Political Science and Political Economy Working Paper

Department of Government

London School of Economics

No. 5/2010

Launchpads and Graveyards: Career Profiles and the Design of Hierarchies

Klaus J. Broesamle (LSE)

Launchpads and Graveyards: Career Profiles and the Design of Hierarchies

Klaus J. Brösamle *

Version: May 2010

This paper presents a simple model showing that a number of institutional design features that are typical for bureaucracy-like public and private sector organizations may be conducive to efficient leadership selection. The practice of having positions with starkly varying workloads, task difficulties, and stress levels at comparable hierarchical levels—termed *asymmetric job design*—induces agents with unknown levels of career commitment and ability to self-sort into different career tracks. In equilibrium, committed and able employees choose "tough" entry-level positions. These turn into so called *launchpad posts* because the employer finds it optimal to recruit leaders primarily from these posts.

An extension of the model investigates the puzzle why bureaucracy-like organizations often retain and even lavishly aliment seemingly incapable employees by assigning them to obviously useless positions—often called *elephants' graveyards*: In the context of the presented sorting model this practice can sustain a separating equilibrium when this would be prohibitively costly without the use of graveyard posts. This result may be viewed as a rationale for tenure, generous early retirement schemes, and automatic promotion, all of which are still surprisingly common in public bureaucracies.

*I am thankful for helpful discussions with and support from Ian Jewitt, Meg Meyer, Lindsey Richardson, Charles Roddie, Gilles Serra, and seminar participants in Nuffield College and the Oxford Economics Department. All errors are of course my own. Financial support during the time of conducting this research from the the German National Merit Foundation and Oxford DPIR is gratefully acknowledged. E-mail: klaus.broesamle@nuffield.ox.ac.uk.

"Only the most committed diplomats apply for a stressful position, like head of policy planning for example, but afterwards they usually get very far."[1]

1. Introduction

This paper is motivated by explaining a set of observations about career profiles in large organisations and their 'hierarchical design'—that is, the structure of their jobs, ranks, and compensation. The first observation is that careers of successful employees—those who make it to the topmost positions—often lead via a limited set of particular posts. The opening quote mentions such a position in the UK Foreign and Common Wealth Office; most other hierarchical public or private sector organizations have their equivalents. I term these positions *launchpad posts*, because holding such a post usually launches a successful career. The second observation is that positions which are formally at the same hierarchical may differ substantially in their "toughness" in terms of workload, responsibility, exposure, and stress levels—think of personal assistants to Ministers compared to research or even back-office positions. I label this phenomenon *asymmetric job design* or *asymmetric hierarchies* and it implies the parallel co-existence of positions of differing toughness at similar hierarchical levels. The third phenomenon I want to draw attention to is *early career path differentiation*; a situation in which already the early career paths of successful employees resemble one another while being distinctly different from the early-phase career patterns of less successful employees. Again, a good example are diplomatic corps, where hardly any senior diplomat has not held an earlier assignment in one of the 'golden circle embassies'[2], while less outstanding employees often have a record of postings in more exotic countries.[3] Not surprisingly, the three phenomena are closely related and this paper presents a formal-theoretical explanation for why this may be.

To start with, one may ask which existing theories could explain any of the three observations observations. Launchpad posts and early career path differentiation are consistent with explanations based on human capital acquisition: managers-to-be may

[1]Quote from an anonymous senior UK diplomat, December 2008, Oxford.

[2]For the UK these are Washington, New York, Paris, Brussels, Moscow, Berlin and more recently also Beijing and Delhi.

[3]While I here only make brief reference to anecdotal evidence for these phenomena, more rigorous evidence is included in some of my empirical work which is still in progress.

simply require certain skills for their leadership task which they only acquire on one or a sequence of particular “training camp” posts. I cannot and do not want to challenge the validity of this view, but as this paper suggests it is surely not the only possible explanation. Launchpads and differentiated careers are also consistent with a situation where certain post are more sought-after than others—for whatever reason. Many employees will apply for them and the employer can “cream skim” (provided she has some way of screening applicants). As a result, whoever ends up on such a post must be suitable for future challenges. In other words: only rockets that fly make it to the launchpad. This explanation, however, is unlikely to be the whole truth because it is inconsistent with the general tendency that launchpad posts are usually the tougher and harder jobs. So while cream skimming may account for launchpad effects, the asymmetries in job attractiveness necessary for effective cream skimming point in the opposite direction of what is empirically plausible: working at the golden circle embassies is hardly known to be a holiday.

Asymmetries in job design are in line with human capital accumulation explanations, which can then also predict launchpad post effects and differentiated career paths. These explanations contend that employees learn leadership skills when doing particular jobs and therefore future leaders are recruited from these positions. However, for this story to work asymmetries are usually exogenously imposed. Some posts simply have to be better leadership training grounds than others and there are no obvious reasons why this should evolve endogenously. Moreover, human capital accumulation explanations can only account for *early* career path differentiation if on-the-job training asymmetries are stark already on the lowest levels of the hierarchy.

There are other theoretical explanations but they hardly explain all three phenomena in a coherent way. For example, the well-developed career tournament literature is silent on launchpads and differentiation: It does not predict in any way that tournament winners should be more likely to come from any particular post. If promotions in organizations were decided solely based on the outcomes of pure tournaments, there would be no launchpad posts and no career path differentiation. Regarding asymmetries, however, there exist tournament-based explanations which suggest the optimality of asymmetric job design in the form of ‘biased contests’ (Meyer 1991). This literature predicts that promising candidates should find promotion decisions biased against their less promising competitors—that is, life is made *easier* for likely future leaders. So just as with the cream skimming story above, the asymmetry created by optimally biased tournaments points in the opposite direction of what one observes empirically.

This paper provides an explanation for the joint existence of launchpad posts, asymmetric job designs and early differentiated career paths. I present a model of the careers of two employees in a simple hierarchy. It is a pure selection model: The employer seeks to design institutions and offer contracts that allow her to pick the best candidate for promotion to an important leadership position.

I first show that in the presence of weak contracting between the employer and employees, specifically if the employer cannot credibly commit to wages and promotion policies, hierarchies will be designed symmetrically, leadership selection will not be particularly precise, and career patterns of successful employees do not differ markedly from those of less-successful ones.

I then allow for a minor change in the contractability assumption, namely that the employer can credibly commit to attaching a wage to each job, similar to pay grade systems common in public and private companies. This changes the picture dramatically. The employer now finds it optimal to design entry-level jobs *asymmetrically*, which induces the heterogenous agents to self-sort into endogenously differing career tracks. Career path *differentiation* results already in the first period. Employees that choose the tougher career path are compensated with better career prospects subsequent to their high-workload assignment, which has the effect of turning the tough positions into *launchpad posts*.[4] Because in equilibrium the launchpad post is always chosen by the more able and committed types and the quieter career is favoured by the less eager types, also the quality of leadership selection is enhanced substantially.

However, such an equilibrium with self-sorting need not obtain under every circumstances. The cost of compensating the recruited managers for the strains of their earlier career may outweigh the economic benefits that it creates. Alternatively, it may be that managers' wages are regulated to be below a level that would provide sufficient sorting incentives. In an extension I investigate these situations and find that yet another set of quite common institutions may be suitable solutions to this problem: The practice of sending long-served employees into temporary retirement, or "pro"-moting them to so called *elephants' graveyard posts*—posts where long-served but not so committed or less able employees can be "buried" while apparently keeping their high-status.

There are not many convincing explanations for why this practice is so common, especially given its cost. The analysis here shows that retaining even less able employees may be optimal because it can help induce self-sorting in a more cost-effective manner than sacking them. Somewhat controversially this suggests that the tendency of public bureaucracies to retain and even lavishly aliment seemingly incapable bureaucrats

[4]Importantly, the employer finds it optimal to increase workloads on these positions even if this workload is "artificial" in that it has no actual positive output effects.

occupying obviously useless positions may in fact serve a purpose. It is just living up to a promise that was necessary to make them separate from the most capable bureaucrats earlier on—and that way they at least do not do any damage in positions that are actually important.

The model and its extension also make some predictions about optimal inter-temporal wage profiles. In any sorting equilibrium that obtains, high-ability employees are underpaid earlier in their career and overpaid later on—relative to the payoff they could realize on the outside labour market. In line with existing models that predict delayed compensation and efficiency wages (see e.g. Lazear 1981), early underpayment prevents high-ability types from quitting after period one. In this setting here, overpayment in later periods not only has an efficiency wage function but is even necessary to sustain the self-sorting of agents.

More of a corollary of the model is that the existence of launchpad posts and differentiated career paths make it rational for the employer to ignore any information on agents' abilities which she may obtain from performance appraisals or other screening mechanisms. Instead, she finds it optimal to tie promotion decisions solely to current assignments. The optimality of such a performance-blind promotions mechanism is interesting in public sector settings where performance monitoring is generally seen to be much harder. Another point to note here concerns the crucial difference between launchpad post effects and fast tracks effects: fast track effects, as they exist in the literature, are linked to the individual while launchpad effects are tied to particular positions. This difference is relevant for empirical testing.

Before I start, perhaps a few words are appropriate about what this paper does *not* do. The model does not try to explain why hierarchies exist. This question is treated in the vast literature on boundaries of the firm, and principal agent models more generally. I resort to presupposing that there is a management post for which it is worthwhile to strive for the best available leadership selection mechanism. Furthermore, the model does not genuinely explain substantial features of 'internal labour markets' in Doeringer and Piore's (1971) sense. While the model can be viewed as providing some insight into why wages in public sector organizations are tied to jobs which are then organized in grades, it does not directly speak to other features of internal labour markets, e.g. why mid-career entry and the external staffing of leadership positions tend to be uncommon. Answers to these questions have been addressed, albeit scarcely. Demougin and Siow (1994) derive these results in a scenario where firms train new workers internally. A non-formal treatment specifically about the public sector is Horn (1995: Ch. 5). Rauch (2001) shows more formally why long-term careers and internal promotions are useful in the presence of corruption.

The rest of the paper is organized as follows: Section 2 reviews the related literature, which is largely theoretical work in intra-organizational labour economics. Section 3 presents the main model and derives the findings on asymmetric hierarchies, launchpad posts, and early career path differentiation. Section 4 presents extensions including a variant of the model showing when using early retirement and elephants' graveyard posts may be optimal. Section 5 concludes after discussing predictions of the model that are empirically testable. While the model is quite intuitive, some algebra is relegated to the Appendix.

2. Related Literature

Viewed most generically, this paper's modeling is in the spirit of Spence's seminal paper on job market signalling (1973), where heterogenous agents choose different actions because they face different expected net utilities from them. Here, an additional actor is an employer whose problem it is to design a mechanism—that is a structure of the hierarchy, wage offers, and a promotion policy—which implement the self-sorting of the heterogeneous agents, because this allows the employer to pick the best one for an important management position. Such games where an uninformed actor moves before the informed agents are also known as screening games. I use the term sorting.

Viewed in a more nuanced way, the paper can be related to quite a number of sub-literatures within theoretical personnel economics. Miyazaki (1977) is related in that he makes the same informational assumptions as this paper: complete information asymmetry about agent types, whereby agents are informed perfectly and the employer knows nothing. Miyazaki's main interest then lies in explaining why a single firm simultaneously offers differentiated contracts to more and less able workers, rather than just selecting the best ones. He derives this result using a Wilson equilibrium concept and thereby provides an early explanation for the puzzle of one firm offering differentiated contracts (without any exogenously imposed task heterogeneity that would require the hiring of different workers). This can be viewed as a form of endogenous asymmetry which I am also interested in here. Miyazaki further claims to provide a rationale for 'internal labour markets' in the sense of Doeringer and Piore (1971). But in fact his paper does not directly speak to why one should observe an internal labour market with one or few points of entry at low levels and subsequent upwards-mobility of employees.[5]

For these phenomena, three lines of argument have been proposed: skill acquisition

[5] See also Miyazaki's discussion (1977: 415–416).

by employees (Becker 1962; Jovanovic and Nyarko 1997; Gibbons and Waldman 1999), screening of employees over time (related are Waldman 1984; Ricart i Costa 1988; Bernhard 1995), and internal training (as e.g. in Demougin and Siow 1994).

The result of job progression with increasing task difficulty features centrally in Jovanovic and Nyarko (1997) and Gibbons and Waldman (1999). Gibbons and Waldman produce the finding of employees moving up a job-ladder by assuming that experience is gained with tenure. Not too surprisingly, upwards movement results when the assumption is made that optimal initial assignments are at the lowest hierarchical level. Jovanovic and Nyarko develop what they call a 'stepping stone mobility' model. The stepping stone explanation's main mechanism is due to some jobs allowing to acquire skills which can be transferred on to the next assignment. By showing that there exist situations in which there is upward mobility in the sense that individuals choose simpler jobs first before moving on to more complex tasks, Jovanovic and Nyarko may be read as essentially laying the base for explaining career profiles based on human capital accumulation.

A third line of explanations for internal upward mobility is based on internal training. Demougin and Siow (1994) analyse a setting where a firm employs many young employees in order to turn one of them into a manager. They show that in some cases employers find it optimal to train only part of their young employees, who then have a high chance to get promoted if training is successful. Their predictions are closely related to this paper because they can explain early career path differentiation as well as the existence of launchpad posts. The main difference to my explanation of launchpad posts is that in Demougin and Siow's model launchpads would coincide with the most training-intensive positions, whereas I show that it should be the toughest posts which launch successful leaders—even in the absence of any training effects.

The practice of singling out some workers and providing them special training early on, as shown to be optimal by Demougin and Siow, is related to the debate about picking future leaders earlier or later in their career, and also to that on fast tracks of various forms.

A key theoretical paper in the 'pick early or late' debate, which also provides a good summary of the empirical background, is Prendergast (1992). More recently Chiappori, Salanié and Valentin (1999) and Ariga, Ohkusa and Brunello (1999) re-addressed the issue empirically. Chiappori, Salanié and Valentin (1999) use panel data on executive pay in a French state-owned firm to show that pay today depends largely on pay in the period before (working as a performance indicator). More importantly, they find evidence for the theoretical prediction of a learning model, whereby 'late beginners' have higher wages later in their career. Their findings are also relevant for the public sector because, as they point out, the theoretical model they use serves best the scenario of a closed career system with compensation according to seniority and rank. The model proposed here can be viewed as supporting the case for early leadership selection: If agents perfectly know their type, then making them reveal their type through self-selection cannot be worse than delaying the revelation of that information. However, this is only a clear-cut case because the model makes a stark informational assumption, abstracts from human capital acquisition effects and neglects issues of moral hazard, all of which may speak against 'pick and separate early' policies.

To turn to fast tracks, there are several reasons for why they may emerge: either because someone's early promotion speed signals high ability (see e.g. Bernhard 1995), because they have done well in earlier career tournaments (this will be discussed below), or because more able agents acquire human capital more quickly (Gibbons and Waldman 1999).[6] In any case, fast tracks and the concept of launchpads presented here are related. The key difference is that fast tracks effects are linked to individuals in most of the named explanations. In contrast, the career boosting effect of launchpad posts is different in that it is tied not to the individual but to a particular post.[7]

Lastly, one related sub-stream of literature should be mentioned which has a lot to say about endogenous asymmetries in job design, and which has been interpreted as a rationale for fast-tracks and special treatment. It mainly consists of two papers discussing 'biased contests' in repeated tournament settings (Meyer 1991, 1992). Meyer (1991) provides a rationale for institutionalizing fast tracks on the grounds of the employer trying to infer agents' abilities from coarse information. She shows that for an organization trying to infer agent ability from coarse information in a two-round career tournament, it is optimal to bias the second round of the tournament in *favour* of the first-round winner.[8] This has the effect that the second-round results can yield additional information, especially if the first-round loser wins the second round. This result is hardly challengeable on theoretical grounds. But as I argued already in the introduction, in real-world organizations it appears to be more common that promising employees (i.e. first-round winners), are actually *dis*advantaged, tested, and assigned to posts with direct exposure to the scrutiny of superiors. Admittedly, if bias takes the form of varying the levels

[6] For empirical treatments see Ariga, Ohkusa and Brunello (1999); Belzil and Bognanno (2004).

[7] The only other paper that can be viewed to predict career acceleration tied to particular positions is the one by Demougin and Siow discussed above, where among several homogenous employees some are randomly singled out to receive training and better subsequent career chances.

[8] In her second paper Meyer (1992) establishes a comparable result, however not only considering optimal selection but also considering whether this type of bias helps to induce optimal effort.

of supervision or monitoring, indeed Meyer's mechanism would imply that promising candidates should receive more attention, which is surely in line with launchpad posts often being very visible jobs such as personal assistant to a minister,[9] but the challenge remains to show why even for the best employees "working one's ass off" is a necessary evil in any high-flying career.

3. A Model of Hierarchical Design and Career Profiles

3.1. Model Setup

I model the careers of two risk neutral employees in a two-level hierarchy, who are hired for at least one and up to two periods. These periods represent the early and the later career phase in their professional life. The two agents differ in type $\theta \in \{H, L\},\ H > L$, which is best thought of as commitment to a high-flying career, leadership suitability, or simply ability. While the agents know their own type, it is initially unknown, not directly observable, and not verifiable to others. At the start of the game the two agents can be thought of as having been pre-selected for a career in the organization and now awaiting their first internal assignment to one of two entry-level jobs, A and B. Before they are assigned, the employer has discretion over the design of these jobs. Specifically, she can impose different workloads to make one tougher than the other. The employer also announces a promotion policy which specifies rules about promoting one of the agents to an important management position (C) in period two, and the wage that will be paid on that post. Importantly, the announced wages and promotion policy can be made contingent on observations that the employer makes throughout the game. (At this stage, I leave it open whether the employer can commit to the announced wages and promotion policy, that is in how far they are contractable.) Having observed the workload choices (and the offered wage and promotion policy, if commitment is assumed), both agents submit applications in which they state preferences for one of the two jobs or whether they are indifferent. First-round assignments take place, the agents work and are paid period-one wages. Production remains unobserved so that agent types remain unknown. But before promotions a performance appraisal takes place from which the principal imprecisely learns which agent is assigned to which post. The principal then promotes one agent to post C where he works and receives a wage. The other agent

[9] Viewed that way, the biased contest finding is very much complementary to the sorting mechanism I propose. Sorting will work best to select on employee characteristics which employees actually have superior information about. Running repeated and biased tournaments is helpful where even employees do not know their (relative) qualities.

leaves the organization.

The following paragraphs describe all the steps of the game in greater detail, and introduce somewhat more concise notation.

Players' Payoff Functions The employer's utility is given by the organization's output—which is additively separable in the outputs produced on individual posts—minus the wage bill:

$$U^E = aY_A - w_A + bY_B - w_B + cY_C - w_C,$$

where Y_Λ is output on post Λ, which is given by $f(\theta) = \theta$ for all posts. Parameters a, b, c indicate the importance to the employer of posts A, B, C, respectively, whereby $0 < a, b < c$. Thus, by assumption C is an important management position. w_Λ are wages paid by the employer for holding and working on post Λ. In line with much of the related literature the implicit assumption here is that wages are tied to jobs rather than negotiated individually.[10]
The employees' utility depends on wages and on the cost of working on the assigned post. Specifically,

$$U^\theta = w_{\Lambda 1} - \phi(\lambda_{\Lambda 1}, \theta) + w_{\Lambda 2} - \phi(\lambda_{\Lambda 2}, \theta)$$

where $\phi(\lambda_{\Lambda t}, \theta)$ is the cost that an agent of ability θ incurs from working on post Λ in period t, which has workload $\lambda_{\Lambda t}$.[11] I assume that for all θ $\phi(0, \theta)$ is some non-negative constant ϕ_0, and that first derivatives $\phi'_\lambda(L) > \phi'_\lambda(H) > 0 \ \forall \ \lambda$. In words this means that both agent types incur a cost of working that is strictly increasing in workloads, and that the rate of increase in λ is always higher for the L-type than the H-type agent. This also implies that for any given workload λ the L-type agent incurs a cost of working that is higher than that of the H-type agent.

Reservation Utilities on the Outside Labour Market Without limiting the insights gained from the model the outside labour market remains largely unmodelled. It is real-

[10]See e.g. Prendergast (1993). The discussion of why this wage-job link exists probably started with Williamson, Wachter and Harris (1975), who argued that attaching wages to jobs rather than individuals minimizes transaction cost in wage bargaining. Other rationales can be found in Waldman (1984) and Bernhard (1995) (asymmetric employer learning) or Lazear and Rosen (1981) and Malcomson (1984) (promotions as prizes in tournaments).

[11]Utility being linear in wages is admittedly a simplistic form of the utility function. While additive separability is required to keep the algebra manageable, for example allowing utility to increase in wages but at a decreasing rate would not substantially alter the results.

istic to assume that agents' private sector net reservation utilities, that is paid wages minus cost of working, are $R_t^H \geq R_t^L > 0$ for $t = 1, 2$ and $R_2^H > R_1^H$, while $R_2^L <> R_1^L$. The (strict) progression in H-type reservation utilities may be due to the market (symmetrically) learning agent types over time, or positive effects from human capital acquisition or seniority. Not restricting the over-time development of the L-type's reservation utility allows for various assumptions regarding what the outside labour market learns about the agents during period one. For example, one could imagine asymmetric learning so that after observing that the L-type agent was not promoted—as will often be the case in this model—outside opportunities may well worsen.[12]

Choice of the Career Path Design As already mentioned, the game starts with the employer having to make an important organizational design choice, which is non-reversible and observed by all actors. It concerns the workloads λ_A and λ_B on posts A and B, respectively. Their relative magnitude indicates whether the hierarchy is *symmetric* ($\lambda_A = \lambda_B$) or *asymmetric* ($\lambda_A \neq \lambda_B$). Workloads have no productivity or output effect for the employer. They solely serve as a strategic (negative) incentive device. However, posts A and B have a minimum workload of $\bar{\lambda}_A > 0$ and $\bar{\lambda}_B > 0$, respectively. While being quite realistic in any case, assuming minimum workloads mainly ensures that equilibrium wages and workloads are non-zero. The employer has no discretion over λ_C; it is always a parameter so that no special notation for its minimum value is needed.

An innocent but handy assumption is that the employer will always opt for a symmetric job design, unless she is strictly better off by choosing $\lambda_A \neq \lambda_B$. Furthermore, without any loss I restrict asymmetric choices to $\lambda_A < \lambda_B$. This restriction has the sole effect of saving me to characterize every asymmetric equilibrium twice, one being the mirror image of the other. Thus, from now on whenever one post is "tougher" than the other it is always post B, but it could just as well be A.

Agents' Applications and Assignments The process of agents applying for posts A and B and their first assignment is as follows: After having observed workload choices and the wage and promotion policy, each agent submits an application π^θ which can state *'application to A'*, *'application to B'*, or *'open application'*, that is in short $\pi^\theta \in \{A; B; O\}$. If an agent is *a priori* indifferent between posts I assume he always submits an open application. Agents submit applications simultaneously, and applications are observed by the principal. Since she does not know agent types, she observes both

[12] In fact, an earlier version of the model considered asymmetric learning of the outside labour market for both types of agents, but since it did not substantially affect any of the results I abandoned this complication.

applications jointly as an *unordered* subset

$$\Pi(\pi^H, \pi^L) \in \{(A, A); (B, B); (O, O); (A, B); (A, O); (B, O)\}.$$

I refer to the cases when $\Pi = (A, B); (A, O); (B, O)$ as 'separating application behaviour', while $\Pi = (A, A); (B, B); (O, O)$ is labeled 'pooling application behaviour'.

Performance Appraisal The performance appraisal at the end of period one produces an appraisal report Θ for the principal, which tells her—not always accurately—which agent is assigned to which post. Specifically, suppose the actual period-one assignments are $\Lambda^L = A, \Lambda^H = B$. Then with probability $q \in (\frac{1}{2}, 1)$ the principal receives the report $\Theta = [\Lambda^L = A, \Lambda^H = B]$ and with probability $1 - q$ she learns $\Theta = [\Lambda^L = B, \Lambda^H = A]$. If actual assignment are $\Lambda^L = B, \Lambda^H = A$ the signal probabilities are reversed. In other words, q is the probability with which the employer learns agent types *correctly* after period one. q is thus best thought of as a parameter capturing how well performance evaluation is possible in a given organization or task environment, and thereby captures one aspect of the severity of the information asymmetry between employer and employees (with a q close to one describing a weak informational asymmetry).

Promotion Policy, Wages, and Commitment The promotion policy announced by the employer is modelled as a probability $\alpha \in [0, 1]$ specifying whether after period one the agent who was assigned to post A will be promoted to post C (α) or whether the agent on post B is chosen ($1 - \alpha$). Commitment to this promotion policy is impossible in most of the paper, but this will be changed in an extension below.

Wages are tied to positions mainly because output is tied to holding a post for one period. Wages are agreed prior to, but paid after each period. The enforceability of wages agreed for the current period is always assumed. This is not self-evident but means nothing more than agents not having to worry about receiving their promised pay at the end of the period. Assumptions about the enforceability of promised wages for *future* periods, for example if a specific wage w_C is promised at the beginning of the game, will be altered during the paper because changing them yields some interesting insights.

Some Assumptions In order to keep the analysis focused on interesting scenarios I make the following assumptions.

The *exogenous symmetry assumptions* ensure that there are no exogenously imposed differences between posts A and B in the organization, which could drive possible results of endogenous asymmetries. This is important to avoid what many other models rely on: to simply assume differences and asymmetries between jobs. The assumptions concern the importance and workloads on the entry-level posts:

$$a = b,$$
$$\bar{\lambda}_A = \bar{\lambda}_B.$$

The *staffing assumption* states that even when the employer cannot do better than randomly pick one agent for promotion, she still prefers to staff post C rather than leave it unstaffed; that is

$$c\big(pH + (1-p)L\big) \geq R_2^\theta + \phi(\lambda_C, \theta), \tag{1}$$

for all θ and all $p \in [\frac{1}{2}, 1]$.

A crucial set of assumptions are the *careering assumptions* which are essential for actually creating the problem to be analyzed. They state that for any post in the hierarchy with minimum workloads chosen, if a high-type agent is paid at his reservation utility so that he is just indifferent between working in the organization or on the outside market, then the low-type is also at least indifferent between working inside or outside the organization. More formally

$$R_1^H - R_1^L \geq \phi(\bar{\lambda}_\Lambda, L) - \phi(\bar{\lambda}_\Lambda, H), \quad \text{for} \quad \Lambda = A, B \tag{2}$$
$$R_2^H - R_2^L \geq \phi(\lambda_C, L) - \phi(\lambda_C, H). \tag{3}$$

Perhaps the most arbitrary assumption I make is one that ensures *uniform entry-level wages*, that is

$$w_A = w_B \equiv w_1. \tag{4}$$

A pragmatic justification for this is that it greatly simplifies the analysis. More importantly, it will become obvious below that allowing entry-level workloads *and* wages to be chosen by the employer gives the model one excess degree of freedom. So restricting one of them is handy. Third, it is an empirically plausible assumption. The main explanation put forward by the literature for why wages are often tied to jobs or ranks rather than being individually negotiated is that employers economize on negotiation cost. This appears a suitable point for public bureaucracies.[13]

[13] Admittedly, it would be desirable to endogenously derive entry wage symmetry rather than assume

3.2. Solving the Game

As stated initially, the existence and properties of equilibria in this game depend on the assumptions one makes about the way in which the employer can commit to the period-two wage strategy. I will consider two scenarios:

1. *Minimal commitment:* The employer cannot make any credible promises about future promotions, assignments, or wages. Commitment is only possible for the immediate wage payment in each period.

2. *Pay-grade commitment:* The employer can commit to paying a wage w_Λ to anyone holding a particular post Λ, even if the assignment lies in the future. However, she cannot credibly commit to the promotion policy α.

Solution under Minimal Commitment Under the assumption of minimal contractability, the solution concept is to look for perfect Bayesian equilibria only. A strategy for the employer is

$$S^E(\lambda_A;\ \lambda_B;\ w_1;\ w_C|\Pi,\Theta;\ \alpha|\Pi,\Theta),$$

where α is the probability with which the employer promotes whichever agent is occupying post A. In other words, a strategy for the employer is to choose a set of workloads, wages, and a promotion rule α. Importantly, w_C and α can be conditioned on observing a particular pattern of agent applications, Π, and the performance appraisal report, Θ. A strategy for an agent is

$$S^\theta(\pi|\theta,\lambda_A,\lambda_B,w_1), \quad \pi\in\{A;B;O\}.$$

The employer and both agents choose their strategies so as to maximise expected utility. Let r and p be the probabilities that the H-type agent is assigned to post A, and that he is promoted to post C, respectively, both of which are to be determined in equilibrium,[14] then the players' maximisation problems are

$$\max_{S^E}\ E(U^E) = ra(H_A+L_B)+(1-r)a(L_A+H_B)+c\big[pH_C+(1-p)L_C\big]-2w_1-w_C,$$
$$\max_{S^H}\ E(U^H) = w_1-r\phi(\lambda_A,H)-(1-r)\phi(\lambda_B,H)+p\big[w_C-\phi(\lambda_C,H)\big]+(1-p)R_2^H,$$
$$\max_{S^L}\ E(U^L) = w_1-r\phi(\lambda_B,L)-(1-r)\phi(\lambda_A,L)+(1-p)\big[w_C-\phi(\lambda_C,L)\big]+pR_2^L.^{15}$$

it, but below I will discuss why it would most likely emerge in any case.

[14]Throughout the paper I will refer to p as the 'precision' or 'quality of leadership selection'.

In addition, the following *agent participation constraints* (APCs) must be satisfied:

$$APC^H: \quad E(U^H) \geq R_1^H + R_2^H, \tag{5}$$

$$APC^L: \quad E(U^L) \geq R_1^L + R_2^L. \tag{6}$$

Solving leads to a first proposition.

PROPOSITION 1: *Without commitment to future compensation, there exists a unique pooling perfect Bayesian equilibrium in which the employer chooses a symmetric job design ($\lambda_A^* = \lambda_B^* = \bar{\lambda}_A$). Both agents submit open applications ($\Pi^* = (O, O)$) so that their application behaviour is uninformative and each agent has equal chance to be staffed to either post ($r^* = \frac{1}{2}$). The employer promotes according to the performance appraisal report so that the quality of leadership selection is equal to that of the appraisal mechanism ($p^* = q$).*

Proof: *See Appendix, if not obvious.*

The full equilibrium strategies are in the Appendix, but let me briefly state the equilibrium wages. The wage on post C is

$$w_C^* = R_2^H + \phi(\lambda_C, H) \;\; \forall \;\; \Pi, \Theta. \tag{7}$$

This is because without credible commitment to future wages, any wage promised by the employer that is higher than what is needed to just retain the H-type, i.e. his reservation utility $R_2^H + \phi(\lambda_C, H)$, is not credible. Entry-level wages follow from (5) and (7):

$$w_1^* = R_1^H + \phi(\bar{\lambda}_A, H), \tag{8}$$

which by the careering assumption (2) means that the L-type agent also participates.

There are a few insights to be gained from this first scenario. First of all, it may appear surprising that no separating equilibrium obtains. Given the assumptions about the functional form of the cost of working ϕ, one would think the employer could simply raise the workload on post B until only the high-type wants to do this job at the offered wage. This cannot occur because the employer cannot credibly offer a w_C above

[15] Here I have retained post suffices for post outputs H and L. This is only to facilitate orientation in the algebra. Output on a post always equals θ, meaning that e.g. $aH_A + bL_B + cH_C$ is the same as $(a + c)H + bL$.

the H-type agent's reservation utility such that he actually prefers the post sequence $B \rightarrow C$. Second, in environments with weak contractability the quality of leadership selection in organizations is strictly limited by the ability of the employer to effectively screen employees. Here, this is captured by parameter q, indicating the severity of the information asymmetry. Third, note that in this setting there is no difference in the *a priori* probability with which future managers should originate from post A or post B ($r^* = \frac{1}{2}$). In other words, there are no signs of career path differentiation, and launchpad posts do not exist.

Treating this first scenario as a benchmark case, I now analyse under which circumstances it may be possible to enhance leadership selection beyond q. A slight change in the contractability assumptions allows for agent self-sorting into posts with differing workloads. The necessity of commitment for this result to obtain is interesting because it suggests another reason why (public) bureaucracies often commit to pre-defined pay scales.

Solution under Pay-Grade Commitment In this section I assume that the employer can commit today to the wages she is willing to pay tomorrow. Such commitment may be possible for various reasons. The most obvious is effective contract enforcement by the courts. But also more subtle incentives such as reputational concerns of the employer can achieve this effect.[16] The analysis will show that the ability of the principal to overpay the promoted agent allows to induce self-sorting into asymmetrically designed career paths. The credibly promised high wage for high-level management positions can now compensate the promoted agent for the pains of the high cost of working earlier in his career.

The setup of the game remains completely unchanged in this section, except for the new commitment assumption. Importantly, this does not change the fact that the employer can make wages and promotions contingent on observing the performance appraisal and application behaviour—her choices just no longer need to be consistent with her posterior beliefs. For this reason, the agents's strategies are now also conditional on the announced wage w_C:

$$S^\theta(\pi|\theta, \lambda_A, \lambda_B, w_1, w_C) \quad \pi \in \{A; B; O\}.$$

Solving under the new commitment assumption leads to a second proposition.

[16] To model this more rigorously, an infinitely repeated game could be constructed but the commitment assumption is an efficient way to avoid such complexity.

PROPOSITION 2: *For realistic parameter values there exists a separating Bayesian Nash equilibrium in which the employer chooses an asymmetric job design* ($\lambda_B^* > \lambda_A^* = \bar{\lambda}_A$)*, the H-type agent is staffed to the tougher post B* ($r = 0$) *and subsequently always promoted to post C* ($p = 1$)*, while the L-type agent leaves the organisation after working on post A for one period.*

Proof: *See Appendix.*

The complete equilibrium strategies are given in the Appendix due to their length. Here I settle for giving an intuition and deriving the equilibrium wages and workloads. The main clue which creates this separating equilibrium lies in the employer's strategy. She makes the certainty of promotion from post B strictly conditional on observing separating application behaviour. More formally, she only chooses $\alpha = 0$ conditional on $\Pi = (A, B); (A, 0); (B, 0)$, although $(A, 0); (B, 0)$ never occur. For all other Π promotion is made dependant on what the performance appraisal report recommends. That way, as soon as either agent deviates from applying to the post that he is "supposed" to choose, future career prospects change and make everyone worse off.

Regarding equilibrium workloads and wages, a first finding is that λ_A^* is always set at its minimum $\bar{\lambda}_A$. Any higher workload on post A would only drive up the overall wage bill and at same time offset any separation-inducing increase in λ_B.

Given that $\lambda_A^* = \bar{\lambda}_A$, the wages and workload that the employer chooses in her equilibrium strategy are now given by the combination of w_1, w_C, λ_B, which jointly minimizes the wage bill of the organization, $2w_1+w_C$, whereby five conditions have to be met. These are two conditions ensuring that the agents prefer separating application behaviour over pooling application behaviour, the two agent participation constraints, and a condition stating that sorting actually makes the employer better off than she would be under the pooling equilibrium:

$$w_C \geq \frac{\phi(\lambda_B, H) - \phi(\lambda_A, H) + 2\epsilon}{2 - 2q} + \phi(\lambda_C, H) + R_2^H, \tag{9}$$

$$w_C \leq \frac{\phi(\lambda_B, L) - \phi(\lambda_A, L) - 2\epsilon}{2 - 2q} + \phi(\lambda_C, L) + R_2^L, \tag{10}$$

$$R_1^L \leq w_1 - \phi(\bar{\lambda}_A, L), \tag{11}$$

$$R_1^H + R_2^H \leq w_1 - \phi(\lambda_B, H) + w_C - \phi(\lambda_C, H), \tag{12}$$

$$c(1 - q)(H - L) \geq 2(w_1 - w_1^{p*}) + (w_C - w_C^{p*}), \tag{13}$$

where w_C^{p*} and w_1^{p*} are the equilibrium wages that obtain in the pooling equilibrium above (equations 7 and 8). ϵ is some small value which is merely introduced to avoid that agents are entirely indifferent between the outcomes from sorting or non-sorting behaviour, in which case they would submit an open application.

The optimal workload λ_B^{s*} results from equating the right hand sides of (9) and (10), so that λ_B^{s*} is chosen such that the following holds:

$$\begin{aligned} \phi(\lambda_B, L) - \phi(\lambda_B, H) = (2-2q)\big[R_2^H - R_2^L - \phi(\lambda_C, L) + \phi(\lambda_C, H)\big] + \\ \phi(\bar{\lambda}_A, L) - \phi(\bar{\lambda}_A, H) + 4\epsilon, \end{aligned} \tag{14}$$

which exists uniquely given the initial assumptions about ϕ, and it must strictly greater than $\bar{\lambda}_A$.[17] Denoting the separating equilibrium with superscript $e*$, the second-period wage w_C^{s*} follows from (9):

$$w_C^{s*} = \frac{\phi(\lambda_B^{s*}, H) - \phi(\bar{\lambda}_A, H) + 2\epsilon}{2-2q} + \phi(\lambda_C, H) + R_2^H, \tag{15}$$

and the optimal entry-level wage from (12) with the lower bound (11), which yields

$$w_1^{s*} = \max\left\{ R_1^H + \frac{\phi(\bar{\lambda}_A, H) - (2q-1)\phi(\lambda_B^{s*}, H) - 2\epsilon}{2-2q} \quad ; \quad R_1^L + \phi(\bar{\lambda}_A, L) \right\}. \tag{16}$$

If finally

$$c(1-q)(H-L) \;\geq\; 2(w_1^{s*} - w_1^{p*}) + (w_C^{s*} - w_C^{p*}), \tag{17}$$

then the equilibrium characterized in Proposition 2 exists, which is the case for large $H - L$, and/or not too large $R_2^H - R_2^L$ and λ_C.

This finding deserves some more interpretation and discussion which the next section provides. It also relates the findings back to organizational reality.

3.3. Asymmetric Job Design, Launchpads, and Differentiated Career Profiles

The model has shown that under plausible assumptions there exist mechanisms which the employer can use in order to make agents self-sort into different career paths according to their privately known career commitment or ability types. The main instruments to achieve this are choosing particular workloads on entry-level jobs, recruiting leaders from very particular positions and a wage profile which exhibits delayed compensation.

[17] Expression (14) essentially means that λ_B is chosen so as to satisfy the Spence-Mirrlees sorting condition of the game.

This creates an organizational setting in which the three phenomena mentioned in the introduction occur jointly. Different jobs at the same hierarchical level are designed *asymmetrically* in terms of their workloads. The equilibrium promotion policy turns the tougher positions into *launchpad posts* because it is the more able and committed agents that choose them. This, in turn, results in career paths being *differentiated* already in period one: All successful employees start their career on post B.

All this has resemblance to institutional arrangements one can observe in real-world organizations, especially public bureaucracies. Many of them have more and less stressful and tough positions at the same hierarchical level—think of personal assistants to Ministers who are often quite junior in terms of age, rank and pay grade, but work much harder than their peers on "normal" positions. This is also reflected strikingly well in the epigraph by a senior UK diplomat: "Only the most committed diplomats apply for a stressful position, like head of policy planning for example...".

The second part of the statement, "...but they usually get very far afterwards", expresses the other key aspect of organizational design which the model explains: the fact that it is usually the tough and stressful positions that launch careers. The employer knows that only agents with high ability, good leadership suitability and high commitment choose these jobs, and therefore she finds it optimal to promote them to higher levels. An obvious empirical correspondence is that there are few successful UK diplomats who never had an assignment in Paris, Washington or New York earlier in their career.

The model yields another insight: The equilibrium promotion policy suggests that it is optimal for the employer to ignore any performance information she obtains from appraisal reports, but instead choose leaders solely on the basis of their earlier assignment. This appears like a corollary but it hints at the potential importance of hierarchical asymmetries to induce self-sorting in organizational settings where talent uncertainty is high and performance is hard to measure. Seemingly performance-blind promotion may therefore not be inefficient at all.

Finally, it is worth commenting on equilibrium wages. From equation (15) one can see that the H-type agent is overpaid in period two—relative to his outside labour market opportunities. In turn, from (16) it follows that he is underpaid in period one.[18] The effect of such a delayed compensation wage profile is retention. High-type agents do not

[18]To see this it helps to rewrite (16) as

$$w_1^{s*} - \phi(\lambda_B^{s*}, H) = R_1^H - \frac{\phi(\lambda_B^{s*}, H) - \phi(\bar{\lambda}_A, H) - 2\epsilon}{2 - 2q}.$$

The H-type is always underpaid unless $w_1^{s*} = R_1^L + \phi(\bar{\lambda}_A, L)$ and $R_1^L + \phi(\bar{\lambda}_A, L) < R_1^H + \phi(\lambda_H^{s*}, H)$.

leave immediately after having been "exploited" early in their career if they can expect higher compensation later on. This is well in line with existing models predicting that wage profiles inside firms are steeper than actual productivity increases (see e.g. Lazear 1981).[19]

Before moving on, a valid plea against this model's story is that compensation for the high workloads necessary to induce sorting need not be promotion in the subsequent period. As in any standard sorting model, the same outcome could be achieved by attaching differentiated wages and workloads to positions A and B in the first period. And indeed, it is not hard to show that there exists a combination of wages and workloads which leads to perfect sorting.[20] In that case it would not even be necessary to make the commitment assumptions introduced earlier.

However, in line with many relevant real-world settings, the situation analyzed here is assumed to require that one of the agents is retained for a second period. Therefore, the problem with offering period-one wages and workloads that ensure sorting would be that they do not exhibit any delayed compensation so that high-type agents have no incentive to stay in the organization. In addition, any promotion offer to the high-type agent will communicate his type to the outside labour market, upon which he receives high wage offers from other employers. As a result, his first employer would have to overpay him in any case, but then without having underpaid him before.

This in fact also provides an *ex post* justification for assuming that entry-level wages are equal. It is due to this assumption that the high-ability agent is underpaid in period one, which facilitates retention. Admittedly, delayed compensation is also possible when $w_A \neq w_B$, but together with considerations of equity and fairness, minimizing transaction cost from wage negotiations, or in the case of public bureaucracies to simply comply with legal requirements, setting $w_A = w_B$ a justifiable restriction.

[19] With a view to the realities of public sector pay, however, this result appears inappropriate. Studies abound that show that, if at all, pay in the public sector is below private sector levels, especially at the management level (see e.g. Hood and Lodge 2006: 39–43). But bearing in mind that the model here makes predictions about net utilities which factor in potential differences in workloads, it is less surprising that pecuniary compensation in the private sector is much higher since private sector workloads also tend to be higher. Other obvious reasons for higher private sector wages are premia for the risk of jobloss or the absence of mission-related utility (see e.g. Besley and Ghatak 1995).

[20] Such a contract would specify $\lambda_A < \lambda_B$ and $w_A < w_B$ such that $U^H(w_B, \lambda_B) > U^H(w_A, \lambda_A)$ and $U^L(w_B, \lambda_B) < U^L(w_A, \lambda_A)$, which exists given the assumptions about the cost function.

4. Extensions

The last section showed that heterogeneous agents can be induced to self-sort into career tracks through a mechanism that uses heavy workloads together with improved career prospects and delayed compensation on selected entry-level positions. This section will present some extensions and refinements. First, I will consider what happens if the maximum workload that an employer can demand from an agent is capped, which may be the case for technical or legal reasons. Then, I will look at how practices such as early retirement, tenure, and demotions to *elephants' graveyard posts* relate to sorting. Finally, I discuss very briefly what happens if human capital acquisition effects are introduced into the model.

4.1. Capped Workloads and Regulated Wages

The occurrence of sorting in the previous section hinged on the ability of the employer to levi heavy workloads onto some posts in the hierarchy. As can be seen from equation (14) the workloads necessary to induce sorting may be excessive, especially when the information asymmetry is severe (q close to $\frac{1}{2}$). Agents then need to be compensated with high wages for the high cost of working they incur on the launchpad posts. Therefore, a situation may arise in which the workloads or wages necessary to induce sorting reach some upper limit of feasibility. This may be for practical/technological or legal reasons. In most sectors working hours are limited at least on paper, and especially in public bureaucracies wages are often set by law. This section briefly considers what happens either if workloads are capped or if wages are regulated.

Consider that the maximum workload the employer can choose is $\hat{\lambda}_B > \bar{\lambda}_B$. Assume further that

$$R_1^H - R_1^L \geq \phi(\hat{\lambda}_B, L) - \phi(\hat{\lambda}_B, H),$$

which is essentially a variant of *careering assumption* (2), now specifying that even if the employer levies the maximum workload on post B the equilibrium characterized in Proposition 2 will not exist.

Alternatively, consider the situation in which the managers' wages w_C are fixed at $\bar{w}_C$, with

$$\bar{w}_C > R_2^H + \phi(\lambda_C, H) \quad \text{but} \quad \bar{w}_C < w_C^{s*},$$

so that again the equilibrium characterized in Proposition 2 cannot exist.

Using the setup from above where commitment to future wages is possible, it turns out that for both scenarios there continues to exist a separating equilibrium in which agents still self-sort into posts A and B. While a detailed formal discussion of these cases can be found in the Appendix, the main differences to above are as follows. Leadership selection and thereby career path differentiation are no longer perfect because the optimal strategy of the employer is now mixed, that is she randomizes over promoting the agent on post A and the one on post B. The equilibrium probability with which the L-type agent is promoted, α^*, is strictly greater than zero. However, the probability that the H-type agent is promoted, $1 - \alpha^*$, is strictly greater than q. Thus, leadership selection is still superior to any pooling equilibrium in which the employer has to rely on the performance appraisal mechanism, but it is inefficient compared to perfect sorting.

Unfortunately, this result hinges on a debatable assumption. In order to induce sorting the employer needs to employ a mixed promotion strategy with $\alpha > 0$. She can only do so if credible commitment to such a strategy is possible because any $\alpha > 0$ is not consistent with her posterior equilibrium beliefs about the assignments of the agents. As a result she sometimes ends up promoting the L type even though she knows he is less able. But committing to this practice is the only way to distinguish the two, and that way she ends up promoting the H-type at least with an *ex ante* probability greater than q.

4.2. Tenure, Pension, and Elephants' Graveyards

A separating equilibrium like the one described in Proposition 2 may not evolve for other reasons. Specifically if $H - L$ is small or $R_2^H - R_2^L$ and λ_C are large then inducing agents to self-sort requires workloads and wages that are simply too expensive compared to the productivity gains from self-sorting.

Perhaps somewhat counter-intuitively, this section shows that a number of seemingly inefficient institutional arrangements which are nevertheless very common in large (public sector) organizations may help to overcome this problem. In particular, institutions like secure tenure, "pro"-motions to *elephants' graveyard posts*, early retirement, and generous pensions even for less able employees can have that effect.[21] This model extension shows that retaining the L-type agent after period one and even paying him a wage above his reservation utility R_2^L can weaken the conditions under which sorting occurs. The employer can then choose less severely asymmetric workloads and thereby

[21] The latter two forms of "retention" are especially common in the Japanese government and they would be better described as *golden* graveyards, because there even very successful civil servants 'descend from heaven' to retire on private sector posts where they earn most of their lifetime income (Ramseyer and Rosenbluth 1993: 117).

lower the overall wage bill. With reference to the previous sub-section, the mentioned institutional instruments can also be useful in situations where sorting is imperfect due to capped maximum workloads or regulated wages. Designing jobs less asymmetrically, paying lower wages to managers, and instead giving tenure to L-type agents can avoid hitting the upper bounds of wages or workloads and the precision of leadership selection is no longer impaired.

To allow for retention of both agents, the model hierarchy is extended by an additional post D, on which wage w_D is paid. For simplicity, I assume the cost of working on this post, $\phi(\lambda_D, \theta)$, to be zero for both agent types. This avoids introducing additional parameters or choice variables into the model. It is even realistic if post D is thought of as early retirement or an elephants' graveyard post where tasks are rarely daunting. The extension requires a slightly altered setup of the game. The employers' utility is now

$$U^E = aY_A - w_A + bY_B - w_B + cY_C - w_C + dY_D - w_D,$$

where is is realistic to assume that d is small. A strategy for the employer is now

$$S^E(\lambda_A;\ \lambda_B;\ w_1;\ w_C|\Pi, \Theta;\ w_D|\Pi, \Theta;\ \delta|\Pi, \Theta),$$

where $\delta \in \{0, 1\}$ is the probability with which the employer promotes the agent on post A to post D. At the same time δ indicates the probability with which the agent on post B is promoted to post C. $1 - \delta$ is then the probability with which the agent on A is promoted to C, and the agent on B is promoted to D.[22]

A new assumption is the *selective retention assumption*, which states that the employer always prefers to fire the agent that was not promoted after period one, rather than to retain him on post D and pay wage w_D. It states

$$dY_D(\theta) - w_D < 0 \ \ \forall\, \theta, \tag{18}$$

which is in line with d being small. This assumption implies that the only reason why the principal may have an interest in offering any period-two assignment and wage to the non-promoted agent is in order to secure gains from self-sorting. The rest of the

[22]Restricting promotion to pure strategies, i.e. $\delta \in \{0, 1\}$, does not affect the equilibrium of the game as long as w_D remains a choice variable. An earlier version of the model assumed that $w_C = w_D$, in which case allowing for mixed promotion strategies is necessary to obtain a separating equilibrium. The basic insights, however, remain the same and the version presented here is somewhat more elegant. The earlier version is available from the author.

game remains the same, including the commitment assumption, which makes a tenured contract (spanning two periods) a credible offer. The key insight from solving is captured in a third and final proposition:

PROPOSITION 3: *For certain parameter values there exists a separating Bayesian Nash equilibrium in which the employer chooses a weakly asymmetric job design ($\lambda_B^* = \bar{\lambda}_A + \epsilon > \lambda_A^* = \bar{\lambda}_A$), the H-type agent is staffed to the tougher post B ($r^* = 0$), and he is always promoted to post C ($p^* = 1$). The L-type agent is first staffed to post A and subsequently retained and paid a wage w_D greater than his reservation utility R_2^L.*

Proof: *See Appendix.*

Again, the full equilibrium strategies are given in the Appendix. Here I first derive the properties of the equilibrium in which the D-post is being used. Then I study when the "graveyard equilibrium" dominates the standard separating equilibrium.

Just as above, because the optimal λ_A is always $\bar{\lambda}_A$, a general solution to the problem can be formulated as any combination of w_1, w_C, w_D, λ_B minimizing the wage bill of the organization, which is now $2w_1 + w_C + w_D$. This happens subject to conditions which differ slightly from those above:

$$w_C \geq \frac{\phi(\lambda_B, H) - \phi(\lambda_A, H) + 2\epsilon}{2 - 2q} + \phi(\lambda_C, H) + R_2^H, \tag{19}$$

$$w_C \leq \frac{\phi(\lambda_B, L) - \phi(\lambda_A, L) - 2\epsilon}{2 - 2q} + \phi(\lambda_C, L) + \frac{w_D - qR_2^L}{1 - q}, \tag{20}$$

$$R_1^L + R_2^L \leq w_1 - \phi(\bar{\lambda}_A, L) + w_D, \tag{21}$$

$$R_1^H + R_2^H \leq w_1 - \phi(\lambda_B, H) + w_C - \phi(\lambda_C, H), \tag{22}$$

$$c(1 - q)(H - L) + dL \geq 2(w_1 - w_1^{p*}) + (w_C - w_C^{p*}) + w_D, \tag{23}$$

where again w_1^{p*} and w_C^{p*} are the equilibrium wages that obtain in the pooling equilibrium characterized in Proposition 1.

When deriving the first equilibrium condition from (19) and (20), one can see that it now contains two choice variables of the principal, representing two instruments to satisfy the condition. The first is an increase in the workload λ_B, just as before. The second is to increase w_D, which means that λ_B can be chosen strictly closer to $\bar{\lambda}_A$. The choice between the instruments depends on their relative cost. In order for the new instrument to be used is must be that increasing w_D by one small unit decreases the wage bill of the organization by more than that. This can work for some parameter values because a higher w_D allows for a smaller λ_B, which in turn means that all the other wages can be lowered. In equations, this is the case if

$$-\frac{dw_D}{d\lambda_B} > \frac{dW}{d\lambda_B}, \tag{24}$$

where $W = 2w_1 + w_C$ is the total wage bill in the normal sorting equilibrium. Calculating this one finds that increasing w_D is more than offset by a lower wage bill when agents operate on a rather flat part of the cost function ϕ, and if q is not too close to one. Specifically, if

$$\frac{\phi'(\lambda_B, L) - \phi'(\lambda_B, H)}{\phi'(\lambda_B, H)} > \frac{3-4q}{1-q}. \tag{25}$$

If (25) does not hold, using tenure or pensions is the more expensive instrument to induce sorting, in which case the separating equilibrium described in Proposition 2 will emerge. The "optimal" w_D will then be smaller than the L-type agent's reservation utility R_2^L so that he would not accept the retirement offer but instead continue to work on the outside labour market.

Note that (25) is only a necessary condition for the use of tenure, graveyards and the like. There are two sufficient conditions which have to be met. The first ensures that the elephants' graveyard equilibrium dominates the pooling one, and the second is that it also dominates the ordinary separating one.[23] To derive these conditions let me first characterize the wages and workloads supposing that a graveyard equilibrium exists.

As always, $\lambda_A^* = \bar{\lambda}_A$, but now λ_B will be chosen just above $\bar{\lambda}_A$, i.e.

$$\lambda_B^{e*} = \bar{\lambda}_A + \epsilon, \tag{26}$$

where superscript $e*$ indicates "*e*lephants' graveyard equilibrium". The optimal w_D is

$$\begin{aligned} w_D^{e*} =& \frac{1}{2}\big[\phi(\bar{\lambda}_A, L) - \phi(\bar{\lambda}_A, H) - \big(\phi(\bar{\lambda}_A + \epsilon, L) - \phi(\bar{\lambda}_A + \epsilon, H)\big)\big] + \\ & (1-q)\big[\phi(\lambda_C, H) - \phi(\lambda_C, L) + R_2^H\big] + qR_2^L + 2\epsilon. \end{aligned} \tag{27}$$

By expressions (19) and (26) the wage on post C is

$$w_C^{e*} = R_2^H + \frac{\phi(\bar{\lambda}_A + \epsilon, H) - \phi(\bar{\lambda}_A, H) + 2\epsilon}{2-2q} + \phi(\lambda_C, H), \tag{28}$$

[23]Dominating here means yielding higher utility to the employer.

and entry-level wages are

$$w_1^{e*} = R_1^H - \frac{\phi(\bar{\lambda}_A + \epsilon, H) - \phi(\bar{\lambda}_A, H) - 2\epsilon}{2 - 2q} + \phi(\bar{\lambda}_A + \epsilon, H). \tag{29}$$

The graveyard equilibrium then dominates the pooling one if

$$c(1-q)(H-L) + dL \geq -2\phi(\bar{\lambda}_A, H) - \frac{\phi(\bar{\lambda}_A + \epsilon, H) - \phi(\bar{\lambda}_A, H)}{2-2q} - \epsilon + w_D^{e*}, \tag{30}$$

and the normal separating one if

$$2w_1^{s*} + w_C^{s*} < 2w_1^{e*} + w_C^{e*} + w_D^{e*}. \tag{31}$$

This becomes

$$dL - w_D^{e*} \geq \frac{3-4q}{2-2q}\big(\phi(\bar{\lambda}_A + \epsilon, H) - \phi(\lambda_B^{s*}, H)\big), \tag{32}$$

so long as the normal separating equilibrium has an interior solution, i.e. the L-type's period-one participation constraint (11) is not binding,[24] which is the most likely case given that the whole analysis is most relevant for workloads on a flatter part of the cost-of-working function. By assumption the left hand side of this condition is negative, and since $\phi(\bar{\lambda}_A + \epsilon, H) - \phi(\lambda_B^{s*}, H)$ is negative too, we know for sure that a graveyard equilibrium cannot dominate the normal sorting one if $q \geq \frac{3}{4}$.

In turn this shows that a graveyard equilibrium can dominate the normal separating one when q is not too close to one. This suggests that the use of elephant's graveyards posts, or similar retention institutions like early retirement, should be expected especially in organizations with noticeable talent uncertainty. In such settings their existence may serve an efficient sorting function, even if these institutions only appear to aliment staff that was retained despite its obvious incapability.

Also here a brief look at wage setting is worthwhile. Equations (29) and (28) reveal that, similar to before, the H-type has a wage profile exhibiting delayed compensation. Interestingly, entry-level wages may now be so low that even the L-type is underpaid. This does not pose a problem because also he now has a credibly tenured contract with higher compensation in period-two. The optimality of retaining, and sometimes even overpaying low-type agents is admittedly a somewhat surprising result. It suggests that bureaucracies' tendency to retain and even lavishly aliment seemingly incapable bureaucrats occupying obviously useless positions may be justified: They deserve it because they have been paid too little earlier in their career, and that way they at least

[24]The other condition is in footnote A.4 in the appendix.

do not end up on really important management positions where they might do damage.

A word of caution is appropriate here though. Leadership selection in this model started on the premise that a choice has to be made between two appropriately pre-selected candidates, which still differ in ability—even if just marginally—so that there is always a better and a "less better" choice. The language of 'high type' and 'low type' is in no way to suggest that the low type should not have been employed by the organization in the first place. The model obviously does not have anything to say about how good mechanisms for the pre-selection of candidates should be designed.

The result that sorting may obtain by paying the L-type agent a wage or pension can also be interpreted with respect to the case of capped workloads. Recall that the main finding in the previous subsection was that leadership selection cannot be perfect because sorting is only possible when a mixed promotion strategy is chosen. This subsection has shown that L-type agents can be kept away from the really important management post C by being "buried" on elephants' graveyard posts or being sent into early retirement. Perfect leadership selection can be restored.

4.3. Adding Productivity and Experience Effects

At the start of the paper, I made the seemingly innocent restriction that only post B can be tougher than A, which on the upside saved me (and the reader) to write (and read) every result twice—perhaps thereby even recurring to the rather confusing habit of presenting one (the other) result outside (inside) parentheses, as I will do for the remainder of this section. It is not hard to see that lifting this restriction only adds a mirror image of every asymmetric separating equilibrium, in which post A is tougher than B. But even if both equilibria are possible, as long as the exogenous symmetry assumptions are in place the model has nothing to say about which post will become the tough launchpad. I made these assumptions quite intentionally to rule out that any kind of comparative advantage or job-suitability effects dilute the result that asymmetric job designs evolve endogenously only to serve a leadership selection function. Setting the job importance parameters $a \neq b$ adds to the analysis the issues that are addressed by the job assignment literature, such as Sattinger (1975) or Rosen (1978). If $a > b$ ($a < b$) the equilibria remain but now the principal has clear preferences over which post to make tougher, namely to make post A tougher than B (B tougher than A). This is because then the asymmetry does not only enhance performance on post C through improved leadership selection but the employer also derives additional utility from the fact that agents self-sort already in period one. By making the more important post tougher the employer has an additional gain of $(b - a)(H - L)$ compared to levying the additional workload on the less important post.

A related reason for why the employer may have a strict preference for making a specific post tough builds on a human capital accumulation argument touched upon in the literature review. For the simplest possible way to show this, suppose that working on post B provides useful training for the task to be performed on post C, so that whoever holds post B in period one has his ability raised by a positive amount σ. σ only has an effect if the agent who gains it subsequently moves to post C. It is therefore obvious that in this setting making post B tougher is the only optimal asymmetric design. Employer utility is raised by $(b-a)(H-L)+c\sigma$ compared to if post A is turned into the launchpad post.

There is a good and a bad implication from the fact that both of these simple extensions weaken the conditions for asymmetric sorting equilibria to occur. The good implication is that such equilibria are in fact quite likely to occur in reality, which makes the results of this paper more robust and empirically more relevant. The less handy implication is, however, that observing asymmetric job designs and career path differentiation empirically need not indicate the presence of sorting but may just be due to other explanations. The conclusion takes up this issue.

5. Conclusion

This paper has presented a model of careers in hierarchies which creates a link between a number of phenomena we observe in reality: First, hierarchies usually exhibit stark asymmetries in the sense that some jobs are much more demanding than others, even if they are formally on the same hierarchical level and have the same pay grade attached to them. Second, having held a tough position often launches an employee into a very successful future career. Third, the early-phase career paths of successful employees often resemble each other, while being very different from the early career phases of less successful employees. Finally, it is a common practice in many large organizations and especially public bureaucracies that employees who are not fit for high-level management positions are retaining and alimented in early retirement arrangements or on so called elephants' graveyards posts. Very simply put, the analysis in this paper suggests that all of these phenomena may be part of efficient institutional arrangements which help to induce self-sorting and thereby improve leadership selection.

These findings yield a number of interesting empirical predictions and define a promising empirical research agenda. First of all, in order to test the plausibility of the model, it should be established more rigorously that indeed career path differentiation exists in that early-phase career patterns are a good predictor of later career success. The scarce anecdotes that this paper was inspired by are hardly convincing. Second, the presence of clearly differentiated career paths in an organization, and especially if the early-phase career patterns of an employee reliably predict his or her later career success, may be a good indicator for good leadership selection. Third and perhaps more surprisingly, the *absence* of any formal performance appraisal mechanism may also indicate that healthy self-selection is going on—provided that career paths are differentiated. This is because if self-selection proves to be a good means to scree leaders, performance assessments may simply be abandoned, never introduced, or at least not taken very seriously.

But there are several caveats with taking any such evidence at face-value for there not being any inefficiencies in a given organization. Reality is obviously more complex than the model presented here. My argument only highlights *one* of many mechanisms that may contribute to good leadership selection, namely self-selection.

First, information asymmetries are likely to be less stark than assumed here. Agents may not have complete information about their talent or type. They might simply overestimate their own ability, or their career commitment may change because they marry and want to spend time with their family. In addition, also employers usually know quite well who is suitable for which job, especially after having observed a worker for some time. All this suggests that in reality career path differentiation should never be as stark as predicted by this model, and the quality of promotion rules based purely on past careers would hardly be perfect. And yet, if there is extreme information asymmetry about one thing then it is how committed someone is and how willing he is to accept the workloads and long hours that a high-profile career demands. And such commitment can be signalled quite credibly by working on a particularly tough post as the model suggests. I am therefore convinced that self-selection is an important aspect of career choices within organizations and since it works in a very invisible way it is likely to make leadership selection better that one would suspect by only investigating the obvious selection devices.

Second, any problems of moral hazard and incentivization of bureaucrats are completely eclipsed in my model. The whole incentive problem is black-boxed into the cost-of-working function, which simply imposes a certain cost for anyone working on a particular assignment—agents have no ‘work or shirk’ choice. The main reason for this simplification is manageability of the model. But it states a problem because if real-world promotion policies were solely based on past assignment, making promotion from a launchpad post *de facto* automatic without any further checks on performance, this would hardly induce the highly-committed careerers to work as hard as they claim to be willing to work. There are countless other models which analyse these problems and the institutions best suited to deal with them, for example performance-related pay or appraisals.

The third complication concerns empirical testing of the proposed hypotheses. A number of the phenomena that I mentioned as possible indicators of an organization having effective self-selection might as well be due to other (efficient or inefficient) mechanisms being at work. As already discussed in the introduction and literature review, if asymmetries in jobs exist for exogenous reasons, then launchpad effects and early career path differentiation may also evolve due to human capital acquisition effects. People acquire skills over time and they learn different things on different posts. The brief extension that incorporated human capital acquisition effects into the model suggested that this does not substantially alter the predictions of the model other than making asymmetries more likely. Unfortunately this makes it difficult to meaningfully differentiate the two mechanisms with any obvious empirical strategy. However, here a lot hinges on *how exactly* human capital is accumulated. On the one hand, in sectors where specialists are required the best training may indeed be holding a very particular sequence of positions, making the predictions of the sorting and human capital acquisition models identical. On the other hand, where generalist managers are needed, career paths that provide a good training ground are likely to be much more heterogeneous.

After all, perhaps the best but admittedly indirect test to knowing whether an organization has effective self-selection institutions is the following: Consider an organization with differentiated career paths and without any institutionalized performance appraisal mechanism at time t. Now suppose that at time $t+1$ such a procedure is introduced. If the career patterns of successful employees who started their career well before t do *not* differ markedly from the career patterns of successful employees who started their career after t, then self-selection is likely to have worked very well already before the introduction of the appraisals. If human capital acquisition effects are predominant, this outcome is unlikely to evolve because the performance appraisals will reflect differences in how well people gain experience and acquire skills over time, which will be positively correlated with general ability. As a result, not everyone having been on a training camp post will make it to the top. The biggest downside: I know of no suitable data that covers a long enough time span to test exactly this.

A. Proofs

Note that throughout the proofs a tilde on top of symbols denoting actors' strategy choices, such as $\tilde{w}$, is used for choices that are conditional on other actors' actions. A star is used only for choices that lie on the equilibrium path.

A.1. Proof of Proposition 1

The agent equilibrium strategies of the equilibrium characterized in Proposition 1 are identical for both agents, namely

$$S^{*\theta}(\tilde{\pi})$$

where

$$\tilde{\pi} = \begin{cases} O & \text{if} \quad \lambda_A = \lambda_B \\ A & \text{if} \quad \lambda_A < \lambda_B. \end{cases}$$

The employer's equilibrium strategy is

$$S^{*E}(\lambda_A^* = \lambda_B^* = \bar{\lambda}_A;\ w_1^*;\ \tilde{w}_C;\ \tilde{\alpha})$$

where

$$\tilde{w}_C = w_C^* = R_2^H + \phi(\lambda_C, H),$$

and

$$\tilde{\alpha} = \alpha^* = \begin{cases} 0 & \text{if} \quad \Theta = [\Lambda^L = A, \Lambda^H = B],\ \forall\,\Pi \\ 1 & \text{if} \quad \Theta = [\Lambda^L = B, \Lambda^H = A],\ \forall\,\Pi. \end{cases}$$

Existence of the equilibrium is proven backwardly. Suppose $\lambda_A = \lambda_B$ and $\Pi = (O, O)$, then $\tilde{\alpha}$ as just specified is the only optimal employer promotion choice because applications are uninformative and the only information that forms the beliefs of the employer about which agent is assigned to which post is Θ which indicates the true assignments with probability $q > \frac{1}{2}$.

The optimal employee response to $\lambda_A = \lambda_B$ is $\pi = O$ if each agent is indifferent between assignment A and B given α^*, which is the case because the L and H-type agents, respectively, find that

$$
\begin{aligned}
&w_1 - \phi(\lambda_A, L) + (1-q)(w_C - \phi(\lambda_C, L)) + qR_2^L \\
&\qquad = w_1 - \phi(\lambda_B, L) + (1-q)(w_C - \phi(\lambda_C, L)) + qR_2^L, \\
&w_1 - \phi(\lambda_A, H) + q(w_C - \phi(\lambda_C, H)) + (1-q)R_2^H \\
&\qquad = w_1 - \phi(\lambda_B, H) + q(w_C - \phi(\lambda_C, H)) + (1-q)R_2^H.
\end{aligned}
$$

The only optimal workload choice for the employer is $\lambda_A^* = \lambda_B^* = \bar{\lambda}_A$. This is a) because any symmetric increase of both workloads above their minimum values $\bar{\lambda}_A = \bar{\lambda}_B$ inflates the wage bill without enhancing productivity. And b) an *asymmetric* workload choice may be justified if it induces sorting. For this to be the case it must be that, given $\lambda_A < \lambda_B$[25], one agent strictly prefers one post while the other agent strictly prefers the other. Note that in such a situation the beliefs of the employer are formed solely on the basis of sorting and thus imply a pure promotion strategy with $\alpha \in \{0, 1\}$, depending on which way around agents self-sort.
Here suppose first that agent H prefers post A to B and agent L prefers post B to A, in which case the optimal $\alpha = 1$. For such a situation to exist it must be that none of the agents has an incentive to apply to the same post as the other agent, which is the case if

$$
\begin{aligned}
&w_1 - \phi(\lambda_A, H) + w_C - \phi(\lambda_C, H) \geq \\
&\qquad w_1 - \frac{1}{2}\big[\phi(\lambda_A, H) + \phi(\lambda_B, H)\big] + q\big(w_C - \phi(\lambda_C, H)\big) + (1-q)R_2^H + \epsilon,
\end{aligned} \tag{33}
$$

and

$$
\begin{aligned}
&w_1 - \phi(\lambda_B, L) + R_2^L \geq \\
&\qquad w_1 - \frac{1}{2}\big[\phi(\lambda_A, L) + \phi(\lambda_B, L)\big] + (1-q)\big(w_C - \phi(\lambda_C, L)\big) + qR_2^L + \epsilon.
\end{aligned} \tag{34}
$$

(Note here that the right hand sides of these two conditions are the expected utilities that obtain if one of the agents deviates from separating application behaviour: First-round assignments are decided by coin throw, the employer's beliefs are formed by the performance appraisal report, and promotion is carried out accordingly.) Substituting

[25] I restricted asymmetry to post B being tougher than A at the outset of the paper just to reduce the number of cases to consider. The proof applies analogously to $\lambda_A > \lambda_B$.

(33) into (34) yields

$$(1-q)\big[R_2^L+\phi(\lambda_C,L)-R_2^H+\phi(\lambda_C,H)\big]-2\epsilon \geq \\ \frac{1}{2}\big[\phi(\lambda_B,H)-\phi(\lambda_A,H)+\phi(\lambda_B,L)-\phi(\lambda_A,L)\big],$$

which—not very surprisingly—is a contradiction by the simplifying restriction $\lambda_A < \lambda_B$ and the careering assumption (3). (This case only shows that "reverse" sorting is impossible.)

Suppose now that agent H prefers post B to A and agent L prefers post A to B, in which case $\alpha = 0$. For such a situation to exist it must be that none of the agents has an incentive to apply to the same post as the other agent, which is the case if

$$w_1-\phi(\lambda_B,H)+w_C-\phi(\lambda_C,H) \geq \\ w_1-\frac{1}{2}\big[\phi(\lambda_A,H)+\phi(\lambda_B,H)\big]+q\big(w_C-\phi(\lambda_C,H)\big)+(1-q)R_2^H+\epsilon, \tag{35}$$

and

$$w_1-\phi(\lambda_A,L)+R_2^L \geq \\ w_1-\frac{1}{2}\big[\phi(\lambda_A,L)+\phi(\lambda_B,L)\big]+(1-q)\big(w_C-\phi(\lambda_C,L)\big)+qR_2^L+\epsilon. \tag{36}$$

These conditions cannot be jointly satisfied. To see that, start by rearranging (35) so that

$$(1-q)\big[w_C-\phi(\lambda_C,H)-R_2^H\big] \geq \frac{1}{2}\big[\phi(\lambda_B,H)-\phi(\lambda_A,H)\big]+\epsilon,$$

which for $\lambda_B > \lambda_A$ can only hold if $w_C > R_2^H + \phi(\lambda_C, H)$. While (36) would be satisfied with such a wage (by the careering assumptions (2)), any second-period wage w_C greater than the reservation wage of the H-type agent is not feasible. Without binding commitment to future wages (as assumed in this scenario), once the employees have revealed their type through self-sorting and worked in period-one, the employer will always defect from her initial promise and pay only $w_C = R_2^H + \phi(\lambda_C, H)$. Condition (35) can thus not be satisfied without commitment to future wages. Proposition 1 including uniqueness of the equilibrium and the entry-level wage as given in the core of the text follow. Profitability of the organization is ensured by the staffing assumption (1). *qed*

A.2. Proof of Proposition 2

The employer's strategy in the equilibrium characterized in Proposition 2 is

$$S^{*E}(\lambda_A^* = \bar{\lambda}_A;\ \lambda_B^* > \lambda_A^*;\ w_1^*;\ \tilde{w}_C;\ \tilde{\alpha})$$

where λ_B^* is given by (14), w_1^* by (16),

$$\tilde{w}_C = \begin{cases} w_C^{s*} \text{ (as given by equ. 15)} \quad \text{if} \quad \Pi = \{A, B\}\ \forall\ \Theta \\ R_2^H + \phi(\lambda_C, H) \quad \text{otherwise,} \end{cases}$$

and

$$\tilde{\alpha} = \alpha^* = \begin{cases} 0 \quad \text{if} \quad \Pi = \{A, B\}\ \forall\ \Theta; \text{ or if } \Pi \neq \{A, B\} \text{ and } \Theta = [\Lambda^L = A, \Lambda^H = B], \\ 1 \quad \text{if} \quad \Pi \neq \{A, B\} \text{ and } \Theta = [\Lambda^L = B, \Lambda^H = A]. \end{cases}$$

The H-type agent's equilibrium strategy is

$$S^{H*}(\tilde{\pi}^H)$$

with

$$\tilde{\pi}^H = \begin{cases} B \quad \text{if } (9), (12) \text{ satisfied} \\ A \quad \text{otherwise}\,. \end{cases}$$

The L-type agent's equilibrium strategy is

$$S^{L*}(\tilde{\pi}^L)$$

with

$$\tilde{\pi}^L = \begin{cases} A \quad \text{if } (10), (11) \text{ satisfied} \\ B \quad \text{otherwise}\,. \end{cases}$$

Expressions (12) and (11) are simply the agent participation constraints. (9) is obtained from the condition stating when the H-type prefers a $B \rightarrow C$ career with the employer choosing $\alpha = 0$, namely

$$\begin{aligned} &w_1 - \phi(\lambda_B, H) + w_C - \phi(\lambda_C, H) \geq \\ &\quad w_1 - \frac{1}{2}\big[\phi(\lambda_A, H) + \phi(\lambda_B, H)\big] + q\big[w_C - \phi(\lambda_C, H)\big] + (1-q)R_2^H + \epsilon. \end{aligned} \tag{37}$$

Analogously, (10) is obtained from the expression stating when the L-type agent prefers an $A \to C$ career with the employer choosing $\alpha = 0$, namely

$$\begin{aligned} &w_1 - \phi(\lambda_A, L) + R_2^L \geq \\ &\quad w_1 - \frac{1}{2}\big[\phi(\lambda_A, L) + \phi(\lambda_B, L)\big] + (1-q)\big[w_C - \phi(\lambda_C, L)\big] + qR_2^L + \epsilon. \end{aligned} \tag{38}$$

Substituting yields to the following joint agent sorting condition:

$$\begin{aligned} \lambda_A, \lambda_B \ : \ &\phi(\lambda_B, L) - \phi(\lambda_A, L) - \big[\phi(\lambda_B, H) - \phi(\lambda_A, H)\big] = \\ &(2-2q)\big[R_2^H - R_2^L - \phi(\lambda_C, L) + \phi(\lambda_C, H)\big] + 4\epsilon, \end{aligned} \tag{39}$$

Because this condition is satisfied solely from a sufficiently large *difference* between λ_A and λ_B, raising λ_A above $\bar{\lambda}_A$ cannot be optimal. Expression (14) and the equilibrium wages derived in the core of the paper follow. *qed*

A.3. Capped Workloads and Regulated Wages

This Appendix sketches the existence of an equilibrium similar to that characterized by Proposition 2 when workloads are capped and/or wages are regulated. A change in assumptions is that commitment to a mixed promotion strategy is possible, even if it is *ex post* suboptimal.

As stated in the core of the paper, the scenario assumes that (37) and (38) cannot be jointly satisfied. As in the main text let $\alpha \in [0, 1]$ be the mixed promotion strategy of the employer. Then sorting can nevertheless be ensured if

$$\begin{aligned} &w_1 - \phi(\lambda_B, H) + (1-\alpha)\big[w_C - \phi(\lambda_C, H)\big] + \alpha R_2^H \geq \\ &\quad w_1 - \frac{1}{2}\big[\phi(\lambda_A, H) + \phi(\lambda_B, H)\big] + q\big[w_C - \phi(\lambda_C, H)\big] + (1-q)R_2^H + \epsilon, \end{aligned} \tag{40}$$

and

$$\begin{aligned} &w_1 - \phi(\lambda_A, L) + \alpha\big[w_C - \phi(\lambda_C, L)\big] + (1-\alpha)R_2^L \geq \\ &\quad w_1 - \frac{1}{2}\big[\phi(\lambda_A, L) + \phi(\lambda_B, L)\big] + (1-q)\big[w_C - \phi(\lambda_C, L)\big] + qR_2^L + \epsilon \end{aligned} \tag{41}$$

are satisfied. Rearranging yields

$$\alpha \leq (1-q) + \frac{\frac{1}{2}\big[\phi(\lambda_B, H) - \phi(\lambda_A, H)\big] + \epsilon}{R_2^H - \big(w_C - \phi(\lambda_C, H)\big)} \tag{42}$$

and

$$\alpha \geq (1-q) + \frac{\frac{1}{2}\big[\phi(\lambda_B, L) - \phi(\lambda_A, L)\big] + \epsilon}{R_2^L - \big(w_C - \phi(\lambda_C, L)\big)} = \alpha^*. \tag{43}$$

From substitution one can see that for some parameter values there exists a wage w_C, which jointly satisfies these conditions. The latter expression gives the employer's optimal promotion strategy, because she wants an α as small as possible in order to minimize the incident of promoting the L type agent from post A. Note that $\alpha^* > 1-q$, which indicates the inefficiency in promotion created by regulated wages or capped workloads.

A.4. Proof of Proposition 3

The employer's strategy in the equilibrium characterized in Proposition 3 is

$$S^{*E}(\lambda_A^* = \bar{\lambda}_A;\ \lambda_B^* = \bar{\lambda}_A + \epsilon;\ w_1^*;\ \tilde{w}_C;\ \tilde{w}_D;\ \tilde{\delta}).$$

where w_1^* is given by (29),

$$\tilde{w}_C = \begin{cases} w_C^{e*} \text{ (as given by equ. 28)} \quad \text{if} \quad \Pi = \{A, B\}\ \forall\ \Theta \\ R_2^H + \phi(\lambda_C, H) \quad \text{otherwise,} \end{cases}$$

$$\tilde{w}_D = \begin{cases} w_D^{e*} \text{ (as given by equ. 27)} \quad \text{if} \quad \Pi = \{A, B\}\ \forall\ \Theta \\ 0 \quad \text{otherwise,} \end{cases}$$

and

$$\tilde{\delta} = \begin{cases} 1 \quad \text{if} \quad \Pi = \{A, B\}\ \forall\ \Theta;\ \text{or if}\ \Pi \neq \{A, B\}\ \text{and}\ \ \Theta = [\Lambda^L = A, \Lambda^H = B], \\ 0 \quad \text{if} \quad \Pi \neq \{A, B\}\ \text{and}\ \ \Theta = [\Lambda^L = B, \Lambda^H = A]. \end{cases}$$

This strategy ensures a) that promotion from B to C occurs if agents self-sort, or if the performance appraisal indicates that H is on B—otherwise the agent on A is promoted, and b) that if no sorting was observed, the non-promoted agent is not rewarded on the D post—he will decline the promotion "offer" because w_D is set to zero.

The H-type agent's equilibrium strategy is

$$S^{H*}(\tilde{\pi}^H)$$

with

$$\tilde{\pi}^H = \begin{cases} B & \text{if } (19),(22) \text{ satisfied} \\ A & \text{otherwise}. \end{cases}$$

The L-type agent's equilibrium strategy is

$$S^{L*}(\tilde{\pi}^L)$$

with

$$\tilde{\pi}^L = \begin{cases} A & \text{if } (20),(21) \text{ satisfied} \\ B & \text{otherwise}. \end{cases}$$

Proposition 3 follows for all parameter values for which the necessary condition (25) and the sufficient conditions (30) and (32) are satisfied.[26] This is most likely to be the case if workloads are such that agents operate on a flat part of ϕ. *qed*

A.5. Graveyards or High Workloads?

This Appendix derives condition (25). Replacing all $\bar{\lambda}_A + \epsilon$ in (27) with λ_B we obtain

$$\frac{d\lambda_B}{dw_D} = \frac{2}{\phi'(\lambda_B, H) - \phi'(\lambda_B, L)}. \tag{45}$$

Taking total differentials of $W = 2w_1 + w_C$ with wages given by (16) and (15) we obtain

$$\frac{dW}{d\lambda_B} = \frac{(3-4q)\phi'(\lambda_B, H)}{2-2q}, \tag{46}$$

[26] For completeness sake, if the normal separating equilibrium has a corner solution with (11) biting, the graveyard equilibrium dominates if

$$\begin{aligned} dL - w_D^{e*} \;\geq\; & \frac{\phi(\bar{\lambda}_A + \epsilon, H) - \phi(\lambda_B^{s*}, H)}{2-2q} + \\ & 2\Big[R_1^H - R_1^L + \phi(\bar{\lambda}_A + \epsilon, H) - \phi(\bar{\lambda}_A, L) - \frac{\phi(\bar{\lambda}_A + \epsilon, H) - \phi(\bar{\lambda}_A, H) - 2\epsilon}{2-2q}\Big], \end{aligned} \tag{44}$$

which can be satisfied. It is unfortunately interpretable only in a less straight-forward manner.

and condition (25) follows from (45) and (46).

References

Ariga, Kenn, Yasushi Ohkusa and Giorgio Brunello. 1999. "Fast track: is it in the genes? The promotion policy of a large Japanese firm." *Journal of Economic Behavior and Organization* 38(4):385–402.

Becker, Gary. 1962. "Investment in Human Captial: A Theoretical Analysis." *Journal of Political Economy* 70(Supplement).

Belzil, Christian and Michael Bognanno. 2004. "The Promotion Dynamics of American Executives." IZA Discussion Paper No 1003.

Bernhard, Dan. 1995. "Strategic Promotion and Compensation." *Review of Economic Studies* 62(2):315–339.

Besley, Timothy and Maitreesh Ghatak. 1995. "Competition and Incentives with Motivated Agents." *American Economic Review* 95(3).

Chiappori, Pierre-André, Bernard Salanié and Julie Valentin. 1999. "Early Starters versus Late Beginners." *Journal of Political Economy* 107(4):731–760.

Demougin, Dominique and Aloysius Siow. 1994. "Careers in Ongoing Hierarchies." *The American Economic Review* 84(5):1261–1277.

Doeringer, Peter B. and Michael J. Piore. 1971. *Internal Labour Markets and Manpower Analysis.* Lexington, Massachusetts: Heath Lexington Books.

Gibbons, Robert and Michael Waldman. 1999. "A Theory of Wage and Promotion Dynamics inside Firms." *Quarterly Journal of Economics* 114(4):1321–1358.

Hood, Christopher and Martin Lodge. 2006. *The Politics of Public Service Bargains: Reward, Competency, Loyalty – and Blame.* Oxford: Oxford University Press.

Horn, Murray. 1995. *The Political Economy of Public Administration: Institutional Choice in the Public Sector.* Cambridge: Cambridge University Press.

Jovanovic, B. and Y. Nyarko. 1997. "Stepping Stone Mobility." *Carnegie-Rochester Series for Public Policy* 46:289–337.

Lazear, Edward. 1981. "Agency, Earning Profiles, Productivity, and Hours Restrictions." *American Economic Review* (71):606–620.

Lazear, Edward P. and Sherwin Rosen. 1981. "Rank-Order Tournaments as Optimum Labour Contracts." *Journal of Political Economy* 89(5):841–864.

Malcomson, James M. 1984. "Work incentives, hierarchies, and internal labour markets." *Journal of Political Economy* 92:486–507.

Meyer, Margaret. 1991. "Learning from coarse Information: biased contests and career profiles." *Review of Economic Studies* 58:15–41.

Meyer, Margaret. 1992. "Biased contests and moral hazard: implications for career profiles." *Annales d'Economie et de Statistique* 25/26:165–187.

Miyazaki, Hajime. 1977. "The rat race and internal labor markets." *Bell Journal of Economics* 8:394–417.

Prendergast, Canice. 1992. "Career Development and Specific Human Capital Collection." *Journal of the Japanese and International Economies* 6:207–227.

Prendergast, Canice. 1993. "The Role of Promotion in Inducing Specific Human Capital Acquisition." *Quarterly Journal of Economics* 108(2):523–534.

Ramseyer, J. Mark and Frances McCall Rosenbluth. 1993. *Japan's Political Marketplace.* Cambridge MA and London: Harvard University Press.

Rauch, James E. 2001. "Leadership Selection, Internal Promotion, and Bureaucratic Corruption in less Developed Polities." *Canadian Journal of Economics* 34(1):240–258.

Ricart i Costa, J. 1988. "Managerial Task Assignment and Promotion." *Econometrica* 56:449–466.

Rosen, S. 1978. "Substitution and Division of Labour." *Economica* 45:235–250.

Sattinger, M. 1975. "Comparative Advantage and the Distributions of Earningns and Abilities." *Econometrica* 43:455–468.

Spence, Michael. 1973. "Job Market Signalling." *Quarterly Journal of Economics* 87(3):355–374.

Waldman, Michael. 1984. "Job Assignments, Signalling, and Efficiency." *Rand Journal of Economics* 15(1):255–267.

Williamson, Oliver, Michael Wachter and Jeffrey Harris. 1975. "Understanding the Employment Relation: The Analysis of Idiosyncratic Exchange." *Bell Journal of Economics* 6(Spring):250–278.