$. Then the perfectly fine partition is optimal. If $X \succ_U Y$, then the perfectly fine partition is the unique optimum.*

For examples and discussion see Dubey-Geanakoplos (2005, 2010).

The set of all optimal partitions turns out to be a lattice. Recall that the join of two partitions is the coarsest partition that refines them both, and that the meet is the finest partition that they both refine. In our case of interval partitions of totally ordered finite sets, the partitions are easily identified with their *cuts*, i.e. the boundary points of the intervals.[16] Then the join of two partitions is defined by the union of their cuts, and the meet is defined by the intersection of the cuts.

Theorem 2 (Lattice Structure): *The optimal partitions form a sublattice, under the join and meet operations, of the lattice of all partitions. Thus there is a unique optimal title partition with the most titles, obtained by taking the join of all the optimal partitions; this maximal optimal partition is the unique optimal partition that displays strict inside domination on each of its cells. There is also a unique optimal title partition with the fewest titles, obtained by taking the meet of all the optimal partitions; this minimal optimal partition is the unique optimal partition that displays strict outside domination. Finally, the sublattice is complete, i.e. it includes all the elements of the lattice between the meet and the join. Indeed, each cell in any partition in the sublattice is the union of consecutive cells C from the maximal optimal partition across which $f(C)/g(C)$ is constant.*

According to Theorem 2, every optimal partition is obtained by consolidating some of the titles of the maximal optimal partition or equivalently by splitting some of the titles of the minimal optimal partition. For example, the fewest titles (in the minimal optimal title partition) might be general, colonel, major, captain, lieutenant, sargeant, corporal, private. The most titles (in the maximal optimal title partition) might be lieutetant general, major general, brigadier general, ..., private first class, private second class and so on.

The proof of Theorem 1 was given in a more general setting in Dubey-Geanakoplos (2005, 2010). For completeness, and because the proof is so much simpler and possibly more instructive in the finite output case considered in this paper, we present it in the Appendix. Theorem 2 is presented here for the first time.

It is worth noting that the lattice is usually a singleton.

[16] Formally speaking, a cut is defined by a pair ab where a is a last element of one interval and b is the first element of the next interval.

Lemma 3 (Unicity of the Lattice): *Regarding f and g as vectors in the finite dimensional set $\mathbb{R}^Q$, the optimal partiton is unique for (Lebesgue) almost all f, g.*

2.1.3 Computing the Minimal and Maximal Optimal Titles Partitions

We shall provide two algorithms for computing the minimal optimal partition $\mathcal{P}_*$. Once we have $\mathcal{P}_*$ it is straightforward to construct $\mathcal{P}^*$.

The First Algorithm for the Minimal Optimal Partition Start with the finest possible partition of Q into singleton cells. Trivially this partition satisfies the inside condition. Proceed inductively as follows.

Given any partition $(...D < C < B < A)$ satisfying the inside condition on each cell, starting from the right look at all pairs of consecutive cells, BA, CB, DC, etc. Take the first pair $\beta\alpha$ in the list for which

$$\frac{f(\beta)}{g(\beta)} \geq \frac{f(\alpha)}{g(\alpha)}$$

If no such pair can be found, then by the inside-outside condition of Theorem 1, we have an optimal partition; and, by the lattice structure of Theorem 2, it is the minimal optimal partition.

Otherwise, combine cells β and α into the bigger cell $\beta \cup \alpha$. By the merger lemma in the Appendix, this new partiton must also satisfy the inside condition on each of its cells. Iterate the process. Since Q is finite, the process must terminate. ■

The Second Algorithm for the Minimal Optimal Partition *Let*

$$\theta_1^* = argmax_{\theta \in Q} \frac{\sum_{x \geq \theta} f(x)}{\sum_{x \geq \theta} g(x)}$$

If there are multiple such maximizers, choose the smallest. Then define the rightmost cell as $C_1 = \{q \in Q : \theta_1^* \leq q\}$.

Given θ_k, define

$$\theta_{k+1}^* = argmax_{\theta < \theta_k} \frac{\sum_{\theta_k > x \geq \theta} f(x)}{\sum_{\theta_k > x \geq \theta} g(x)}$$

Again choose the lowest such maximizers in case there are ties. Then define $C_{k+1} = \{q \in Q : \theta_{k+1}^* \leq q < \theta_k^*\}$. This algorithm terminates in a partition after at most $|Q|$ steps.

It remains to check that the constructed partition satisfies the inside and outside conditions. But this is evident. If

$$\frac{f(C_{k+1})}{g(C_{k+1})} > \frac{f(C_k)}{g(C_k)}$$

then by betweenness, θ_{k+1} would have done better than θ_k in the kth maximization problem, a contradiction. This establishes the outside condition.

For any cut of C_k into $[\theta_k, \theta) \cup [\theta, \theta_{k-1})$ we must have

$$\frac{f([\theta_k, \theta))}{g([\theta_k, \theta))} \geq \frac{f([\theta, \theta_{k-1}))}{g([\theta, \theta_{k-1}))}$$

otherwise, by the betweenness property, θ would do better than θ_k in the kth maximization problem. This establishes the inside condition.■

Algorithm for the Maximal Optimal Partition Given any optimal partition $\mathcal{P}$, it is a simple matter to construct the maximal optimal partition $\mathcal{P}^*$. Simply look at all cuts of any cell C in $\mathcal{P}$. If the cut leaves f/g the same on both sides, make it. By the splitting lemma in the appendix, the new partition satisfies the inside-outside conditions. Continue iterating the process. By Theorem 2 (Lattice Structure) the algorithm can only stop at the maximal optimal partition.■

2.2 Wages Alone

Now we turn to the classical pure wage problem where employees do not care about status, but only about money, and hence must be motivated by wages alone. Surprisingly, we find that to minimize his total wage bill, the owner *must* always pick a wage schedule that is measurable with respect to the maximal optimal partition $\mathcal{P}^*$ for the pure titles problem, and *may* always pick it measurable with respect to the minimal optimal partition $\mathcal{P}_*$. In other words the pure wage solution never pays differently to outputs that are accorded the same title in the pure titles solution $\mathcal{P}^*$. In the pure wage problem people don't compare themselves with each other, and only think about what money can buy. In the pure titles problem they don't care about money, but only about how they rank against others. Nonetheless, the solutions to these diametrically opposed problems are in harmony. Indeed, our characterization of the optimal titles partition vastly simplifies the search for the optimal pure wage schedule.

From now on we shall make the not unrealistic assumption that the disutility of work is high enough that **no employee will work for status alone**

$$d > \sigma I^S$$

Theorem 3 (Compatibility of Pure Wages and Pure Titles): *Any solution* w of *the pure wage problem*

$$\min_{w \in \mathcal{W}} \sum_{x \in Q} f(x) w(x)$$
$$\text{s.t. } I(w) \equiv \sum_{x \in Q} [f(x) - g(x)] u(w(x)) \geq d$$

is measurable with respect to the maximal optimal titles partition $\mathcal{P}^*$*. Furthermore, there exists a solution that is measurable with respect to the minimal optimal titles partition* $\mathcal{P}_*$*.*

The next theorem shows that every optimal pure wage schedule pays the minimal wage for a nontrivial initial segment of outputs.

Theorem 4 (Minimum Wage with Risk Aversion) *Let the maximal optimal titles partition $\mathcal{P}^*$ consist of consecutive cells $C_1, ..., C_K$. Let $f(C_k) \leq g(C_k)$. (Clearly there must be one such k,* since $\sum f(C_k) = 1 = \sum g(C_k)$*). Then, for every solution w of the pure wage problem given above, $w(x) = w_{\min}$ for all $x \in C_1 \cup ... \cup C_k$.*

A surprising Corollary of Theorem 4 is that with strict risk aversion, there is a unique optimal pure wage schedule.

Corollary to Theorem 4 (Uniqueness with Strict Risk Aversion) *If u is strictly concave, then the pure wage problem has a unique solution.*

The next theorem shows that there is always an optimal pure wage schedule that is measurable with respect to the coarsest optimal pure title partition $\mathcal{P}_*$ and pays as before the minimum wage for an initial segment of cells, but is strictly increasing across all cells thereafter.

Theorem 5 (Wage Structure with Risk Aversion): *Let the minimal optimal titles partition $\mathcal{P}_*$ consist of consecutive cells $C_1, ..., C_L$. Let employee utility u be differentiable. Then there is an optimal pure wage schedule w that is measurable with respect to P_*, such that $w(x) = w_{\min}$ on a non-empty initial segment $x \in C_1 \cup ... \cup C_{\ell^*}$ and strictly increasing across cells to the right of C_{ℓ^*}. Furthermore*

$$\frac{u'(w_{\ell+1})}{u'(w_\ell)} = \frac{[f(C_\ell) - g(C_\ell)]/f(C_\ell)}{[f(C_{\ell+1}) - g(C_{\ell+1})]/f(C_{\ell+1})} < 1$$

for all $\ell^* < \ell < L$, and if $f(C_{\ell^*}) > 0$, then

$$\frac{u'(w_{\ell^*+1})}{u'(w_{\ell^*})} \geq \frac{[f(C_{\ell^*}) - g(C_{\ell^*})]/f(C_{\ell^*})}{[f(C_{\ell^*+1}) - g(C_{\ell^*+1})]/f(C_{\ell^*+1})}$$

In the special case when employees are risk neutral towards money (i.e. u is linear), we find that the minimum wage segment stretches all the way to the top cell of $\mathcal{P}_*$. Thus there is an optimal *trigger wage*, which is $w_{\min}$ on every cell in $\mathcal{P}_*$ below the top cell and $w_{\min}$ plus a positive bonus B for all outputs in the top cell. (The lowest element θ of the top cell triggers the bonus).

Park (1995) derived a trigger wage for risk neutral workers under the much stronger hypothesis that f uniformly dominates g. He also did not consider risk averse workers. (He did, however, allow for multiple levels of effort, which under under uniform domination, can be accomodated in our model as well; see the Remark after Theorem 10). Our approach is also different, linking the pure wage problem to the pure titles problem.

Theorem 6 (Trigger Wage with Risk Neutrality): *If u is linear, then every optimal wage schedule must assume the constant value $w_{\min}$ for all outputs below the top cell of $\mathcal{P}_*$. For any optimal titles partition $\mathcal{P}$, there exists an optimal pure wage schedule w of the trigger form that pays $w_{\min}$ for all outputs below the top cell of $\mathcal{P}$ and a bonus $w_{\min} + B$ to outputs in the top cell of $\mathcal{P}$, where*

$$B = \frac{d}{\sum_{x \geq \theta}[f(x) - g(x)]}$$

and θ is the lowest output in the top cell of $\mathcal{P}$. For almost all f and g, this is the unique optimal wage schedule.

The next theorem shows that as the disutility of effort falls, the optimal pure wage schedule becomes more trigger like, as long as u displays increasing relative risk aversion.

Theorem 7 (Trigger-Like Wages with Risk Aversion): *Let employee utility u be differentiable and strictly concave. Suppose w_d, w_e are solutions of the pure wage problem above for disutilities $d < e$. Then $w_d(x) < w_e(x)$ whenever $w_e(x) > w_{\min}$* and *$w_d(x) = w_e(x)$ whenever $w_e(x) = w_{\min}$. Suppose in addition that u displays increasing relative risk aversion, i.e. suppose*

$$-\frac{u''(x)}{u'(x)}x$$

is strictly increasing in x. Let $w_{\min} \geq 0$. Then w_d looks more like a trigger wage schedule than w_e in the sense that $x < y$ implies

$$\frac{w_e(y) - w_{\min}}{w_e(x) - w_{\min}} < \frac{w_d(y) - w_{\min}}{w_d(x) - w_{\min}}.$$

2.3 Titles and Wages Together

Having considered titles and wages separately, we are ready to put them together. One surprise is that the optimal pure titles schedule need not change when wages are added. Imagine a pre-monetary society in which workers were motivated by titles alone. Suppose their disutility of work goes up, requiring further motivation from monetary wages. Then the optimal wage-title deployment would not alter the titles whatsoever, but on the contrary, simply reinforce them by paying wages according to the old titles.

Theorem 8 (Optimal Wage-Title Schedule): *Let wages be secret or public. Let $\sigma I^S < d \leq \sigma I^S + I^M$. (If $d \leq \sigma I^S$, wages are unnecessary for motivation and everybody could be paid $w_{\min}$). Suppose $(\mathcal{P}, w)$ is an optimal wage-title schedule. Then $\mathcal{P}$ is a solution for the pure titles problem and w is a solution for the pure wage problem with $d^* = d - \sigma I(\mathcal{P})$ in place of d. Moreover, for almost all f and g, the partition $\mathcal{P}$* is *uniquely determined.*

Thus with homogenous employees, there is nothing to be gained by keeping wages secret. Moreover, we should in general expect to see far fewer titles given than there are outputs, and far fewer wages than there are titles.

These features come starkly to light when employees are risk neutral. There is only one wage above the minimum, given as a bonus to an elite of top performers. Many title distinctions may occur below the elite, but all of them are paid the minimum wage.

Theorem 9 (Star Wages): *Suppose u is linear. Then, for generic f* and *g, the optimal wage-title schedule w is the trigger wage*

$$w_{(\theta,B)}(x) = \begin{cases} w_{\min} & \text{if } x < \theta \\ w_{\min} + B & \text{if } x \geq \theta \end{cases}$$

where θ is the smallest element of the top cell of $\mathcal{P} = \mathcal{P}^* = \mathcal{P}_*$ and

$$B = \frac{d^*}{\sum_{x \geq \theta}[f(x) - g(x)]}$$

where $d^* = d - \sigma I(\mathcal{P})$.

If furthermore, f uniformally stochastically dominates g, then the optimal title-wage schedule pays the minimum wage to every worker who does not achieve the top-most output, and a giant bonus to those who do.

We are now in a position to examine what happens when σ rises and society becomes more status conscious. Our main result is that in the presence of increasing relative risk aversion, increasing status has the effect of making the optimal wage schedule more star-like.

Theorem 10 (Status Creates Star-Like Wages): *Suppose u is twice differentiable and strictly concave, and displays strictly increasing relative risk aversion. Then as σ rises and the agents become more status conscious, wages fall and become more trigger like in the sense of Theorem 7 (Trigger-Like Wages with Risk Aversion). Indeed, the move from zero status to σ has the same effect as lowering the disutility by σI^S and finding an optimal pure wage schedule for the diminished disutility.*

Remark (Multiple Effort Levels) Suppose utilities are linear and agents have multiple effort levels $e_1, ..., e_{m-1}, e_m$ with corresponding stochastic outputs $X_{1,...,} X_{m-1}, X_m$ and disutilities d_i to switch from effort e_i to maximal effort e_m, for $i = 1, ..., m-1$. Further suppose that X_m uniformly stochastically dominates X_i for $i = 1, ..., m-1$. Then, by Theorem 9 (Star Wages), there exist a bonus B_i at the top-most output that will most efficiently motivate each agent to switch from e_i to e_m. But then $\widetilde{B} = \max\ [B_i : 1 \leq i \leq m-1]$, given as bonus at the topmost output, will be the best way to motivate all the agents to put in maximal effort.

Remark (Participation Constraints) Suppose utilities are concave, not necessarily linear. We could add an ex ante Participation Constraint (PC), over and above the minimal wage requirement, i.e., the (expected) utility any agent gets, from wages and titles combined, should never sink below some stipulated floor u_*. We can still prove that every optimal wage schedule must be measurable with respect to P^*. To see this, first note that, when they all work, their (expected) status utility is zero (by symmetry) and thus the utility they enjoy is just their wage utility. Take any wage schedule w that satisfies the incentive constraint and the participation constraint. Suppose w is not measurable with respect to the maximal optimal titles partition P^*. Then consider w^*, obtained from from w, exactly as in the proof of Theorem 3 (Compatibility of Pure Wages and Pure Titles). By construction, w^* leaves the worker's wage utility unchanged, hence w^* satisfies the participation constraint. The shirker's utility is worsened in w^* compared to w, hence the incentive constraint is satisfied in a stronger manner in w^*. Finally, the total wage bill is not increased. Thus an optimal wage schedule must be measurable with respect to P^*.

3 Disparate Employee Types

We now turn to the other extreme in which employees have disparate abilities. For the sake of a more succinct presentation, output distributions are taken to be continuous. (The discrete case is completely analogous, but the formulae become messier).

We imagine disparate employee types $i = 1, ..., \ell$ arranged in order of ascending abilities, with N_i employees of type i. Assuming all the others work, an employee of type i produces output continuously distributed on the high interval $J_H^i = [a_H^i, b_H^i]$ with density f_i when he works, and on the low interval $J_L^i = [a_L^i, b_L^i]$ with density g_i when he shirks. Conditional on others' working, his output depends only on chance and on his own effort, and is independent of all the others' outputs. We assume that abilities are *disparate*: $J_L^i < J_H^i < J_L^{i+1} < J_H^{i+1}$, i.e., an employee of type $i+1$ is so much more able than an employee of type i, that he always comes out ahead even when he shirks and the other works. In particular, the supports of the densities $g_i, f_i, g_{i+1}, f_{i+1}$ are all disjoint. This corresponds to a situation in which the employees can be clumped into distinct groups with widely different training or experience or expertise. The case when $\ell = 1$ is a special instance of the homogeneous employees we discussed in the last Section.

As in the previous Section with homogeneous employees, it will be useful to examine first the optimal pure title schedule when wages cannot be paid, and then the optimal pure wage schedule when titles cannot be conferred. We pass over these two cases quickly to get to the interesting interplay between wages and titles that was absent in the homogeneous case.

We shall find that the optimal titles partition gives as many titles as there are employee types, far fewer than the continuum of output levels. Typically the optimal pure title partition does not give the highest title to all the employees of the highest type, and allows for the best performers of the lowest type (and in fact all types below the top type) to gain a title equal to the worst performers among those one ability rank higher. When titles and wages are allocated together, the optimal titles partition changes. Secret wages make a difference.

When wages are public, the top ability types should be motivated entirely by wages, and not at all by status (though they get the highest status). A tiny group of elite performers among the top ability type should get astronomical wages. As status becomes more important, the disparity in pay between the highest types and all the other types increases; pay becomes more star like as status becomes more important. When wages are secret, there should be only two titles, so that all the status incentive is concentrated on the lowest ability type, and everyone else is motivated by wages.

Until the very end of this section we shall assume that $\sigma = 1$. As we said earlier, this is without loss of generality, since it can always be achieved be appropriately rescaling utilities.

3.1 Titles Alone

Once again titles will be given on the basis of performance as measured by a partition $\mathcal{P}$ of the output space into consecutive cells, as in the last section with homogeneous workers. Assuming all others are working, the expected status payoff to an employee when he works/shirks is given by the expected number of people he beats (according to $\mathcal{P}$) minus the expected number of people who beat him. His status incentive to work is his expected status payoff when he works minus his expected status payoff when he shirks. We begin by proving a lemma:

Lemma 4 (Cuts): *Suppose there are ℓ disparate types, and a given title partition. Then there is another title partition, with (1) the lowest cut at a_H^1, (2) at most one cut in every J_H^i, $i \geq 1$, and (3) no other cuts, which improves (or leaves unchanged) the status incentive to work of every employee.*

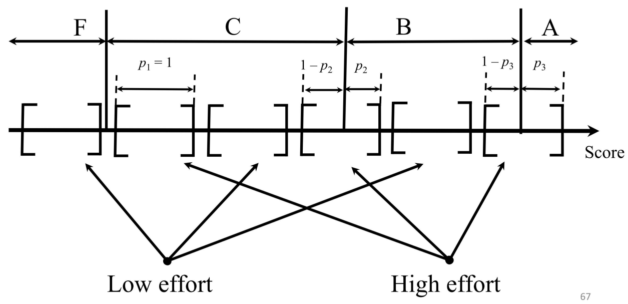
Let us denote the cut in J_H^i by c_i, and let $p_i = \int_{c_i}^{b_i^H} f_i(x)dx$ be the probability of the upper tail $J_H^i(c_i) = \{q \in J_H^i : q \geq c_i\}$. Since the density f_i may be zero on some intervals, there may be several c_i that give upper tails with the same probability. In this case it is harmless to choose the lowest such cut $c_i(p_i) = \min\{q \in J_H^i : \int_q^{b_i^H} f_i(x)dx = p_i\}$.

In view of the cuts lemma, we concentrate our attention on partitions $\mathcal{P}(p)$ given by vector $p = (p_{i_1} = p_1 = 1, p_{i_2} ..., p_{i_k})$ with cuts at $c_{i_j}(p_{i_j})$ in $J_H^{i_j}$, where the first cut comes at $c_{i_1}(p_{i_1}) = c_1(p_1) = c_1(1) = a_H^1$, corresponding to $i_1 = 1$.

Lemma 5 (Optimal Pure Titles): *Suppose there are ℓ disparate types, and a title partition that gives positive status incentive to work for every employee. Then there is another title partition consisting of ℓ cuts, with the lowest cut at a_H^1, and exactly one cut in every J_H^i, $i \geq 1$, which improves (or leaves unchanged) the status incentive to work of every employee.*

If there is exactly one cut per type, we may identify the partition $\mathcal{P}(p)$ with the probabilities $p = (p_1, ..., p_\ell)$, where $p_1 = 1$.

The Optimal Pure Titles Partition $\boldsymbol{P}(p_1, p_2, p_3)$



The number p_i is the probability with which an employee who does put in effort gets a title corresponding to those of his ability class who do work. To the extent $p_i < 1$, the incentive of type i employees to work is reduced. On the other hand, when $p_i < 1$, the status incentive of employees of type $i-1$ is enhanced, because by working then can come equal in status with a fraction $1 - p_i$ of the workers of type i. Every ability type $i \geq 2$ has a substantial status incentive to work because shirking forces them to be classified with the type $i-1$ just below them. But the lowest ability type $i = 1$ does not have that incentive.

Suppose there are $N_1, ...N_\ell$ employees of type $i = 1, ..., \ell$. Given the title partition $p = (p_1, ..., p_\ell)$, the status incentive to work for the ℓ types is

$$
\begin{aligned}
I^1(p) &= p_1[(N_1 - 1) + (1 - p_2)N_2] \\
I^i(p) &= p_i[(N_i - 1) + p_{i-1}N_{i-1} + (1 - p_{i+1})N_{i+1}] \text{ for } 2 \leq i \leq \ell - 1 \\
I^\ell(p) &= p_\ell[(N_\ell - 1) + p_{\ell-1}N_{\ell-1}].
\end{aligned}
$$

When working, an employee of type $2 \leq i \leq \ell - 1$ might get unlucky, with probability $1 - p_i$, and find himself no better off than if he shirked. But with probability p_i he will be lucky, outranking the fraction p_{i-1} of type $i - 1$ he otherwise would be equal with, and coming equal with the fraction $1 - p_{i+1}$ of type $i + 1$ he would otherwise have lost out against. In addition, he either outranks (instead of equalling) or equals (instead of being outranked by) every employee of his own type. This gives the formula $I^i(p)$ for $2 \leq i \leq \ell - 1$. Taking $N_0 = N_{\ell+1} = 0$ gives the formulas for $I^1(p)$ and $I^\ell(p)$.

In the case where all disincentives $d_i = d$, it is natural to maximize the minimum incentive to work, i.e. to choose $p \in [0, 1]^\ell$ to solve

$$\max_{p \in [0,1]^\ell} \min_{1 \leq i \leq \ell} I^i(p)$$

Since the $I^i(p)$ are continuous, and since $[0, 1]^\ell$ is compact, an optimal $\tilde{p}$ clearly exists. For more details, see Section 3 of Dubey-Geanakoplos (2005, 2010), where in particular it is shown that if $N_1 \leq N_2 \leq ... \leq N_\ell$, then $0 < \tilde{p}_i < 1$ for $2 \leq i \leq \ell$. (Of course $\tilde{p}_1 = 1$, for why reward any employee of type 1 for shirking.)

3.2 Wages Alone

As in the homogeneous case of the last section, we assume that all agents have the same concave utility function u for wages. We also assume that u is continuous and strictly monotonic, with $u(x) \longrightarrow \infty$ as $x \longrightarrow \infty$. If titles confer no status, and the owner must motivate his employees only by wages, then he must ensure that the wage incentive to work for each employee of type i is enough to overcome his disutility of working, i.e.

$$\int u(w(x))df_i(x) - \int u(w(x))dg_i(x) \geq d_i$$

It is perfectly clear what needs to be done. The owner would then simply set $p_i = 1$ for all $i = 1, ..., \ell$ and compensate each employee for precisely his disutility when he switches from shirk to work. Since wages must be monotonic in output, this implies that an optimal pure wage schedule is a step function which pays w_0 on $[0, a_H^1)$ and w_i on $[a_H^i, a_H^{i+1})$, where $w_0 < w_1 < ..... < w_l$ are defined recursively as follows, starting with w_0 :

$$\begin{aligned} w_0 &= w_{\min} \\ u(w_i) - u(w_{i-1}) &= d_i \end{aligned}$$

(Our assumptions on u guarantee the existence of such a wage schedule for any $w_{\min}, d_1, ..., d_l$.) Without status, wages rise with ability, but in increments determined entirely by the utility of wages and the disutility of effort. In particular, when employees are risk neutral, we have that the wage staircase starts at $w_0 = w_{\min}$ and jumps by d_i at output level a_H^i.

3.3 Wages and Titles Together

Once again we ask the question: given that the owner can use *both* titles and wages as incentives, how should he deploy them together? As we have just seen, solving for them separately often leads to very different partitions. For example, if $d_i = d$ for all i, cuts are in the interior of J_H^i for pure titles and on its boundary for pure wages. Superposing the wage schedule onto the pure titles schedule, as in the homogenous case, is in general not possible.

When wielding wages and titles together, should the owner use money and status in equal proportions for all employees? Or should he, for example, reserve status mostly for higher employee types? And does our answer depend on whether the wages are secret or public?

We begin by showing that Lemma 4 (Cuts) still applies when employees are also motivated by money.

Lemma 6 (Cuts with Wages): *Suppose there are ℓ disparate types, and an original wage-title schedule (with wages secret or public). Then the total incentive to work of every employee can be improved (or left unchanged) by another wage schedule whose wage bill is unchanged, together with a title partition with the lowest cut at a_H^1, and at most one cut in every J_H^i, $i \geq 1$, and no other cuts. Furthermore, if the original wage schedule was public (i.e. measurable with respect to the original titles partition), then the new wage schedule can also be taken to be public (i.e. measurable with respect to the new titles partition).*

3.3.1 Secret Wages

The owner's optimization problem is given below for the case of secret wages. He seeks to minimize his wage bill, subject to incentivizing everyone to work.

Lemma 6 (Cuts with Wages) already guarantees that we need only consider at most one cut per J_H^i. To allow for the possibility of missing cuts, we consider the k-vector $p = (p_{i_1} = p_1 = 1, p_{i_2} ..., p_{i_k})$, where the length k can vary. This defines the partition $\mathcal{P}(p)$ with cuts at $c_{i_j}(p_{i_j})$ in $J_H^{i_j}$, where the first cut comes at a_H^1, corresponding to $i_1 = 1$. Let $I^i(p)$ denote the status incentive created by the title partition p for employee-type i. Since wages are secret, they can be set independently of the title partition p. But then, on account of the concavity of u and the risk-neutrality of the owner, we might as well take wages to be a constant w_i on J_H^i and as low as possible (while respecting the constraint of monotonicity) on J_L^i, namely[17] w_{i-1}. Thus the owner's optimization problem may be written

$$\min_{p,w} \sum_{i=1}^{\ell} w_i N_i$$

$$\begin{aligned} \text{s.t. } I^i(p) + u(w_i) - u(w_{i-1}) &\geq d_i, \text{ for } 1 \leq i \leq \ell \\ 0 \leq p_{i_j} &\leq 1, \text{for } 2 \leq j \leq k \\ 1 = i_1 \leq i_2 \leq \ldots i_k &\leq \ell \\ w_{\min} = w_0 \leq w_1 &\leq ..... \leq w_l \end{aligned}$$

We require

The Necessity-of-Wages Assumption (Disparate Case):

A*t any feasible wage-title schedule (p, w), each agent must get a positive wage incentive (in addition to his status incentive),*[18].

At first glance one might think that the highest type, who will necessarily wind up with the highest status payoff, ought to be motivated by status, while the lowest type,

[17] Consider any weakly monotonic wage function w' . Let w_i denote the average value of w'on J_H^i and replace w' by the step function $w = (w_0, w_1, ..., w_l)$ as discussed. Then w will not raise the expected wage bill, and will create no less wage incentives, compared to w'.

[18] This is guaranteed if, for example, $N_{i-1} + N_i + N_{i+1} < d_i$ for all i (with $N_0 = N_{\ell+1} = 0$). Would anybody work for free, just for the status of coming ahead of all his peers?

who will necessarily wind up with the least status payoff, will need to be motivated by money. But quite the opposite is true. As we said in the introduction, shame is the flip side of honor. So status considerations apply at both ends. Furthermore it is the change in status payoff (or money payoff) upon switching from shirk to work that counts for incentive, not the absolute payoffs.

Since wages have to be monotonic, giving a raise to the bottom end will push wages up for all, creating a huge wage bill for the owner. It is to his advantage to make the initial rung of the wage staircase as low as possible. He can achieve this by incentivizing the lowest type as much as possible via status, so that the wage incentive needed for the lowest ability employee is small. When wages are secret, the employer can indeed concentrate all the status incentive on the lowest type.

Theorem 11 (Optimal Wage-Title Schedule with Secret Wages and Risk Neutrality) *Suppose there are ℓ disparate risk-neutral types. Then the optimal wage-title schedule, when wages can be kept secret, has just two titles, the low title for outputs below a_H^1 (that would be given to a person of the lowest type were he to shirk) and the high title for all outputs above a_H^1. Thus despite the freedom to hand out titles costlessly, the owner should award every worker the same high title. All workers above those of type 1 are incentivized by the secret wages alone. So the wage schedule is a step function where the jump from w_{i-1} to w_i takes place at a_H^i and*

$$\begin{aligned} w_0 &= w_{\min} \\ w_1 &= w_0 + d_1 - (N_1 - 1 + N_2 + ... + N_l) \\ w_i &= w_{i-1} + d_i \end{aligned}$$

The above optimal schedule is remeniscent of joining a club, all of whose members enjoy the same title, although they may secretly be getting different perqs.

Pure Wage Schedule vs Secret-Wage & Title Schedule The optimal pure wage schedule is a monotonic step function with jumps from w_{i-1} to w_i at at the start a_H^i of every interval $J_H^i = [a_H^i, b_H^i]$, satisfying $w_0 = w_{\min}$ and, with risk neutral employees, $w_i - w_{i-1} = d_i$. When titles are introduced and wages are secret, the optimal wage schedule is exactly the same except that the whole wage staircase is shifted down by $(N_1 - 1 + N_2 + ... + N_l)$, which is equal to the enormous status incentive created for the lowest type by the two titles (the shirker of the lowest type getting the low title and all other outputs the high title). In short, the lowest type is incentivized as much as possible by status, while the others are motivated by wages alone.

3.3.2 Public Wages

The owner's optimization problem is exactly the same as for secret wages, except for the added constraint that wages must be measurable with respect to titles, i.e., must be constant across all outputs that are awarded the same title.

As before, the owner seeks to minimize his wage bill, subject to incentivizing every employee to work. In light of Lemma 6 (Cuts with Wages), we may restrict attention to title partitions with at most one cut in every J_H^i. It is easy to see that every such cut must occur. For if some J_H^i had no cut, then J_L^i and J_H^i would have the same title, hence employees of type i would have no status incentive to work; but then, since wages are public, wages would have to be constant across J_L^i and J_H^i, and then employees of type i would have no wage incentive either. Therefore the title partition is represented by the full vector $p = (p_1, p_2, ..., p_\ell)$. Denote by $\mathcal{W}(p_1, p_2, ..., p_\ell)$ the class of wage schedules that are measurable with respect to the title partition $\mathcal{P}(p_1, p_2, ..., p_\ell)$. One critical aspect of the problem is that we have capped the maximum wage at an arbitrary, but high, level M.[19]

We may then state the employer's optimization problem as follows:

$$\begin{aligned}
&\min_{p,w} \sum_{i=1}^{\ell}[(1-p_i)w_{i-1}+p_i w_i]N_i = \min_{p,w}\left\{\sum_{i=1}^{\ell} w_{i-1}N_i + \sum_{i=1}^{\ell} p_i(w_i - w_{i-1})N_i\right\} \\
&\text{s.t. } \tilde{I}_i \equiv I^i(p) + p_i(u(w_i) - u(w_{i-1})) \geq d_i, \text{ for } 1 \leq i \leq \ell \\
&\qquad\qquad 0 \leq p_i \leq 1 \text{ and } w \in \mathcal{W}(p_1, p_2, ..., p_\ell) \\
&\qquad\qquad w_{\min} = w_0 \leq w_1 \leq \cdots \leq w_\ell \leq M
\end{aligned}$$

The total incentive, $\tilde{I}_i$, of each agent of type i consists, as before, of the status incentive $I_i \equiv I^i(p)$ plus a wage incentive $p_i(u(w_i) - u(w_{i-1}))$.

We shall see that, unlike the case of homogeneous employees, the optimal reward schedule does not arise by a simple combination of the solutions for titles alone and wages alone. There is a more intricate interplay between wages and titles. Wages will now depend on the population distribution of employees $N_1, ..., N_\ell$, as well as the disutilities of effort $d_1, ..., d_\ell$. But some features stand out independent of the N_i and d_i. The most dramatic change to the wage schedule is that now a tiny elite of top performers will be given exorbitant wages. The title partition will also change. For $1 \leq i \leq \ell - 1$, it is optimal to set $p_i = 1$ so that each of those types surely gets a higher title by working. Thus the cells C_i of the optimal title partition are as follows: $C_0 = J_L^1$ and, for $1 \leq i \leq \ell - 2$, $C_i = J_H^i \cup J_L^{i+1}$, and $C_{\ell-1} = J_H^{\ell-1} \cup J_L^\ell \cup [J_H^\ell \backslash J_H^\ell(p_\ell)]$, and $C_\ell = J_H^\ell(p_\ell)$. In the optimal public-wage & title schedule, p_ℓ is not 1 but is close to 0, meaning that only a tiny fraction consisting of the ultra productive employees of type ℓ will be awarded the topmost title when they work, while the majority of them will be pooled with the second best type $\ell - 1$. The optimal public-wage & title schedule is thus vastly different from the optimal pure titles schedule and from the optimal pure wage schedule.

The CEO is picked by lottery from the senior managing directors (type ℓ). Contrary to what one might have guessed, the type ℓ workers are motivated almost entirely by the chance of the huge money payoff of the CEO and not by status. On the other hand, the very next tier of managing directors (type $\ell - 1$) are motivated heavily by titles. Unlike all the other workers, by working hard they will come equal

[19] This helps to keep the problem compact. The bound M may also be interpreted as the degree of inequity aversion in the society (see Fehr-Schmidt (1999)).

not just with all of their own type but also with virtually all of the type above them. To sum up, the top tier is motivated by money, the next is motivated as much as possible by status, and the rest by an even mix of status and money.

To ease the formal exposition of the foregoing discussion, we make the

Differentiability Assumption:
u is continuously differentiable[20] *(in addition to being concave and strictly monotonic).*

More substantially we shall assume that agents become risk-neutral when their wealth is sufficiently large. Precisely, we have the condition below (which is automatically satisfied when u is linear):

Asymptotic Risk Neutrality Assumption:
(1) *u becomes asymptotically risk-neutral ,i.e., the derivative $u'(w) = \lambda$ for some constant λ whenever w exceeds a threshold w^τ.*
Furthermore, at any feasible wage-title schedule (p, w),
(2) $w_{\ell-2} > w^\tau$

Part (2) requires disutilities to be *sufficiently more* than status incentives (cumulatively across the types) so that, by the time the first $\ell - 2$ agents have been given the requisite wage increases, the threshold w^τ is crossed. Note that if $d = (d_1, ..., d_\ell) \longrightarrow \infty$,then $w = (w_1, ..., w_\ell) \longrightarrow \infty$ in order to keep (p, w) feasible; so (2) is automatic for large enough d. Tighter sufficient conditions can easily be stated in terms of the *exogenous* data $d, u, (N_i)_{i=1}^{\ell}$ of the model in order to guarantee (2), but we leave this to the reader.

Theorem 12 (Exorbitant Elite Wages): *Let there be ℓ disparate types of workers, with $N_i \geq 1$ of each type $i = 1, ..., \ell$. Suppose the Necessity of Wages and the Differentiability Assumptions hold. Then at any optimal wage-title schedule (p, w), we have $p_i = 1$ for all $i = 1, ..., \ell - 1$, so that any two workers of the same type below ℓ get the same status and wage.*

Next assume that the *Asymptotic Risk Neutrality Assumption also holds. and that* M is large enough to ensure that $\lambda M > \lambda w_{\min} + \sum_{i=1}^{\ell} d_i$ *(where, recall, λ denotes the constant value of the derivative $u'(w)$ for $w > w^\tau$).*

Then there is a unique optimal status partition and wage schedule (p, w), with $p_\ell \leq (\lambda w_{\min} + \sum_{i=1}^{\ell} d_i)/\lambda M < 1$, and $w_\ell = M$. Thus for large M, p_ℓ is very small and a tiny elite $p_\ell N_\ell$ out of the highest type ℓ is paid the exorbitant salary M, while the rest of their type obtain the same status and pay as type $\ell - 1$. Thus type ℓ employees are motivated almost entirely by wages alone.

Theorem 12 gives an explanation for the exorbitant pay often seen at the very top of some real world hierarchies. It is cheaper to incentivize the managing directors of type $\ell - 1$ as much as possible via status rather than wages. To achieve this they

[20] This is not essential and Theorem 12 below holds with just concavity of u. Its proof is exactly the same but with left (right) derivatives of u used to estimate decreases (increases) in u.

must be able to get the same status as most of the senior managing directors of type ℓ, if they work hard. This fixes the wage of the latter group at the managing director's level. In order to incentivize the senior managing directors, they are given to understand that the CEO will be chosen from among their rank, and even though the chance of getting selected is small, the salary is huge. (Denoting the probability of getting the top CEO title by ϵ, the status incentive of type l is $\epsilon(N_{l-1} + (1-\epsilon)N_l)$ which is negligible compared to the wage incentive ϵM, where M is the huge bonus.)

This stratagem of paying a huge salary to the tiny fraction of top performers in a group is counterproductive at any level below ℓ, because monotonicity would force the employer to pay *all* workers of higher type at least as much.

The conclusion that the top ability group is motivated by wages alone has the consequence that as status grows in importance (σ rises above 1) the pay of everyone is be reduced, except for the elite performers, who continue to get the same maximum M, since every group below the top ability group was getting some status incentive and now can get more. The difference in pay between the elite performers and the rest must therefore grow, and the fraction of employees getting the same elite wage must shrink. The wage schedule gets more star like.

3.3.3 Wage Differentials for Disparate Employees

The conclusions about exorbitant pay for the CEO and $p_i = 1$ for all $i = 1, ..., \ell - 1$ are quite robust; they hold regardless of the distribution of abilities $N_1, ..., N_\ell$, or the disutilities of work $d_1, ..., d_\ell$.

But the wage differentials $w_i - w_{i-1}$ for $i < \ell$ do depend on the N_j's and d_j's as we shall see. *For simplicity let us assume $u(w) = w$ for the rest of this section.* Our analysis is based on the following corollary:

Corollary to Theorem 12: *Under the conditions of Theorem 11 and with risk neutrality ($u(w) = w$), at the optimum wage-title schedule the title incentives are $I^1 = N_1 - 1 < I^2 = N_1 + N_2 - 1$; $I^i = N_{i-1} + N_i - 1$ for $i = 3, ..., \ell - 2$. Also, $I^{\ell-1} = N_{\ell-2} + N_{\ell-1} + (1 - p_\ell)N_\ell - 1 \approx N_{\ell-2} + N_{\ell-1} + N_\ell - 1$. Finally, $I^\ell = p_\ell(N_{\ell-1} + N_\ell - 1) \approx 0$.*

Thus for $2 \le i \le \ell - 2$,

$$(w_i - w_{i-1}) - (w_{i-1} - w_{i-2}) = (d_i - d_{i-1}) + N_{i-2} - N_i.$$

A natural case to consider is the one where the population N_i declines in size as the ability type increases. If disutilities do not fall as fast (i.e., if $N_{i-2} - N_i > d_{i-1} - d_i$, which occurs for example, if disutilities are constant), then we conclude from the corollary that wage differentials escalate as we go up the ability ladder from $i = 2$ to $i = \ell - 2$.

Another natural case arises in a population that is bell-shaped around the mean ability. When $N_i - N_{i-2} > d_i - d_{i-1}$ for small i and $N_{i-2} - N_i > d_{i-1} - d_i$, for large i, we get a wage schedule which is first concave and then convex.

The simplest case is when $N_i = N\ \forall i$ and $d_i = d\ \forall i$. Then the wage rises steadily by a fixed step of $d+1-2N$ until $w_{\ell-2}$, then rises by only $d+1-3N$ to $w_{\ell-1}$, then jumps astronomically to $w_\ell = M$.

Remark (When public wages confer status): We could have postulated, instead of titles, that status is conferred by wages themselves: w_i confers higher status than w_j if, and only if, $w_i \geq w_j + \delta$ for some threshold $\delta > 0$. Then our last constraint in the owner's optimization problem would read: $w_{\min} = w_0$, $w_i + \delta \leq w_{i+1}$ for $1 \leq i \leq \ell - 1$, $w_\ell \leq M$. It is worth noting that as $d = (d_1, ..., d_\ell) \to \infty$, $p_i(w_i - w_{i-1}) \to \infty$ since the status incentive terms $I_\ell^i(p)$ are bounded by $N_1 + \cdots + N_\ell$. Thus the constraints $w_i + \delta \leq w_{i+1}$ are automatically satisfied for large enough d (given any δ), and our analysis remains intact.

Participation Constraints with Risk Neutrality In the case of disparate agents, observe that no matter what the underlying partition for the wage-title schedule may be, the participation constraint (PC) is met by everyone if, and only if, it is met by an agent of type 1. This is so because his wage utility is never more than that of the others, on account of the monotonicity of the wages in terms of the output; nor is his status utility more, since titles are also monotonic and so render it impossible for him to outrank any higher type. Thus it suffices to maintain the PC for type 1. With this in mind, consider the proof of the Exorbitant Elite Wages Theorem. Start with any wage-title schedule which incentivizes everyone to work, while also meeting the PC for an agent of type 1. Now read the entire proof without change.We need only check that the PC for this agent is maintained throughout. But this is straightforward. Raising p_1 boosts both his wage utility and his status utility, so the PC continues to hold for him. Next, when we raise p_2, his status utility does go down in the amount εN_2 . But the subsequent increase of w_1 to $\tilde{w}_1$ raises his wage utility in precisely the same amount, so that the PC is still not violated for him. The rest of the proof proceeds without at all impacting agents of type 1. Hence the Exorbitant Elite Wages Theorem, and its proof, hold exactly as before, with just one amendment: in the optimal wage-title schedule that we wind up with, it may be that w_1 is escalated to ensure $u(w_1) - N_2 - .... - N_l = u_*$ (thereby meeting the PC for agents of type 1). There is no other change. The titles-partition and the wage differentials, starting from w_1, stay exactly the same.

4 Relative Wages and Titles

One might wonder whether it would be easier to motivate employees by paying them relative wages, i.e., wages and titles based on how their performance ranks relative to their rivals. We can formalize this by a sequence $K = (K_Z, ..., K_B, K_A)$, where the top K_A performers get the highest wage and title, the next K_B get the next highest wage and title, and so on. Ties are broken randomly with equal probability.

The answer is no.

4.1 Homogeneous Employees

Consider any money payment scheme in which the money payment to an employee is a function of his output and the output of all the others. Conditional on his own output, the worker thereby obtains a certain expected utility of the forthcoming money payment, which is equivalent to getting some wage for certain. (Since he is risk averse, this certainty equivalent wage is actually smaller than the expected money payment the owner is making, conditional on the worker's output.) If this certainty equivalent wage (thought of as a function of the worker's output) gives him the incentive to work, then it must cost the employer at least as much as the optimal pure wage schedule derived in section 3. Thus in our model, absolute wages cannot be beaten by relative wages or any other wage scheme, when status considerations are absent.

On the other hand, with pure status, the optimal absolute partition of the Proposition beats any relative schedule $K = (K_Z, ..., K_B, K_A)$, by Dubey-Geanakoplos (2005,2010). But as we saw in the Optimal Wage-Titles Theorem, this same partition also serves for the optimal wage schedule. Thus the same absolute partition gives more status incentive than any relative scheme, and also gives more wage incentive than would be generated by any relative scheme. Since status incentives and wage incentives are additive, this absolute wageititle schedule is better than any relative schedule.

4.2 Disparate Employees

Consider a general population $N = (N_1, ..., N_\ell)$ of ℓ disparate types, and any relative wage-title schedule given by $K = (K_A, K_B, ..., K_Z)$, $K_A + \cdots + K_Z = N_1 + \cdots + N_\ell$. (Recall that we don't need to worry about ties since outcomes are continuously distributed). We can find an absolute wage-title schedule that creates at least the same incentives (from status and money combined), while handing out the same amount of money.

Define absolute grade intervals by the intervals $[x_A, \infty), [x_B, x_A)$ and so on, where the cuts x_α are defined by the *maximum* values solving the equations

$$K_A + \cdots + K_\alpha = \text{Expected number of people with scores in } [x_\alpha, \infty)$$

assuming everybody works. Award the relative wages and titles on these absolute intervals. It is easy to check that the absolute wage-title schedule we have defined, costs the same and creates (using the concavity of u) at least the same incentives.

5 Incomplete Information

We have assumed so far that every player knows the precise characteristics of every other player, in addition to his own. Our analysis can be modified very easily to accomodate *incomplete information, i.e.,* when each player knows his own characteristics precisely, but has only a probability distribution on those of others.

First consider our model of N homogeneous employees, i.e., each produces random output with the same probability density f, g if he works, shirks (independently of the effort chosen by the others). In order to introduce incomplete information, let us suppose that the disutility of effort can take on many possible values $d_1 < .... < d_k$. Nature moves first, independently picking a disutility level for everyone and revealing to each only his own. An optimal reward schedule must motivate every employee to work no matter what his disutility level may be. This is clearly equivalent to motivating an employee to switch from shirk to work when his disutility is the highest possible (i.e., is d_k) and when the remaining $N-1$ employees are working. Thus the optimal reward schedule we have constructed in the complete information case, when all employees have the common disutility d_k, is also the optimal schedule with incomplete information.

Next consider the case of l disparate ability-types, with disutility d_i and disjoint performance intervals $J_L^i < J_H^i$ for type i, as before. Here the natural game of incomplete information (that we have in mind) is as follows. Nature moves first, assigning type $i = 1, ..., l$ randomly to everyone with probabilities that are i.i.d [21] across the employees, say type i is picked with probability θ_i. Each employee comes to know his own type and not those of the others, before choosing his effort level. But the game of course is common knowledge, so each is cognizant of the probability distribution on types.

The status incentive for any employee of type i is linear in the expected number of rivals of each type. When there is complete information, these numbers are deterministically given by the vector $(N_1, ..., N_{i-1}, N_i - 1, N_{i+1}, ..., N_\ell)$. When there is incomplete information, this vector does not depend on i, and is always given by the expected numbers $(\theta_1(N-1), ..., \theta_\ell(N-1))$. Based on this observation, the entire analysis of the disparate case can be transported from complete information to incomplete information as follows. Lemma 4 (Cuts) , Lemma 5 (Optimal Pure Titles), and Lemma 6 (Cuts with Wages) all hold mutatis mutandis. Theorem 11 (Optimal Wage-Title Schedule with Secret Wages and Risk Neutrality) holds with a slight alteration. There is an optimal secret wage-title schedule (though no longer necessarily unique) that still consists of just one cut at a_H^1. The formula for wages is also just the same, except that

$$w_1 = w_0 + d_1 - (N-1)$$

To see this, define ν_{i_j} just as in the proof of Theorem 11 but with each N_i replaced by $\theta_i N$ (i.e.,the expected number of employees in the region V_{i_j}) and define $\nu_{i_j}^* = \theta_i(N-1) = [(N-1)/N]\nu_{i_j}$ (the expected number of *others* in V_{i_j}, conditional on one employee — of any type —- standing aside). Then re-read the proof of Theorem 11 with the following amendments: the changes in status incentives are given by the same formulae replacing ν_{i_j} by $\nu_{i_j}^*$ throughout and dropping "-1"(thus $\nu_{i_1} + \nu_{i_2} - 1$ is replaced by $\nu_{i_1}^* + \nu_{i_2}^*$, etc.). Repeating the maneuver of wage changes as we move to the single cut at a_H^1, the total change in wage bill is *no more than*

[21] What is important is that the probabilities be independent across the employees. We suppose that they are identically distributed only for ease of notation.

$$
\begin{aligned}
&- (\nu_{i_1} + \nu_{i_2} + \nu_{i_3} + ... + \nu_{i_k})[\nu^*_{i_2} + \nu^*_{i_3} + ... + \nu^*_{i_k}] \\
&+ (\nu_{i_2} + \nu_{i_3} + ... + \nu_{i_k})[\nu^*_{i_1} + \nu^*_{i_2}] \\
&+ (\nu_{i_3} + ... + \nu_{i_k})[\nu^*_{i_2} + \nu^*_{i_3}] + ... \\
&+ (\nu_{i_k})[\nu^*_{i_{k-1}} + \nu^*_{i_k}]
\end{aligned}
$$

But since $\nu^*_{i_j} = [(N-1)/N]\nu_{i_j}$ we may undo the stars in the above display (scaling the expression by $(N-1)/N$), which reveals that the displayed expression is 0 as before. though it may no longer be the unique optimal schedule (as was the case with complete information).

Finally consider the case of public wages. A variant of Theorem 12 (Exorbitant Wages) also remains intact: all optimal partitions must have $p_i = 1$ for $1 \leq i \leq l-1$ and — with the Asymptotic Risk Neutrality Assumption — there exists *an* optimal schedule with $p_l \leq (\lambda w_{\min} + \sum_{i=1}^{l} d_i)/M < 1$ and $w_l = M$. As with secret wages, we can no longer assert *uniqueness* of the optimal wage-title schedule.

Let us outline the changes needed in the proof of Theorem 12 for establishing this variant. Notice first that we must once again have exactly one cut in each interval J^i_H, for if such a cut were missing then wages would have to be the same for J^i_L and J^i_H as these two intervals get the same title; and thus i would have no incentive to work whatsoever. Now, as pointed out earlier, the formulae for wage bill (resp. status incentive) are preserved if we replace N_i (resp. N_i and $N_i - 1$) by $\overline{N}_i = \theta_i N$ (resp. $N^*_i = \theta_i(N-1) = [(N-1)/N]\overline{N}_i$). With these substitutions we can literally repeat the proof of Theorem 11; indeed, the estimates for *changes* in the wage bill, as we go through the wage-schedule modifications prescribed in that proof, will be the same *exact* expressions as before, replacing $N_i N_{i-1}$ (or, $N_i N_{i+1}$) in the proof by $[(N-1)/N]\overline{N}_i\overline{N}_{i-1}$ (or, $[(N-1)/N]\overline{N}_i\overline{N}_{i+1}$) throughout. The reason is that changes in the wage bill are the product of two terms:

(a) changes in the wage (which compensate for changes in status incentive, and therefore involve terms N^*_i); and

(b) the expected number of workers for whom that change is occuring (which involve $\overline{N}_i$)

But products like $\overline{N}_i N^*_{i-1}$are equal to $[(N-1)/N]\overline{N}_i\overline{N}_{i-1}$. This summarizes the main changes, and the rest of the argument proceeds exactly as before. It shows that positive reductions are achieved in the wage bill whenever we increase $p_i < 1$ to $p_i + \varepsilon$ for $1 \leq i \leq l-1$, hence such $p_i = 1$ as claimed; and when we lower p_l, the wage bill is unaffected (instead of being strictly reduced), establishing the claim regarding p_l and w_l.

Thus exorbitant wages *must* occur with complete information and constitute *one* of the feasible optima if there is incomplete information. This leads us to conjecture that exorbitant wages become necessary for any information regime that is in between the two. The modeling of such information regimes and the precise formulation of the result is left to future research.

6 References

Auriol E. and R. Renault (2008) "Status and Incentives", *RAND Journal of Economics*, vol 39(1), pp. 305-326.

Brown, G., and J. Gardner, A. Oswald, and J. Qian (2008) "Does Wage Rank Affect Employees Well Being", Industrial Relations: A Journal of Economy and Society, Volume 47, Issue 3, pages 355–389.

Cole, Harold L., George J. Mailath, and Andrew Postlewaite (1992). "Social Norms, Savings Behavior, and Growth," *Journal of Political Economy*, 100(6): 1092–1125.

——— (1995). "Incorporating Concern for Relative Wealth into Economic Models," *Federal Reserve Bank of Minneapolis Quarterly Review*, 19(3): 12–21.

——— (1998). "Class Systems and the Enforcement of Social Norms," *Journal of Public Economics*, 70: 5–35.

Corneo, Giacomo and Oliver Jeanne (1997). "On Relative Wealth Effects and the Optimality of Growth," *Economics Letters*, 54: 87–92.

Dubey, P. and Chien-wei Wu (2001). "When Less Scrutiny Induces More Effort," *Journal of Mathematical Economics*, 36(4): 311–336.

Dubey, P. and J. Geanakoplos (2005). "Incentives in Games of Status: Marking Exams and Setting Wages," Cowles Foundation Discussion Paper No. 1544, Yale University.

Dubey, P. and J. Geanakoplos (2010). "Grading Exams: 100, 99, ..., 1 or A, B, C?" *Games and Economic Behavior*, 69, pp 72-94.

Dubey, P. and O. Haimanko (2003). "Optimal Scrutiny in Multiperiod Promotion Tournaments," *Games and Economic Behavior*, 42(1): 1–24.

Duesenberry, James S. (1949). *Income, Saving and the Theory of Consumer Behavior*. Cambridge: Harvard University Press.

Fehr, E. and K.M. Schmidt (1999). "A Theory of Fairness, Competition and Cooperation," *Quarterly Journal of Economics*, 114: 817–868.

Frank, Robert H. (1985). *Choosing the Right Pond: Human Behavior and the Quest for Status*. New York: Oxford University Press.

Green, J.R. and N.L. Stokey (1983). "A Comparison of Tournaments and Contracts," *Journal of Political Economy*, 91: 349–364.

Hopkins, Ed and Tatiana Kornienko (2003). "Ratio Orderings and Comparative Statics," Working Paper, University of Edinburgh.

Lazear, E. and S. Rosen (1981). "Rank-Order Tournaments as Optimal Labor Contracts," *Journal of Political Economy*, 89: 841–864.

Moldovanu, B. and A. Sela (2001). "The Optimal Allocation of Prizes in Contests," *American Economic Review*, 91(3): 542–558.

Moldovanu, B., A. Sela and X. Shi (2007). "Contests for Status" Journal of Political Economy, vol 115, pp 338-363.

Park, E. (1995) "Incentive Contracting Under Limited Liability", Journal of Economics and Management Strategy, 4, 3, 477-490.

Pollak, Robert (1976). "Interdependent Preferences," *American Economic Review*, 66(3): 309–320.

Robson, Arthur (1992). "Status, the Distribution of Wealth, Private and Social Attitudes to Risk," *Econometrica*, 60(4): 837–857.

Shaked, M. and J.G. Shanthikumar (1994). *Stochastic Orders and their Applications*. San Diego: Academic Press.

Veblen, Thorstein (1899). *The Theory of the Leisure Class*. New York: Macmillan.

7 Appendix

We begin with a series of lemmas that will be helpful for proving both theorems.

Recall that for any subset $A \subset Q$, we let $f(A) = \sum_{x \in A} f(x)$ and $g(A) = \sum_{x \in A} g(x)$.

Incentive Lemma: *Let a cell $C = L \cup R$ in a partition P be the union of two consecutive intervals $L < R$. If*

$$\frac{f(L)}{g(L)} < \frac{f(R)}{g(R)}, \frac{f(L)}{g(L)} = \frac{f(R)}{g(R)}, \frac{f(L)}{g(L)} > \frac{f(R)}{g(R)}$$

then the incentive to work is strictly improved, left unchanged, strictly worsened (respectively) by splitting C into L and R.

Proof: Splitting C changes the incentive to work by $f(R)g(L) - f(L)g(R)$.■

Merger Lemma: *Suppose inside domination holds separately on two consecutive intervals $L < R$. If*

$$\frac{f(L)}{g(L)} \geq \frac{f(R)}{g(R)}$$

then inside domination holds on the single cell $L \cup R$.

Proof: Take an arbitrary cut of R into $R_- < R_+$. By inside dominaton $Y \succsim X$ on R,

$$\frac{f(R_-)}{g(R_-)} \geq \frac{f(R_+)}{g(R_+)}$$

Putting this together with the hypothesized inequality immediately gives

$$\frac{f(L \cup R_-)}{g(L \cup R_-)} \geq \frac{f(R_+)}{g(R_+)}$$

If the cut occurs inside L, an analogous proof works. If the cut divides L from R, there is nothing to prove.■

Splitting Lemma: *Suppose inside dominaton holds on a cell $C = L \cup R$ that is the union of two consecutive intervals $L < R$. If*

$$\frac{f(L)}{g(L)} \leq \frac{f(R)}{g(R)}$$

then inside dominaton holds separately on each of the cells L, R.

Proof: Take an arbitrary cut of R into $R_- < R_+$. By inside dominaton $Y \succsim X$ on $L \cup R$,

$$\frac{f(L \cup R_-)}{g(L \cup R_-)} \geq \frac{f(R_+)}{g(R_+)}$$

It follows from this and the inequality hypothesized that

$$\frac{f(R_-)}{g(R_-)} \geq \frac{f(R_+)}{g(R_+)}$$

proving that R satisfies inside dominaton. A similar argument applies to L.■

Though uniform domination is stronger than domination, the two become equivalent when they are opposed:

Constant Ratio Lemma: If X *uniformly dominates* Y *on the interval* $[c, d]$ and Y *dominates* X *on* $[c, d]$, *then*

$$\frac{f(a)}{g(a)} = \frac{f(b)}{g(b)}$$

for all $a, b \in [c, d]$.

Proof: Since $X \succsim_U Y$ on $[c, d]$,

$$\frac{f(b)}{g(b)}$$

is weakly increasing in $b \in [c, d]$. However, since $Y \succsim X$ on $[c, d]$, taking a cut just before the last element d gives

$$\frac{\Pr(X \in [c, d))}{\Pr(Y \in [c, d))} \geq \frac{f(d)}{g(d)}$$

These two conditions are compatible only if $f(b)/g(b)$ is constant for $b \in [c, d]$. ■

Constant Incentive Lemma: *Suppose the partition $\mathcal{P}$ satisfies inside dominaton and that the partiton $\mathcal{P}'$ satisfies outside domination. If $\mathcal{P}'$ is a refinement of $\mathcal{P}$, then $\frac{f(A)}{g(A)}$ is a constant across the cells A of P' that subdivide any cell C of $\mathcal{P}$; consequently both partitions give the same status incentive.*

Proof: On any cell C of $\mathcal{P}$, Y dominates X by the inside domination hypothesis on $\mathcal{P}$. But across the cells of $\mathcal{P}'$ that subdivide this cell C of $\mathcal{P}$, X uniformly dominates Y by the outside domination hypothesis on $\mathcal{P}'$. Hence, by the Constant Ratio Lemma, $\frac{f(A)}{g(A)}$ is a constant across the cells A of $\mathcal{P}'$ that subdivide C. Therefore by the Incentive Lemma, incentives are the same for $\mathcal{P}$ and $\mathcal{P}'$.■

Tail Lemma: *Let P be a partition of Q into consecutive cells $\{..., E < B, ...\}$ that satisfies inside and outside domination with respect to X and Y. Suppose that $D \subset E$ is a right tail segment of E and $C \subset B$ is a left tail segment of B. Then*

$$\frac{f(D)}{g(D)} \leq \frac{f(C)}{g(C)}$$

Proof: From outside domination we know that

$$\frac{f(E)}{g(E)} \leq \frac{f(B)}{g(B)}$$

From inside domination we know that

$$\frac{f(D)}{g(D)} \leq \frac{f(E)}{g(E)} \text{ and } \frac{f(B)}{g(B)} \leq \frac{f(C)}{g(C)}$$

proving the lemma.■

Join-Meet Lemma: *Suppose the partitions P and P' each satisfy inside and outside domination. Then so do their join and meet.*

Proof: Let $B^{\vee} < A^{\vee}$ be any two consecutive cells of the join $\mathcal{P} \vee \mathcal{P}^{'}$. The cut between them must come from one of the partitions, hence by the Tail Lemma

$$\frac{f(B^{\vee})}{g(B^{\vee})} \leq \frac{f(A^{\vee})}{g(A^{\vee})}.$$

Thus f/g is increasing over the cells of $\mathcal{P} \vee \mathcal{P}^{'}$ as we move to the right, proving outside domination for $\mathcal{P} \vee \mathcal{P}^{'}$. By the Splitting Lemma, inside domination holds for each cell of $\mathcal{P} \vee \mathcal{P}^{'}$ contained in any cell of $\mathcal{P}$ or $\mathcal{P}^{'}$, and hence it holds in every cell of $\mathcal{P} \vee \mathcal{P}^{'}$. Thus inside and outside domination hold for the join.

By the Constant Ratio Lemma, inside domination on $\mathcal{P}$ and inside domination on $\mathcal{P}'$ then imply that f/g is constant over all the cells from $\mathcal{P} \vee \mathcal{P}'$ that lie in the same cell of $\mathcal{P}$ and over all the cells from $\mathcal{P} \vee \mathcal{P}'$ that lie in the same cell of $\mathcal{P}'$. Hence f/g is constant over all the cells from $\mathcal{P} \vee \mathcal{P}'$ that lie in any cell of $\mathcal{P} \wedge \mathcal{P}'$. Hence by the merger lemma, the inside condition holds on each cell of $\mathcal{P} \wedge \mathcal{P}'$. The outside condition for the meet follows from the simple fact that f/g is rising across the cells of the join.■

Proof of Theorem 1 (Inside-Outside Condition):

Proof of Necessity: If the inside condition is violated when some cell C of an optimal partition $\mathcal{P}$ is cut into two consecutive cells $L < R$, then (by the incentive lemma), splitting C improves incentives, contradicting the optimality of $\mathcal{P}$. Similarly, if f/g strictly falls across two consecutive cells of $\mathcal{P}$, then (by the incentive lemma) merging them strictly increases incentives. Hence f/g must be (weakly) increasing across all cells of $\mathcal{P}$.■

Proof of Sufficiency: Suppose $\mathcal{P}$ satisfies the inside and outside conditions. Since Q is finite, there must trivially exist an optimal partition $\bar{P}$. By the necessity proof, $\bar{P}$ satisfies the inside and outside conditions. By the Join-Meet Lemma and the Constant Incentive Lemma, the meet $\mathcal{P} \wedge \bar{\mathcal{P}}$ and the join $\mathcal{P} \vee \bar{\mathcal{P}}$ give the same incentive as both $\mathcal{P}$ and $\bar{P}$, proving that $\mathcal{P}$ is also optimal.■

Proof of Theorem 2 (Lattice Structure): By Theorem 1 and the Join-Meet lemma, the optimal title partitions form a lattice as claimed. Let C be a cell in the maximal optimal partition $\mathcal{P}^*$. If some cut of C into consecutive intervals $L < R$ leaves the ratios $f(L)/g(L) = f(R)/g(R)$, then (by the incentive lemma) it also leaves incentives unchanged when C is split into L and R, contradicting the maximality of $\mathcal{P}^*$. Hence the inside condition must always hold strictly on cells of $\mathcal{P}^*$.

Next, let $L < R$ be consecutive cells of the minimal optimal partition $\mathcal{P}_*$. Again, if $f(L)/g(L) = f(R)/g(R)$, then (by the incentive lemma) it also leaves incentives unchanged when L and R are merged into $C = L \cup R$, contradicting the miminality of $\mathcal{P}_*$. Hence the outside condition must always hold strictly on cells of $\mathcal{P}_*$.

Consider a partition $\mathcal{P}$ obtained via any subset of the cuts of the join $\mathcal{P}^*$ that includes all the cuts of the meet $\mathcal{P}_*$. By the constant ratio lemma, f/g is constant across the cells of $\mathcal{P}^*$ that lie in the same cell of $\mathcal{P}_*$ Hence by the merger lemma, the new partition $\mathcal{P}^*$ must also satisfy the inside condition on each of its cells. It obviously inherits the outside condition from $\mathcal{P}^*$. Hence $\mathcal{P}$ is optimal by Theorem 1.■

Proof of Lemma 1 (Coarse Partition): If there were a cut between x and y, then by the Tail Lemma we would have

$$\frac{f(x)}{g(x)} \leq \frac{f(y)}{g(y)}$$

a contradiction.■

Proof of Lemma 2 (Fine Partition): This is an immediate corollary of Theorem 1.■

Proof of Lemma 3 (Unicity of the Lattice): For almost all f and g, it is clear that

$$\frac{f(A)}{g(A)} \neq \frac{f(B)}{g(B)}$$

for any two unequal intervals A, B of Q. By the constant ratio lemma, any optimal partition that is finer than the minimal optimal partition $\mathcal{P}_*$ must produce an equality on cells that lie in the same cell of $\mathcal{P}_*$.■

For what follows it will be useful to recall a standard property of stochastic dominance (see e.g. Shaked-Shanthikumar 1994).

Lemma (Dominance Increases Expectation) : Suppose that Y dominates X on the interval C, and that $f(C)g(C) > 0$. If $\phi : C \to \mathbb{R}$ is any monotonic function, then

$$\frac{1}{g(C)} \sum_{x \in C} \phi(x)g(x) \geq \frac{1}{f(C)} \sum_{x \in C} \phi(x)f(x)$$

where the inequality is strict if ϕ is not constant, and Y strictly dominates X.

Proof: The proof is by induction on the number of values ϕ takes on. If it takes on just one value, there is nothing to prove. So assume the theorem is true if ϕ takes on k values. Now consider a ϕ which takes on $k+1$ values $c_1 < ... < c_k < c_{k+1}$. Let C' be the right tail of C on which ϕ takes its maximal value c_{k+1}. Define ϕ' by leaving it unchanged on $C \backslash C\prime$ and reducing ϕ on $C\prime$ from c_{k+1} to c_k.. From the domination hypothesis

$$\frac{1}{g(C)}(c_{k+1} - c_k)g(C') \leq \frac{1}{f(C)}(c_{k+1} - c_k)f(C')$$

where the inequality is strict if g strictly dominates f. By the inductive assumption

$$\begin{aligned}\frac{1}{g(C)} \sum_{x \in C} \phi(x)g(x) &= \frac{1}{g(C)} \sum_{x \in C} \phi'(x)g(x) + \frac{1}{g(C)}(c_{k+1} - c_k)g(C') \\ &\leq \frac{1}{f(C)} \sum_{x \in C} \phi'(x)f(x) + \frac{1}{f(C)}(c_{k+1} - c_k)f(C') = \frac{1}{f(C)} \sum_{x \in C} \phi(x)g(x)\end{aligned}$$

where again the inequality is strict if g strictly dominates f.■

Proof of Theorem 3 (Compatibility of Pure Wages and Pure Titles) Let $w : Q \to [w_{\min}, \infty)$, be any wage schedule in $\mathcal{W}$ that is not measurable wrt the maximal optimal titles partition $\mathcal{P}^*$. We shall construct another wage schedule $w^* \in \mathcal{W}$ that is measurable wrt the maximal optimal titles partition $\mathcal{P}^*$ and creates a strictly higher incentive to work than w does.

For each cell C of $\mathcal{P}^*$, define the constant wage w_C such that

$$u(w_C) = \frac{1}{f(C)} \sum_{q \in C} f(q) u(w(q))$$

(If $f(C) = 0$, set $w_C = \max\{w(x) : x < C\}$; otherwise, w_C exists because of the intermediate value theorem and the continuity of u.) Since u is concave and monotonic,

$$w_C \leq \frac{1}{f(C)} \sum_{q \in C} f(q) w(q)$$

for every $C \in \mathcal{P}^*$. Hence the function w^*, made by patching w_C across all $C \in \mathcal{P}^*$, is no more costly for the employer than w. By construction, the worker gets the same utility payoff from both wage schedules. It remains to show that the shirker gets a strictly lower utility payoff in w^* than in w, implying that w^* creates strictly more incentive to work than w.

From Theorem 2 (Lattice Structure) , we know that g strictly dominates f on C. Since u and w are both monotonic, so is $u(w(q))$, and hence by the Lemma above (Dominance Increases Expectation) when $f(C)g(C) > 0$

$$u(w_C) = \frac{1}{f(C)} \sum_{q \in C} f(q) u(w(q)) \leq \frac{1}{g(C)} \sum_{q \in C} g(q) u(w(q))$$

with strict inequality on cells C on which w is not constant. So the change in the shirker's utility payoff on any C by moving from w to w^* is

$$\sum_{q \in C} g(q)(u(w_C) - u(w(q)) =$$

$$g(C)u(w_C) - \sum_{q \in C} g(q) u(w(q)) \leq 0$$

from the above when $f(C)g(C) > 0$, and trivially when $g(C) = 0$, and also when $f(C) = 0$, because in this last case $w_C \leq w(q)$ for all $q \in C$. The change is strict on every cell C on which w is not constant. Thus if w is not measurable with respect to $\mathcal{P}^*$, it is not optimal.

To prove the second half of the theorem, note first that feasible wage schedules form a compact subset of the finite dimensional space $\mathbb{R}^Q$. By continuity of the total wage bill, an optimal solution w exists, which by our proof is measurable with respect to $\mathcal{P}^*$. By averaging as above over cells of $\mathcal{P}_*$ instead of $\mathcal{P}^*$, we obtain a wage schedule w_* on $\mathcal{P}_*$. By construction the expected utility to the worker remains the same, and the wage bill does not go up. But since f/g is constant over the cells of $\mathcal{P}^*$ which constitute any given cell of $\mathcal{P}_*$, the expected utility to the shirker has also remained constant. Thus the incentive to work is unchanged. This proves that w_* is also an optimal wage schedule.■

Proof of Theorem 4 (Minimum Wage with Risk Aversion): By the outside condition, $f(C_i)/g(C_i)$ is weakly increasing in i. Hence if $f(C_i) > g(C_i)$, then for all $j > i$, $f(C_j) > g(C_j)$. Hence $f(C_i) \leq g(C_i)$ for all $i \leq k$. As just shown, the owner is paying a constant wage w_i on each cell C_i. Reducing all w_i to $w_{\min}$ for all $i \leq k$ reduces the wage bill and increases the incentive to work, unless already $w_i = w_{\min}$ for all $i \leq k$.■

Proof of Corollary to Theorem 4 (Uniqueness with Strict Risk Aversion) From Theorem 3 (Compatibility of Wages and Titles) every optimal schedule w musr be constant on the cells $C_1, ..., C_L$ of $\mathcal{P}^*$. Hence we may write the owner's optimization problem as

$$\min_{w \in \mathbb{R}^L} \sum_{\ell=1}^{L} f(C_\ell) w_\ell$$

$$\text{s.t.} \sum_{\ell=1}^{L} [f(C_\ell) - g(C_\ell)] u(w_\ell) \geq d$$

$$w_{L+1} \equiv \infty > w_L \geq ... \geq w_1 \geq w_{\min} \equiv w_0$$

By Theorem 4 there is a k such that $w_i = w_{\min}$ for all $i \leq k$; and $f(C_\ell) > g(C_\ell)$ for all $\ell > k$. If there are two distinct solutions, then at least one of them has $w_L > w_{\min}$, otherwise both would be identically $w_{\min}$. Now the half-half convex combination w^* (of the two solutions) trivially satisfies the bottom sequence of linear inequalities and leaves the minimand unchanged. Furthermore, since the two solutions agree (and are equal to $w_{\min}$) at every ℓ with $f(C_\ell) - g(C_\ell) \leq 0$, the combination w^* must satisfy the first (incentive) constraint strictly on account of the strict concavity of u. We can lower slightly all the wages in w^* that are strictly above $w_{\min}$, maintaining all the constraints but lowering the wage bill, a contradiction.■

Proof of Theorem 5 (Wage Structure with Risk Aversion): From Theorem 3 (Compatibility of Wages and Titles) there is indeed an optimal schedule w that is constant on each cell C_ℓ of $\mathcal{P}_*$. Hence we may write the owner's optimization problem exactly as in the proof of the Corollary to Theorem 4, but with $\mathcal{P}^*$ replaced with $\mathcal{P}_*$. Furthermore, by the same logic applied to $\mathcal{P}_*$ as was applied to $\mathcal{P}^*$ in the proof of Theorem 4, we need only consider paying $w_i > w_{\min}$ for cells with $f(C_i) > g(C_i)$.

By the Kuhn-Tucker theorem, there is $\mu > 0$ such that for all $1 \leq \ell \leq L$ with $f(C_\ell) > g(C_\ell)$,

$$\begin{aligned} f(C_\ell) - \mu[f(C_\ell) - g(C_\ell)]u'(w_\ell) &= 0 \text{ if } w_{\ell-1} < w_\ell < w_{\ell+1} \\ &\geq 0 \text{ if } w_{\ell-1} \leq w_\ell < w_{\ell+1} \\ &\leq 0 \text{ if } w_{\ell-1} < w_\ell \leq w_{\ell+1} \end{aligned}$$

It follows that for all $2 \leq i < j \leq L$, if $w_{i-1} < w_i \leq w_j < w_{j+1}$, then

$$\frac{u'(w_j)[f(C_j) - g(C_j)]}{f(C_j)} \leq \frac{1}{\mu} \leq \frac{u'(w_i)[f(C_i) - g(C_i)]}{f(C_i)}$$

The left hand term is the gain in incentive to work per expected dollar spent on the margin in cell j. If the inequality were violated, then by increasing w_j by very small $\varepsilon > 0$ and decreasing w_i by slightly more than $\varepsilon f(C_j)/f(C_i)$, the owner could reduce his expected wage bill and increase the incentive to work. From the fact that $[f(C_\ell) - g(C_\ell)]/f(C_\ell) = 1 - g(C_\ell)/f(C_\ell)$ is strictly increasing in ℓ, because $f(C_\ell)/g(C_\ell)$ is strictly increasing across the cells of the minimal partition $\mathcal{P}_*$, it follows that $u'(w_j) < u'(w_i)$ and so $w_i < w_j$. Thus not more than one cell can give the same wage $w > w_{\min}$ (for call the first one i and call the last one j). The first order conditions are then all equalities and the formula given by the theorem follows from taking $i = \ell = j - 1$ and rearranging terms. ■

Proof of Theorem 6 (Trigger Wage with Risk Neutrality):: Now suppose in the last theorem that u is linear. Then $\frac{u'(w_{\ell+1})}{u'(w_\ell)}$ can never be less than 1, and so the optimal wage schedule given there pays $w > w_{\min}$ only on a single cell $\ell^* + 1$.

If C_i lies below the top cell of $\mathcal{P}_*$, then by Theorem 2 (Lattice Structure), $f(C_i)/g(C_i) < f(C_k)/g(C_k)$ and the modified patched wage would be strictly better if $w_{C_i} > w_{\min}$, establishing the second claim of the theorem.

Finally, if the finite set of numbers $f(q_j), g(q_j)$ are chosen randomly according to the Lebesgue measure over the nonnegative numbers, then with probability one there will be a unique optimal titles partition, as mentioned earlier, and so $P_* = P^*$, implying that the optimal wage must be the one indicated by the theorem. ■

Proof of Theorem 7 (Trigger-Like Wages with Risk Aversion): Since u is strictly concave, by the Corollary to Theorem 4, w_d and w_e are uniquely defined. Let ℓ_e^* denote the last cell of $\mathcal{P}_*$ on which $w_e = w_{\min}$. Suppose that on cell $\ell_e^* + 1$, $w_d \geq w_e$. Then by Theorem 5 (Wage Structure with Risk Aversion), $w_d \geq w_e$ for all cells of $\mathcal{P}_*$ above $\ell_e^* + 1$. The reason is that for two such consecutive cells $\ell, \ell + 1$,

$$\frac{u'(w_{d,\ell+1})}{u'(w_{d,\ell})} = \frac{u'(w_{e,\ell+1})}{u'(w_{e,\ell})}$$

and so by the strict monotonicity of u', if $w_{d,\ell} \geq w_{e,\ell}$, then $w_{d,\ell+1} \geq w_{e,\ell+1}$. But this would imply that $w_d(x) \geq w_e(x)$ for all x, contradicting $d < e$. It follows that $w_{d,\ell_e^*+1} < w_{e,\ell_e^*+1}$. By precisely the previous logic, $w_{d,\ell} < w_{e,\ell}$ for all cells $\ell > \ell_e^*$. Moreover, we cannot have $w_{d,\ell_e^*} > w_{\min} = w_{e,\ell_e^*}$, for that would imply that $f(C_{\ell_e^*}) > g(C_{\ell_e^*}) \geq 0$, and that

$$\frac{u'(w_{e,\ell_e^*+1})}{u'(w_{e,\ell_e^*})} > \frac{u'(w_{d,\ell_e^*+1})}{u'(w_{d,\ell_e^*})} = \frac{[f(C_{\ell^*}) - g(C_{\ell^*})]/f(C_{\ell^*})}{[f(C_{\ell^*+1}) - g(C_{\ell^*+1})]/f(C_{\ell^*+1})} \geq \frac{u'(w_{e,\ell_e^*+1})}{u'(w_{e,\ell_e^*})}$$

a contradiction. Thus $w_d = w_{\min}$ on all cells of $\mathcal{P}_*$ below ℓ_e^*, concluding the proof of the first part of the theorem.

From the first part of the theorem and Theorem 5, whenever $x < y$ and $w_d(x) > w_{\min}$,

$$\frac{u'(w_d(x))}{u'(w_d(y))} = \frac{u'(w_e(x))}{u'(w_e(y))}.$$

Consider now for $\lambda > 1, w > 0$ that

$$d[\log u'(w) - \log u'(\lambda w)]/dw = \frac{u''(w)}{u'(w)} - \frac{u''(\lambda w)}{u'(\lambda w)}\lambda > 0$$

where the last inequality follows from strictly increasing relative risk aversion, as can be seen by multiplying the middle expression by w. It follows that if

$$\frac{w_d(x)}{w_d(y)} = \frac{w_e(x)}{w_e(y)}$$

then we would have

$$\frac{u'(w_d(x))}{u'(w_d(y))} < \frac{u'(w_e(x))}{u'(w_e(y))}$$

hence in order for there to be equality in the last expression, we must have

$$\frac{w_d(x)}{w_d(y)} < \frac{w_e(x)}{w_e(y)}$$

or

$$\frac{w_d(x)}{w_e(x)} < \frac{w_d(y)}{w_e(y)}$$

from which the theorem follows (keeping in mind that $w_d(y) > w_d(x)$ and $w_e(y) > w_e(x)$.■

Proof of Theorem 8 (Optimal Wage-Title Schedule): Consider the pair $(\mathcal{P}^*, w^*)$ where $\mathcal{P}^*$ is the maximal optimal titles partition and w^* is any solution to the pure wage problem with $d^* = d - I_\sigma(\mathcal{P}^*)$ in place of d. Then by Theorem 5 (Wage Structure with Risk Aversion), w^* is measurable with respect to $\mathcal{P}^*$. Then $(\mathcal{P}^*, w^*)$ is feasible, hence the total wage bill in w^* is at least as high as the total wage bill in the optimal w. It follows that $I_\sigma(w) \leq I_\sigma(w^*)$, since w^* is an optimal pure wage schedule. $\mathcal{P}$ generates status incentive $I_\sigma(\mathcal{P}) \leq I_\sigma(\mathcal{P}^*)$, since $\mathcal{P}^*$ is an optimal titles partition. Since $(\mathcal{P}, w)$ is optimal, we must have that the joint incentive $I_\sigma(\mathcal{P}) + I(w) \geq d = I_\sigma(\mathcal{P}^*) + I(w^*)$, and hence that $I_\sigma(\mathcal{P}) = I_\sigma(\mathcal{P}^*)$ and $I(w) = I(w^*)$. Thus $\mathcal{P}$ solves the pure titles problem and w solves the pure wage problem for $d^* = d - I_\sigma(\mathcal{P}^*) = d - I_\sigma(\mathcal{P})$.

For generic f and g, Lemma 3 (Unicity of the Lattice) implies that $\mathcal{P} = \mathcal{P}^* = \mathcal{P}_*$.■

Proof of Theorem 9 (Star Wages): Immediate from Theorem 8 (Optimal Wage-Title Schedule) and Theorem 6 (Trigger Wage with Risk Neutrality).■

Proof of Theorem 10 (Status Creates Star-Like Wages): Raising σ is tantamount to decreasing d in the pure wage problem, because the required wage incentive is given by $d - \sigma I^S$. The result now follows from Theorem 7 (Trigger-Like Wages with Risk Aversion).■

Proof of Lemma 4 (Cuts): Let us begin with a title partition defined by a finite set of cuts. Since in equilibrium there is nobody in any of the intervals (b_H^i, a_H^{i+1}), any cut in such an interval can be moved to a_H^{i+1} without hurting the status payoff of any worker, and leaving unchanged or perhaps hurting the status payoff of the unilateral shirker of type i. If there is not a cut at a_H^1, add it. This reduces (or leaves unchanged) the status payoff of the shirker of type 1 without changing the status payoff (under work or shirk) of anybody else, because in equilibrium there is nobody below a_H^1.

Now suppose there is a cut at a_H^1 and at least two cuts in some J_H^i. Remove the highest of all the cuts in J_H^i. Notice first that this does not affect the status payoff of any (worker or shirker) of type $j < i$, or of the shirker of type i, since they all come below the second highest cut in J_H^i anyway. Thus the status incentives of employees of types $j < i$ are unaffected. The status payoff of any shirker of type $j > i$ must go down by at least as much as that of the worker of the same type; hence their status incentives cannot decrease. Finally, the status payoff of a worker of type i can only go up. Against workers of his own type, he always gets expected status payoff of zero (by symmetry), and eliminating the cut increases (or leaves unchanged) his probability of coming equal with workers of higher type. This proves the lemma by iteratively removing all but the lowest cut from each interval J_H^i.■

Proof of Lemma 5 (Optimal Pure Titles): From Lemma 4 (Cuts) we can already assume that all the cuts are in the J_H^i, and that no J_H^i has more than one cut. If any J_H^i had no cuts, then the status incentive to work of employees of type i would be 0.■

Proof of Lemma 6 (Cuts with Wages): Suppose there is at least one cut in the interval (b_H^{i-1}, a_H^i). Take the topmost such cut and move it right to a_H^i. Set the wage for outputs between the old topmost cut and a_H^i equal to the wage on the left of the old topmost cut, and leave all other wages the same. This restores the measurability of wages. Moreover, this does not raise the wage of any unilaterally deviant shirker of type i, nor does it lower the wage of the worker of type i. At the same time, the status of the deviant shirker stays the same or goes down, while the status of the worker of type i stays the same. Thus the status incentive to work for type i is also not hurt. By iteratively moving cuts in this manner, we may assume that there are no cuts in any of the intervals (b_H^{i-1}, a_H^i).

From this point we can repeat the argument in the proof of Lemma 4 (Cuts) and show that in the new partition given there, the status incentive to work of every employee is improved (or held constant). If wages are secret, they need not change, and so total incentives have gone up or stayed the same. If wages are public, then the removal of the top cut in J_H^i, (as in the proof of Lemma 4) might require a change in the wage schedule to maintain measurability with respect to the titles partition. Replace the wages on the cells just below and above the removed cut by the average per capita wage over those two cells. This restores measurability of the wage schedule and leaves the wage bill unchanged. The expected wage of a worker of type i stays the same or increases. Hence his expected utility of working must go up or stay the same by concavity of his utility. His wage if he shirks is unaffected, hence his wage incentive to work rises or stays the same. For any employee of type $j > i$, his wage if he shirks either stays the same (in which case his working wage does too) or falls. If his working wage wage falls at all, it must have been the same as his shirking wage, and must fall by the same amount (with probability at most 1). Hence his incentive to work cannot go down. $\blacksquare$

Proof of Theorem 11 (Optimal Wage-Title Schedule with Secret Wages and Risk Neutrality): As was said before, Lemma 6 (Cuts with Wages) already guarantees that we need only consider at most one cut per J_H^i, defined by the vector $(p_{i_1} = p_1 = 1, p_{i_2} ..., p_{i_k})$ giving rise to cuts $c_{i_j}(p_{i_j})$ in $J_H^{i_j}$, where the first cut comes at a_H^1, corresponding to $i_1 = 1$

Define the expected number of people v_{i_j} in the region V_{i_j} from each cut c_{i_j} to the next cut c_{i_j+1} (assuming everybody works) by

$$\nu_{i_1} = N_1 + N_2 + ... + N_{i_2-1} + (1 - p_{i_2})N_{i_2}$$

and, for $2 \leq j \leq k-1$

$$\nu_{i_j} = p_{i_j} N_{i_j} + N_{i_j+1} + ... + N_{i_{j+1}-1} + (1 - p_{i_j+1})N_{i_{j+1}}$$

and, finally

$$\nu_{i_k} = p_{i_k} N_{i_k} + N_{i_k+1} + ... + N_{i_\ell-1} + N_{i_\ell}$$

Suppose now that we eliminate all the cuts except the one at a_H^1. The status payoff of the shirker and the worker of type $j \notin \{i_1 = 1, i_2, ..., i_k\}$ is unchanged.

The status payoff of the shirker of type 1 is unchanged, but the status payoff (and hence the incentive) of the worker of type 1 goes up by the expected number of workers above cut c_{i_2}

$$\nu_{i_2} + \nu_{i_3} + ... + \nu_{i_k}$$

since now when a type 1 employee works, he comes equal with all these other people.

For any employee of type $i_j \in \{i_2, ..., i_k\}$, the status incentive after the cuts are removed is zero! Prior to the removal, the status incentive of i_j was

$$p_{i_j}(\nu_{i_{j-1}} + \nu_{i_j} - 1)$$

because when an employee of type i_j worked, with probability $(1 - p_{i_j})$ he ended up with the same status as a shirker, and with probability p_{i_j} he gained status by outranking all the people in region V_{i_j-1} and coming equal with all the people in region V_{i_j} (not counting himself). Thus the loss in status incentive is $p_{i_j}(\nu_{i_{j-1}} + \nu_{i_j} - 1)$.

Given these changes in status incentives, it is possible to change the wages, in fact to lower the total wage bill, and yet leave all the employees with the same total incentive (i.e., status incentive plus wage incentive). First, recall that for outputs below $c_{i_1} = c_1 = a_H^1$, the wage is at $w_{\min}$. For outputs above c_{i_1} lower all wages by $\nu_{i_2} + \nu_{i_3} + ... + \nu_{i_k}$. This restores the original total incentive of all employees of types below i_2 and continues to leave unchanged the total incentive of each types $j \notin \{i_1 = 1, i_2, ..., i_k\}$. By assumption, the resulting wages must still be strictly above $w_{\min}$. Otherwise, the employees of type 1 would now be incentivized to work without a positive wage incentive (or indeed despite a negative wage incentive) contradicting our assumption that status incentive alone can never overcome the disutility of working).

For outputs above c_{i_2}, now raise all wages by $\nu_{i_1} + \nu_{i_2} - 1 < \nu_{i_1} + \nu_{i_2}$. This restores the total incentive of employees of type i_2 and leaves unchanged all other incentives. Successively raise all wages for outputs above c_{i_j} by $\nu_{i_{j-1}} + \nu_{i_j} - 1 < \nu_{i_{j-1}} + \nu_{i_j}$. As before, this restores the total incentive of employees of type i_j without changing any other incentives. Thus the new wage schedule gives all employees precisely the same total incentive as before.

We now show that the new wage schedule has a smaller total wage than the original. We compute the change in the wage bill by multiplying the number of workers by the change in their wages. The change in the total wage bill is thus strictly less than

$$\begin{aligned}
&- (\nu_{i_1} + \nu_{i_2} + \nu_{i_3} + ... + \nu_{i_k})[\nu_{i_2} + \nu_{i_3} + ... + \nu_{i_k}] \\
&+ (\nu_{i_2} + \nu_{i_3} + ... + \nu_{i_k})[\nu_{i_1} + \nu_{i_2}] \\
&+ (\nu_{i_3} + ... + \nu_{i_k})[\nu_{i_2} + \nu_{i_3}] + ... \\
&+ (\nu_{i_k})[\nu_{i_{k-1}} + \nu_{i_k}] \\
&= 0
\end{aligned}$$

This shows that the original wage-title schedule can be strictly improved by another wage-title schedule in which the title partition has just one cut at a_H^1.

But given a title partition with just one cut at a_H^1, it is evident (in view of our Necessity-of-Wages Assumption) that the optimal secret wage schedule is as stated.■

Proof of Theorem 12 (Exorbitant Elite Wages):: It will be useful to keep in mind throughout that raising p_i has the effect of raising $\tilde{I}_i$ and $\tilde{I}_{i+1}$ and lowering $\tilde{I}_{i-1}$ without disturbing other incentives. Similarly, raising w_i raises $\tilde{I}_i$ and lowers $\tilde{I}_{i+1}$, with no other effect.

We shall show inductively, starting with $i = 1$, that $p_i = 1$ for all $i = 1, ..., \ell - 1$. First note that, thanks to the Necessity of Wages Assumption, $p_1 > 0$ and $w_1 > w_0$. Suppose $p_1 < 1$. Define $\tilde{w}_1$ by $\tilde{w}_1 = (1 - p_1)w_0 + p_1 w_1$. Let the employer raise p_1 to 1 and lower w_1 to $\tilde{w}_1$, leaving all other p_i and w_i unchanged. Clearly this does not affect the wage bill. At the same time the status incentive of type 1 does not go down (indeed it goes up, unless $N_1 = 1$ and $p_2 = 1$ when it remains the same); both the status incentive and the wage incentive of type 2 go up (the first on account of the rise in p_1, and the second on account of the fall in w_1); the incentives of players of type $3, ..., l$ are undisturbed; and the wage incentive of type 1 does not go down (indeed it goes up if u is strictly concave) as the following calculation shows:

$$u(\tilde{w}_1) - u(w_0) \geq (1 - p_1)u(w_0) + p_1 u(w_1) - u(w_0) = p_1 [u(w_1) - u(w_0)]$$

To sum up, the employer's maneuver improves both status and wage incentives of type 2, without hurting any other incentives and without raising the wage bill. Next, the employer can decrease w_2 by a small ϵ, and thus lower the wage bill. The decrease of w_2 has just two effects on incentives: it raises the wage incentive of type 3 and lowers the wage incentive of type 2 and, other than this, has no effect on any other wage or status incentives (including the status incentive of type 2). For small enough ε, the incentive of type 2 will not fall below his original (pre-maneuver) level. Thus the principal does better, generating incentives that are no worse, for a lower wage bill, a contradiction. We conclude that $p_1 = 1$.

Inductively assume that $p_1 = \cdots = p_{i-1} = 1$ for $i < \ell$. If $p_i < 1$, we shall reach a contradiction by finding a cheaper way of providing the same incentives.

In what follows we shall be making *small* changes in wages to get from w_j to $\tilde{w}_j$,i.e., $|w_j - \tilde{w}_j| < \kappa\varepsilon$ for some constant κ and infinitesimal ε.So, denoting the derivative $u'(w_j) = \lambda_j$, we shall write[22] $u(w_j) - u(\tilde{w}_j) = \lambda_j(\ w_j - \tilde{w}_j)$ (see the Differentiability Assumption). Note that $\lambda_1 \geq \lambda_2 \geq ..... \geq \lambda_\ell$ since u is concave and since $w_1 \leqslant w_2 \leqslant .... \leqslant w_\ell$.

Set $\tilde{p}_i = p_i + \varepsilon$ and set $\tilde{w}_{i-1}$ to satisfy $\lambda_{i-1}(\tilde{w}_{i-1} - w_{i-1}) = \varepsilon N_i$, i.e., $\tilde{w}_{i-1} = w_{i-1} + \lambda_{i-1}^{-1}\varepsilon N_i$. Then the status incentive of $i-1$ goes down by $\varepsilon N_i p_{i-1} = \varepsilon N_i$, but his wage incentive goes up by the same amount: since $p_{i-1} = 1$, we have $p_{i-1}(u(\tilde{w}_{i-1}) - u(w_{i-2})) = u(w_{i-1}) - u(w_{i-2}) + u(\tilde{w}_{i-1}) - u(w_{i-1}) = u(w_{i-1}) - u(w_{i-2}) + \lambda_{i-1}(\lambda_{i-1}^{-1}\varepsilon N_i)$ $= p_{i-1}(u(w_{i-1}) - u(w_{i-2})) + \varepsilon N_i$.

Also, the status incentive of i goes up by

$$\Delta I_i(\varepsilon) \equiv \varepsilon[(N_i - 1) + (1 - p_{i+1})N_{i+1} + N_{i-1}].$$

This allows us to reduce his wage incentive by the same amount. So, set $\tilde{w}_i$ to satisfy $\tilde{p}_i \left[u_i(\tilde{w}_i) - u_i(\tilde{w}_{i-1})\right] -$

$$\tilde{p}_i\lambda_i(\tilde{w}_i - \tilde{w}_{i-1}) \equiv p_i\lambda_i(w_i - w_{i-1}) - \Delta I_i(\varepsilon), \text{i.e.,}$$
$$\tilde{p}_i(\tilde{w}_i - \tilde{w}_{i-1}) \equiv p_i(w_i - w_{i-1}) - \frac{\Delta I_i(\varepsilon)}{\lambda_i}$$

For small ε, $\Delta I_i(\varepsilon)$ is small, so $p_i(w_i - w_{i-1}) > 0$ implies that $\tilde{p}_i(\tilde{w}_i - \tilde{w}_{i-1}) > 0$, which in turn implies $\tilde{w}_i > \tilde{w}_{i-1}$, retaining the monotonicity of the revised wages. We shall be assuming ε small enough to guarantee monotonicity in all future wage revisions, without explicitly saying so.

Note that, since $w_i > w_{i-1}$ and $p_i + \varepsilon < 1$ (if $p_i < 1$ and ε is small) and $\lambda_{i-1} \geqslant \lambda_i$

[22] This to be understood as a first-order approximation, ignoring all higher-order effects. Strictly speaking we should replace λ_j with a number between $u'(w_j)$ and $u'(\tilde{w}_j)$. But the reader may easily check that our argument below holds, mutatis mutandis,with these strictly correct $\lambda's$ in place of ours (to represent exactly, rather than approximately, the changes in wage-utilities).

and $N_{i-1} \geq 1$,

$$\begin{aligned}
\tilde{w}_i &= \frac{p_i}{\tilde{p}_i} w_i + \tilde{w}_{i-1} - \frac{p_i}{\tilde{p}_i} w_{i-1} - \frac{1}{\lambda_i \tilde{p}_i} \Delta I_i(\varepsilon) \\
&= \frac{p_i}{\tilde{p}_i} w_i + w_{i-1} + \frac{\varepsilon N_i}{\lambda_{i-1}} - \frac{p_i}{\tilde{p}_i} w_{i-1} - \frac{1}{\lambda_i \tilde{p}_i} \Delta I_i(\varepsilon) \\
&\leq \frac{p_i}{p_i + \varepsilon} w_i + \left(1 - \frac{p_i}{p_i + \varepsilon}\right) w_{i-1} + \frac{1}{\lambda_i} \left(\varepsilon N_i - \frac{1}{p_i + \varepsilon} \Delta I_i(\varepsilon)\right) \\
&< w_i + \frac{1}{\lambda_i} \left(\varepsilon N_i - \varepsilon[(N_i - 1) + (1 - p_{i+1}) N_{i+1} + N_{i-1}]\right) \\
&\leq w_i - \frac{1}{\lambda_i} (\varepsilon (N_{i-1} - 1)) \leq w_i
\end{aligned}$$

Finally, the status incentive of $i+1$ goes up by

$$\Delta I_{i+1}(\varepsilon) \equiv \varepsilon p_{i+1} N_i.$$

Therefore the wage incentive of $i+1$ can be reduced by the same amount. So set $\tilde{w}_{i+1}$ to satisfy

$$\begin{aligned}
p_{i+1}\lambda_{i+1}(\tilde{w}_{i+1} - \tilde{w}_i) &= p_{i+1}\lambda_{i+1}(w_{i+1} - w_i) - \Delta I_{i+1}(\varepsilon), \text{i.e.,} \\
p_{i+1}(\tilde{w}_{i+1} - \tilde{w}_i) &= p_{i+1}(w_{i+1} - w_i) - \frac{\Delta I_{i+1}(\varepsilon)}{\lambda_{i+1}}.
\end{aligned}$$

Since $\tilde{w}_i < w_i$, clearly $\tilde{w}_{i+1} < w_{i+1}$. Hence recursively setting

$$\tilde{w}_j - \tilde{w}_{j-1} = w_j - w_{j-1} \text{ for } j > i+1$$

further lowers wages without changing incentives.

It remains to show that the wage bill defined in the owner minimization problem has gone down. The only terms that increase are

$$w_{i-1} N_i \text{ and } p_{i-1}(w_{i-1} - w_{i-2}) N_{i-1}$$

while many terms are reduced, including

$$p_i(w_i - w_{i-1}) N_i \text{ and } p_{i+1}(w_{i+1} - w_i) N_{i+1}.$$

The increases add up to

$$\frac{1}{\lambda_{i-1}}(\varepsilon N_i N_i + \varepsilon N_i N_{i-1}),$$

while just these two reductions add to

$$\begin{aligned}
&\frac{1}{\lambda_i} \Delta I_i(\varepsilon) N_i + \frac{1}{\lambda_{i+1}} \Delta I_{i+1}(\varepsilon) N_{i+1} \geq \frac{1}{\lambda_{i-1}} (\Delta I_i(\varepsilon) N_i + \Delta I_{i+1}(\varepsilon) N_{i+1}) \\
&= \frac{1}{\lambda_{i-1}} (\varepsilon N_i [(N_i - 1) + (1 - p_{i+1}) N_{i+1} + N_{i-1}] + \varepsilon p_{i+1} N_i N_{i+1}) \\
&= \frac{1}{\lambda_{i-1}} (\varepsilon N_i^2 + \varepsilon N_i N_{i-1} + \varepsilon N_i (N_{i+1} - 1)).
\end{aligned}$$

(the inequality following from the fact that $\lambda_{i-1} \geq \lambda_i \geq \lambda_{i+1}$ on account of the concavity of u). Since $N_{i+1} \geq 1$, the reduction is at least as big as the increase. But we have ignored many other strictly positive reductions (for example in $w_{i+1}N_{i+1}$). This contradiction proves that $p_i = 1$, for $i = 2, ..., \ell - 1$ and establishes part (a) of the theorem.

Now suppose $w_\ell < M$. Since we assumed $p_\ell(w_\ell - w_{\ell-1}) > 0$, clearly $p_\ell > 0$. Lower p_ℓ by ε. This raises the status incentive of type $\ell - 1$ workers by εN_ℓ, enabling us to lower the wage incentive for type $\ell - 1$ by the same amount.

Recalling that $p_{\ell-1} = 1$, and (by the Asymptotic Risk Neutrality Assumption) $w_{\ell-2} > w^\tau$ so that $\lambda_{\ell-1} = \lambda_\ell = \lambda$, set $\tilde{w}_{\ell-1}$ to satisfy

$$(\tilde{w}_{\ell-1} - w_{\ell-2}) = (w_{\ell-1} - w_{\ell-2}) - \frac{\varepsilon N_\ell}{\lambda}.$$

This drop in p_ℓ unfortunately lowers the status incentive of type ℓ by $\varepsilon(N_\ell - 1 + N_{\ell-1})$. Therefore we must raise the wage incentive of ℓ, choosing $\tilde{w}_\ell$ to solve

$$(p_\ell - \varepsilon)(\tilde{w}_\ell - \tilde{w}_{\ell-1}) = p_\ell(w_\ell - w_{\ell-1}) + \frac{\varepsilon(N_{\ell-1} + N_\ell - 1)}{\lambda}.$$

Fortunately, there is no group $\ell + 1$ to be affected by the change in p_ℓ, which is why it will turn out to be optimal to *lower* p_ℓ as long as $w_\ell < M$, whereas it was shown to be optimal to *raise* p_i all the way to 1 for any $i < \ell$.

Indeed the terms in the wage bill that change are

$$w_{\ell-1}N_\ell + p_{\ell-1}(w_{\ell-1} - w_{\ell-2})N_{\ell-1} + p_\ell(w_\ell - w_{\ell-1})N_\ell.$$

The net change in those terms, by our estimates above, is

$$\begin{aligned}&\frac{1}{\lambda}(-\varepsilon N_\ell^2 - \varepsilon N_\ell N_{\ell-1} + \varepsilon(N_{\ell-1} + N_\ell - 1)N_\ell)\\ &= -\frac{1}{\lambda}\varepsilon N_\ell < 0\end{aligned}$$

showing that the wage bill can be reduced, a contradiction. This proves that $w_\ell = M$.

Having proved that $p_i = 1$ for any $1 \leq i \leq \ell - 1$, it follows that the status incentives for $1 \leq i \leq \ell - 2$ are given by $I_i = I_\ell^i(1, ..., 1, p_\ell) = N_i + N_{i-1} - 1$. Hence the wages are recursively determined (starting from $w_0 = w_{\min}$) for $1 \leq i \leq \ell - 2$ by the equation

$$u(w_i) - u(w_{i-1}) = d_i - I_i = d_i - (N_i + N_{i-1} - 1).$$

Next, it will be convenient to scale the money by λ. Accordingly denote $w_j^* = \lambda w_j$, $M^* = \lambda M$ etc. Then we also have

$$w_{\ell-1}^* - w_{\ell-2}^* = d_{\ell-1} - [N_{\ell-2} + N_{\ell-1} + (1 - p_\ell)N_\ell - 1]$$

andϵ

$$M^* - w_{\ell-1}^* = \frac{d_\ell}{p_\ell} - [N_{\ell-1} + N_\ell - 1]$$

(recalling that $w_\ell \equiv M$,hence $w^*_\ell \equiv M^*$, in the last equation). We now show that there is a unique solution $w^*_{\ell-1}$, p_ℓ of these two simultaneous equations, so that the optimal wage schedule is determined uniquely. To do this, we add the two equations to get a convex quadratic in the single unknown p_ℓ. We then show that it has a positive value at $p_\ell = 0$ and a negative value at $p_\ell = 1$ and therefore a unique solution p_ℓ in $(0, 1)$.

More precisely, multiplying each equation by p_ℓ and then adding them yields

$$-p_\ell(M^* - w^*_{\ell-2} - d_{\ell-1}) - p_\ell(N_{\ell-2} + N_{\ell-1} - 1) - N_\ell p_\ell(1 - p_\ell) + d_\ell - p_\ell(N_{\ell-1} + N_\ell - 1) = 0$$

i.e. $f(p_\ell) = 0$, say. Clearly $f(0) = d_\ell > 0$. Also, since $\lambda w_i - \lambda w_{i-1} \leq u(w_i) - u(w_{i-1}) \leqslant d_i$ for $1 \leq i \leq \ell - 1$ (the first inequality following from the concavity of u, the second from the fact wage incentive *plus* title incentive equals disutility), we have

$$w^*_i \leq w^*_{\min} + d_1 + ... + d_i$$

for all $1 \leq i \leq \ell - 1$, and in particular for $i = \ell - 2$. This, in conjunction with $M^* > d_1 + ... + d_\ell$ and $N_i \geq 1$, implies $f(1) < 0$.

Finally, since the wage incentive of ℓ is at most d_ℓ, we have

$$p_\ell(M^* - w^*_{\ell-1}) \leq d_\ell,$$

hence

$$p_\ell \leq \frac{d_\ell + p_\ell w^*_{\ell-1}}{M^*} \leq \frac{d_\ell + w^*_{\ell-1}}{M^*} \leq \frac{w^*_{\min} + \sum_{i=1}^{\ell} d_i}{M^*}.$$

■

Proof of Corollary to Theorem 12: The incentive formulae are trivially generated by plugging $p_i = 1$ for $1 \leq i \leq \ell - 1$ into the status incentives for each agent, and by observing that $p_\ell \leq [w_{\min} + \sum_{i=1}^{\ell} d_i]/M \approx 0$ if M is large.

The wage differentials were explicitly computed in the proof of Theorem 12.■