Providing Benefits to Uninformed Workers:
A Dynamic Search Model with Hidden Attributes*

Tomasz Sulka†

23 March, 2026

Abstract

This paper develops a dynamic search model in which certain "hidden attributes" are revealed only after acceptance of an offer and may trigger continued search in the following period. The model is applied to study how workers' imperfect information about pecuniary workplace benefits (such as employer-sponsored pension and health insurance plans) during job search, and the subsequent realization of these benefits on the job, affect the multidimensional compensation packages offered in equilibrium by profit-maximizing firms. I find that unobservability of benefits prior to acceptance distorts firms' incentives toward providing inefficiently low benefits, despite the fact that lower benefits induce higher worker turnover. Furthermore, when workers differ in strategic sophistication, and therefore hold different beliefs about unobservable benefits, there exist equilibria with spurious differentiation in compensation packages. In these equilibria, the wage differential is bounded from above by the benefit differential. The model demonstrates how imperfect information about workplace benefits can explain several empirical puzzles, including inefficiently low benefit provision and large between-firm dispersion in benefits.

JEL: D83, D91, J31, J32, J33

Keywords: Exploitative Contracting; Hidden Attributes; Job Search; Workplace Benefits; Compensating Differentials

*I would like to thank Ş. Nuray Akin, Daniel Garcia, Paul Heidhues, Maarten Janssen, Heiko Karle, Bernhard Kasberger, Philipp Kircher, Mats Köster, Eirik Gaard Kristiansen, Cristina Lafuente, Simon Martin, Eeva Mauring, David Ronayne, Heiner Schumacher, Marco Schwarz, Heidi Thysen, Georg Weizsäcker, Irenaeus Wolff and participants in a number of seminars and conference presentations for helpful comments and discussions. Financial support from the Deutsche Forschungsgemeinschaft (DFG) through CRC TRR 190 (project number 280092119) is gratefully acknowledged.

†Humboldt-Universität zu Berlin, Germany. Email: tomasz.sulka@hu-berlin.de.

1 Introduction

Workplace benefits account for a large and increasing share of total compensation. For example, the average employee in the US receives 31% of their total compensation in the form of various pecuniary benefits, most importantly health insurance, paid leave, and retirement benefits.[1] Moreover, because such benefits are more unequally distributed than wages, wage-based measures of income inequality underestimate the extent of total compensation inequality (Kristal et al., 2020; Ouimet and Tate, 2023; Pierce, 2001).

At the same time, many employees have a poor understanding of the benefits provided by their employer, for example in the domain of workplace pensions (Agnew et al., 2012; Gustman and Steinmeier, 2005; Mitchell, 1988).[2] The evidence also indicates, however, that understanding of one's own benefits improves with tenure, which could be driven by social interactions among colleagues (Duflo and Saez, 2002, 2003).

In this paper, I depart from the standard assumption that workers are fully informed about firms' compensation packages and instead study how workers' imperfect information about pecuniary benefits during job search, and their subsequent realization, affect multidimensional compensation packages offered in equilibrium. Methodologically, I build on and extend the literature on exploitative contracting (Heidhues and Kőszegi, 2018) by incorporating dynamic effects that arise when an agent observes the hidden attribute of a chosen alternative and can subsequently search again.

More precisely, I consider a setting in which each compensation package consists of a *simple attribute* ("wage") and a *complex attribute* ("benefits"). Benefits can take one of two values: high or low. Firms' provision costs and workers' preferences are homogeneous and such that offering high benefits to every worker is efficient. I focus on this clear-cut case to isolate the effects of imperfect information about workplace benefits.

A worker searching for a job samples wage offers, but observes the value of associated benefits only after accepting a particular offer. Thus, I effectively model the benefits component of a compensation package as an experience good.[3] In terms of their beliefs about benefits associated with a given wage offer, I distinguish between "sophisticated" workers, who make correct inferences based on the joint distribution of wages and benefits, and "naïve"

[1] Compiled from the data of US Bureau of Labor Statistics, accessed 10 September 2024.

[2] More broadly, there is extensive evidence that individuals struggle with evaluating complex financial products, see the reviews by Beshears et al. (2018) and Lusardi and Mitchell (2014).

[3] Throughout the paper, I abstract from the possibility that a firm may educate workers and credibly advertise its benefits. Failures to educate are apparent in the prevalence of dominated plan choices, even in settings where simplified information is provided (e.g., Bhargava et al., 2017; Choi et al., 2010, 2011). Furthermore, pension and health insurance plans tend to differ across multiple dimensions and thus understanding one plan in detail is unlikely to allow a worker to compare it accurately with other plans.

workers, who correctly perceive the marginal distribution of benefits but mistakenly treat benefits as independent of the wage offer.[4] After realizing the value of benefits on the job, a worker decides whether to remain in current employment or to leave and search again.

I characterize equilibria of a two-period model where the firms commit to a single compensation package for both periods and the workers adopt perceived-optimal search strategies. Because both worker types are forward-looking, they never accept an offer in period 1 that would induce them to search again in period 2 even if the realized benefits were high (Corollary 1). Thus, in any equilibrium, providing low benefits leads to weakly higher turnover.

First, I show that unobservability of benefits prior to acceptance distorts firms' incentives toward providing inefficiently low benefits. For a given wage offer that induces acceptance, deviating to high benefits raises the firm's costs without improving hiring or retention, whereas deviating to low benefits reduces costs while maintaining hiring and, if search costs in period 2 are sufficiently high, retention. Consequently, the efficient outcome where all firms provide high benefits is more difficult to sustain in equilibrium than an outcome where all firms provide inefficiently low benefits (Proposition 1). Since the surplus and therefore profits are strictly greater in the high-benefits equilibrium, the unobservability of benefits by workers may be detrimental to the firms.

Second, the coexistence of sophisticated and naïve workers generates additional equilibria with spurious differentiation in compensation packages (Proposition 2). Because these worker types interpret the wage-benefit trade-off differently, firms can profitably offer distinct compensation packages, even in the absence of any heterogeneity in preferences or costs. Although these equilibria are qualitatively similar to the prediction of the classical compensating differentials theory (Rosen, 1974, 1986), the underlying logic and their efficiency properties are starkly different. Importantly, the model predicts that in such equilibria the wage differential is bounded from above by the benefit differential (Corollary 2). In particular, equilibrium wages may undercompensate workers for the actual differences in benefits across firms.

Beyond its theoretical contribution, the model helps explain several empirical puzzles related to workplace benefits, specifically: (i) inefficiently low benefit provision (Cole and Taska, 2023), (ii) the link between benefits and turnover (Ouimet and Tate, 2023), (iii) large between-firm dispersion in benefits (Ouimet and Tate, 2023), and (iv) the challenges in interpreting compensating differential estimates (Lavetti, 2023). These patterns are difficult

[4] Invoking correlation neglect (Enke and Zimmermann, 2019) provides a parsimonious way to capture workers who hold misspecified beliefs about the joint distribution of compensation packages while remaining forward-looking and anticipating future learning. This approach is closely related to an established game-theoretic literature modeling strategic mistakes through misspecified beliefs (e.g., Eyster and Rabin, 2005; Jehiel, 2005).

to rationalize with standard models assuming perfect information. The model further suggests that policies aimed at increasing benefit provision should target firms' incentives (e.g., through tax advantages) rather than rely on workers' ability to evaluate benefits and sort accordingly.

This paper contributes to several strands of the literature. First, by modeling the benefits component of a compensation package as unobservable during job search, I build on and extend the theoretical literature in behavioral industrial organization studying the design of products with hidden attributes, see Heidhues and Kőszegi (2018) for a review. In related recent papers, Heidhues et al. (2021) and Gamp and Krähmer (2022, 2023) introduced limited attention and misperceptions, respectively, into the model of consumer search.

To the best of my knowledge, no previous work accounts for the possibility that an agent eventually observes the hidden attribute chosen by the provider and might search again for a better alternative. Thus, the methodological contribution of this paper is to analyze the equilibrium effects of observing the hidden attribute and possible re-contracting.[5]

Second, I contribute to the literature in behavioral labor economics (Dohmen, 2014) by modeling a multidimensional compensation package consisting of an observable wage and unobservable benefits. Existing work analyzes the effects of workers' present bias (Bubb and Warren, 2020; DellaVigna and Paserman, 2005; Englmaier et al., 2023a,b; Fahn and Seibel, 2022), cost of gathering information (Wu, 2024), misperceptions of either occupational risk (Rea, 1981; Viscusi, 1980), job finding probability (Caliendo et al., 2015; Spinnewijn, 2015), or value of the outside option (Jäger et al., 2024), and complexity of incentives (Abeler et al., 2023; Schumacher and Thysen, 2022) on job search, effort provision, and compensation.

Third, the theoretical results of this paper complement the recent empirical literature on workplace benefits (e.g., Cole and Taska, 2023; Ouimet and Tate, 2023) and compensating differentials (Lavetti, 2023) by providing a unified explanation for several puzzling patterns, see section 4 for details.

The remainder of this paper is structured as follows. In section 2, I characterize boundedly rational workers' search behavior for a given distribution of offers. In section 3, I endoge-

[5] Schumacher (2024) analyzes a model in which a boundedly rational consumer correctly anticipates the value of the hidden attribute on the equilibrium path, but potentially misperceives it for an alternative she never selects. In terms of other dynamic aspects of competition in markets with hidden attributes, Johnen (2019) and Murooka and Schwarz (2019) analyze the design of automatic-renewal contracts when consumers differ in their propensity to passively accept the automatic renewal after realizing their (exogenous) utility from continued consumption. Johnen (2020) studies competition between firms who learn private information about their customers' naiveté. The consumers remain unaware that they are being exploited, however. Heidhues et al. (2023) analyze a model where firms offer an initial price and a switching price. Consumers observe all current and future prices perfectly, but may procrastinate on switching.

nize the distribution of wages and benefits offered in equilibrium. In section 4, I discuss the empirical relevance of the model's predictions. Section 5 concludes.

2 Search with simple and complex attributes

2.1 Setup

Consider a two-period search model in which agents eventually observe the complex attribute of a chosen alternative. In period 1, an agent enters the labor market and samples job offers, performing random, sequential search with perfect recall, as in the seminal model of McCall (1970). The first offer is sampled for free, while each additional draw imposes a fixed cost of $c_1 > 0$.[6] Each offer consists of a simple and a complex attribute. When sampling job offers, the agent observes only their simple attributes. The complex attribute associated with a particular offer is observed only after the offer is accepted and the agent derives utility from being employed. The agent may also take up her outside option, the value of which is normalized to 0.

In period 2, an employed agent decides whether to remain in her current employment or to leave and search again. If she stays, she derives the same (known) utility from being employed as in period 1. Otherwise, she searches at a fixed cost of $c_2 \geq c_1$ until she accepts a new offer.[7] The complex attributes of sampled offers remain unobservable prior to acceptance.

The form of unobservability of the complex attribute assumed here, essentially modeling it as an experience good (Nelson, 1970), can be micro-founded with costly information acquisition or verification. Moreover, such costs should be firm-specific, so that discovering the true value of a complex attribute at one firm does not make the agent an expert in evaluating other offers.[8] In the context of evaluating workplace benefits, there are important idiosyncratic features that indeed may require independent costly verification across firms, such as assessing the quality of a medical provider network in the case of health insurance (Gruber and McKnight, 2016) or a menu of investment options in the case of pension plans (Ayres and Curtis, 2015). Additionally, information acquisition costs may reflect complex or opaque pricing schemes that differ across providers due to "exploitative innovation" (Heidhues et al., 2016). When such information is sufficiently costly, an agent may optimally rely on delayed

[6] Assuming that the first offer is sampled for free is standard in the literature, as it permits equilibria with trade even though the Diamond paradox obtains. See Stiglitz (1979) for a discussion.

[7] It is standard in the literature to assume that search costs increase on the job, either due to classical economic reasons such as a higher opportunity cost of time (e.g., Burdett, 1978) or psychological factors such as time-inconsistent preferences or loss aversion (Heidhues et al., 2023; Karle et al., 2023).

[8] In that vein, Heidhues et al. (2021) analyze when a boundedly rational individual prefers to browse multiple offers superficially rather than study a single offer in detail.

but free informative signals about the complex attribute, for example obtained through peer interactions, information letters, or personal experience. For concreteness, I assume that such a signal materializes with probability 1 and is perfectly informative of the true value of the complex attribute.

Denoting the simple attribute by w ("wage") and the complex attribute by b ("benefits"), utility of agent i employed at firm k for a single period is:

$$u_{ik} = w_k + b_k.$$

Moreover, suppose that the benefits chosen by the employer can be either "high" or "low", i.e. $b_k \in \{\underline{b}, \bar{b}\}$ for some $\bar{b} > \underline{b} \geq 0$. Workplace benefits are assumed to be non-negative purely for exposition.[9]

When the agent stays in the labor market, her objective is to maximize the expected utility from being employed in two periods, net of search costs. For notational simplicity, I omit time discounting.

2.2 Beliefs and search behavior

In this section I analyze the search behavior of an agent facing an exogenously given distribution of multidimensional job offers. While I restrict attention to boundedly rational agents for whom the complex attribute of an offer remains unobservable prior to acceptance, there are several degrees of freedom in specifying how the agents form their beliefs about b_k. In this paper, I distinguish between "(strategically) sophisticated" and "(strategically) naïve" workers, defined as follows. A *sophisticated worker* holds beliefs about benefits associated with a particular wage offer that are consistent with the joint distribution of wages and benefits. A *naïve worker*, in contrast, takes into account the correct marginal distribution of benefits, but mistakenly perceives those as independent of the wage offer.[10]

Importantly, both sophisticated and naïve workers internalize the fact that they will observe the value of the complex attribute after acceptance. Thus, I depart from the approach prevalent in the behavioral industrial organization literature where it is standard to assume

[9] The literature accounts for the possibility that a worker might discover the idiosyncratic match quality, say ϵ_{ik}, on the job (Rogerson et al., 2005). As long as ϵ_{ik} is i.i.d. with mean zero and the worker's utility is additive, including this feature would not fundamentally affect the current model's predictions. Nevertheless, its mechanics are radically different when the unobservable component is a choice variable of the firm.

[10] For evidence on correlation neglect and its relation to complexity see Enke and Zimmermann (2019). Formally, correlation neglect can be modeled as manifestation of cursed beliefs (Eyster and Rabin, 2005) or analogy-based reasoning (Jehiel, 2005). See also Eyster (2019) for a taxonomy of errors that people make in strategic environments, summary of the lab and field evidence, and an overview of game-theoretic models incorporating those errors.

that a myopic agent does not anticipate the hidden component of an offer, see Heidhues and Kőszegi (2018). In section 3.3, I discuss the role of specific assumptions in generating the model's predictions.

Solving the model by backward induction, I characterize the search strategies adopted by both types for a given distribution of offers. For expositional simplicity, I assume throughout this section that wage offers are distributed continuously, which results in a clean illustration of the key intuition but is not essential for the results.

Period 2. Worker i employed at firm k understands the utility from continued employment to be $u_{ik} = w_k + b_k$. If she leaves to search again, her value function of sampling a wage offer w is:

$$v_2^j(w) = \max\,\{w + \mathbb{E}^j[b \,|\, w]\ ;\ v_2^j - c_2\},$$

where $j \in \{S, N\}$ denotes the worker's type (sophisticated or naïve), $v_2^j = \int_w v_2^j(w)\phi(w)dw$, and $\phi(w)$ is a density function characterizing the distribution of wage offers. The first expression inside the curly brackets represents the expected utility from accepting the offer. The second expression represents the expected utility from rejecting the offer and sampling again.

Notice that $\mathbb{E}^S[b \,|\, w] = \mathbb{E}[b \,|\, w]$, while $\mathbb{E}^N[b \,|\, w] = \mathbb{E}[b]$. Then, since $d\,\mathbb{E}^N[b \,|\, w]\,/\,d\,w = 0$, $v_2^N(w)$ equals $v_2^N - c_2$ for low values of w, and $w + \mathbb{E}[b]$ for sufficiently high values of w. Consequently, the perceived-optimal search strategy of a naïve worker always takes the form of a cutoff rule which prescribes to continue searching until the first offer with $w \geq R_2^N$ is sampled (McCall, 1970). In turn, a sophisticated worker follows a cutoff rule with reservation wage R_2^S provided that $d\,\mathbb{E}[b \,|\, w]\,/\,d\,w > -1$ for all w, so that $w + \mathbb{E}[b \,|\, w]$ is strictly increasing. For expositional simplicity, in this section I assume that this is indeed the case.[11]

Then, for all $w < R_2^j$, $v_2^j(w) = v_2^j - c_2$, while for all $w \geq R_2^j$, $v_2^j(w) = w + \mathbb{E}^j[b \,|\, w]$. Combining this with indifference at the cutoff, $R_2^j + \mathbb{E}^j[b \,|\, R_2^j] = v_2^j - c_2$, yields:

$$\begin{aligned}\int_w v_2^j(w)\phi(w)dw &= (R_2^j + \mathbb{E}^j[b \,|\, R_2^j]) \cdot \mathbb{P}[w < R_2^j] \ + \ \int_{w \geq R_2^j} (w + \mathbb{E}^j[b \,|\, w])\phi(w)dw \\ &= R_2^j \ + \ \mathbb{E}^j[b \,|\, R_2^j] \ + \ c_2,\end{aligned}$$

which solves for:

[11] If instead $d\,\mathbb{E}[b \,|\, w]\,/\,d\,w \leq -1$, then a sophisticated worker may deliberately search for low wages which in expectation are combined with high benefits and thus provide higher total utility. This case is indeed relevant for the characterization of certain equilibria in the subsequent section. However, abstracting from it here does not affect the qualitative insights of Lemmas 1 and 2 while streamlining the exposition.

$$c_2 = \int_{w \geq R_2^j} \left(w + \mathbb{E}^j[b \mid w] - R_2^j - \mathbb{E}^j[b \mid R_2^j]\right)\phi(w)dw. \tag{1}$$

Intuitively, the optimal reservation wage R_2^j equalizes the marginal cost of search with the marginal benefit of search. Since the right-hand side of (1) is strictly decreasing in R_2^j, the optimal cutoff is uniquely determined whenever the solution is interior.

How does the agent's type affect her search strategy? Intuitively, whether a sophisticated or a naïve worker adopts a higher reservation wage in period 2 should depend on whether the correlation between wages and expected benefits is positive or negative. For instance, if wages were negatively correlated with benefits, a naïve worker would overestimate the benefit of sampling a high wage offer, relative to her sophisticated counterpart, and would adopt a higher cutoff as a result.

This intuition is confirmed in the sense that the sign of $d\,\mathbb{E}[b \mid w] \,/\, d\,w$ pins down the ordering between the reservation wages adopted by sophisticated and naïve workers:

Lemma 1: *If $d\,\mathbb{E}[b \mid w] \,/\, d\,w \lesseqgtr 0$ for all w, then $R_2^N \gtreqless R_2^S$.*

A short proof is provided in Appendix A.

Period 1. The optimal search strategy in period 1 should depend on the anticipated behavior once in employment. While both types internalize the fact that they will observe the value of benefits on the job, the difference in beliefs about the conditional distribution of benefits may again result in divergent search rules.

Accounting for this, the value function of sampling an offer with wage w in period 1 is:

$$v_1^j(w) = \max\,\{u_1^j(w), v_1^j - c_1\},$$

where:

$$\begin{aligned} u_1^j(w) = \left(w + \mathbb{E}^j[b \mid w]\right) \;&+\; \mathbb{P}^j[w + b \geq v_2^j - c_2 \mid w] \cdot \left(w + \mathbb{E}^j[b \mid b \geq v_2^j - c_2 - w, w]\right) \\ &+\; \mathbb{P}^j[w + b < v_2^j - c_2 \mid w] \cdot \left(v_2^j - c_2\right) \end{aligned}$$

and $v_1^j = \int_w v_1^j(w)\phi(w)dw$. Conditional on accepting the offer, an agent expects to obtain the utility of $w + \mathbb{E}^j[b \mid w]$ in period 1. Furthermore, if the realized utility from being employed turns out to exceed the type-dependent cutoff $R_2^j + \mathbb{E}^j[b \mid R_2^j] = v_2^j - c_2$, she expects to remain in the same employment in period 2. Otherwise, she expects to search again, which yields expected utility of $v_2^j - c_2$.

The optimal search strategy in period 1 is also characterized by a cutoff rule as long as $d\,u_1^j(w)/d\,w > 0$. Differentiating yields:

$$\begin{aligned} d\,u_1^j(w)/d\,w \;=\;& \big(1 + d\,\mathbb{E}^j[b \mid w] \;/\; d\,w\big) \\ &+\; \mathbb{P}^j[w + b \geq v_2^j - c_2 \mid w] \cdot \big(1 + d\,\mathbb{E}^j[b \mid b \geq v_2^j - c_2 - w, w] \;/\; d\,w\big) \\ &+\; d\,\mathbb{P}^j[w + b \geq v_2^j - c_2 \mid w] \;/\; d\,w \cdot \big(w + \mathbb{E}^j[b \mid b \geq v_2^j - c_2 - w, w] - (v_2^j - c_2)\big). \end{aligned}$$

Clearly $d\,u_1^N(w)/d\,w > 0$, while for a sophisticated worker a sufficient condition for $u_1^S(w)$ to be strictly increasing in w is again $d\,\mathbb{E}[b \mid w] \;/\; d\,w > -1$. Then, proceeding as above yields the condition for the optimal cutoff in period 1:

$$c_1 = \int_{w \geq R_1^j} \big(u_1^j(w) - u_1^j(R_1^j)\big)\phi(w)dw. \tag{2}$$

Observe, however, that even imposing the sign of $d\,\mathbb{E}[b \mid w] \;/\; d\,w$ does not lead to a straightforward comparison between $d\,u_1^N(w)/d\,w$ and $d\,u_1^S(w)/d\,w$, and therefore between R_1^S and R_1^N. For example, if $d\,\mathbb{E}[b \mid w] \;/\; d\,w < 0$ for all w, then there are two opposing forces that affect how sophisticated workers search, relative to their naïve counterparts. On the one hand, sophisticated workers realize that higher wages are associated with lower expected benefits, which makes them less inclined to search intensely for high-wage offers. On the other hand, $d\,\mathbb{E}[b \mid w] \;/\; d\,w < 0$ implies $R_2^S < R_2^N$, which means that the initial choice is more likely to be "sticky" in period 2, thus increasing the sophisticated worker's incentive to search intensely for a good offer in period 1. Which of these two considerations dominates ultimately depends on a specific distribution of offers. This will therefore be endogenized in the next section. For now, note the following:

Lemma 2: *Even if $\mathbb{E}[b \mid w]$ is strictly monotonic in w, the sign of $d\,\mathbb{E}[b \mid w] \;/\; d\,w$ does not generally determine the ordering between R_1^N and R_1^S.*

A specific parametric example showing this indeterminacy is provided in Appendix B. Lemma 2 highlights that dynamic incentives have a major impact on how correlation neglect affects search behavior, relative to the static problem that applies in period 2. It therefore demonstrates the importance and added insight of accounting for dynamics induced by the realization of a hidden attribute.

Before proceeding to equilibrium analysis, it is useful to point out certain search strategies that can never be optimal. The following observation states that a forward-looking worker never accepts a wage offer in period 1 that, according to her beliefs, would induce her to search again in period 2 with certainty. This is a consequence of the assumption of weakly increasing search costs and the anticipation thereof.

Corollary 1: *A worker of type $j \in \{S, N\}$ rejects any wage offer $\hat{w}$ in period 1 such that $\hat{w} + \bar{b} < v_2^j - c_2$.*

A short proof is provided in Appendix C. An important implication of Corollary 1 is that a worker of either type would only ever leave her current employment in period 2 when she discovers low benefits. Therefore, the perceived-optimal search behavior in period 1 implies that, conditional on accepting an offer, low benefits lead to a (weakly) higher turnover.

3 Equilibrium analysis

In this section I characterize equilibria of the following game. There is a unit mass of identical firms that simultaneously choose the wage and benefits to offer to a unit mass of workers, committing to a single compensation package for both periods 1 and 2. Fraction $\lambda \in (0, 1)$ of workers are naïve, while $(1 - \lambda)$ are sophisticated. The worker's problem is as outlined in the preceding section.

Offering a compensation package (w, b) costs the firm $w + (1 - \tau)b$ per period when accepted by a worker. Here, $\tau \in (0, 1)$ captures advantages of providing workplace benefits, such as tax advantages.[12] Since the workers' utility function aggregates received wages and benefits, it is efficient to offer high benefits to all workers.

Each firm produces $y > 0$ units of a numeraire good per hired worker using a constant-returns-to-scale technology.[13] Thus, the profit from hiring a worker for a single period is:

$$\pi = y - w - (1 - \tau)b.$$

Firms maximize the expected sum of profits over the two periods.

The equilibrium concept I apply is weak perfect Bayesian equilibrium (PBE) adapted to allow naïve workers to display correlation neglect. Formally:

Definition 1: *Workers' strategies and beliefs, together with firms' strategies, constitute* PBE with correlation neglect *if:*

1. *(Consistency) The sophisticated workers' beliefs about the joint distribution of wages and benefits are derived from the firms' strategies via Bayes' rule.*

[12] For example, all OECD countries offer preferential tax treatment for private pension contributions, including employer contributions to qualifying workplace plans (OECD, 2020).

[13] The assumption of constant returns to scale is commonly made in the job search literature for tractability (Rogerson et al., 2005), but it has also received considerable empirical support (Petrongolo and Pissarides, 2001).

2. *(Naïve consistency) The naïve workers' beliefs about the marginal distributions of wages and benefits are derived from the firms' strategies via Bayes' rule, but they display correlation neglect in that they perceive those two components as independently distributed.*

3. *(Sequential rationality) Given their beliefs, each worker type adopts a perceived-optimal search strategy.*

4. *(Firms' best response) Given workers' strategies and other firms' offers, each firm chooses its compensation package to maximize expected profits.*

Workers' beliefs off the equilibrium path, specifically their beliefs about the benefits associated with a wage offer that deviates from equilibrium, determine the profitability of potential deviations. To maintain logical consistency with the belief formation of sophisticated and naïve workers on the equilibrium path, I impose the following additional assumptions. Sophisticated workers have *pessimistic beliefs*: upon observing a deviation, they suppose that the offer is associated with low benefits. Naïve workers, in turn, have *passive beliefs*: they expect the same marginal distribution of benefits to apply independently of a specific wage offer. This ensures that, off the equilibrium path, sophisticated workers do not become less strategically sophisticated and naïve workers do not become more sophisticated.[14] Furthermore, since the probability of re-sampling a specific firm's offer is zero, neither type updates beliefs about the overall distribution of offers upon observing a deviation by a single firm.[15] Thus, I supplement Definition 1 with the following assumption regarding off-path beliefs:

Assumption 1: *Let W^* denote the set of wages offered in equilibrium. Upon observing a deviation to $w_k \notin W^*$, sophisticated workers hold pessimistic beliefs about b_k (i.e., $\mathbb{E}^S[b \mid w_k] = \underline{b}$) and naïve workers hold passive beliefs about b_k (i.e., $\mathbb{E}^N[b \mid w_k] = \mathbb{E}[b]$). Neither type updates their beliefs about the distribution of offers.*

3.1 Equilibria with a degenerate distribution of benefits

Consider first equilibria in which all firms offer the same benefit, say b^*. Then, correlation neglect plays no role and workers of both types hold correct beliefs about the offered benefits. Furthermore, there is nothing new to discover on the job and therefore every worker anticipates that she will stay in the initially accepted employment for two periods.

[14] McAfee and Schwartz (1994) consider also *wary beliefs*, under which an agent infers what hidden action would be optimal from the deviator's perspective. This requires, however, imposing strong additional assumptions on the agent's higher-order beliefs, specifically beliefs about other agents' reactions to the deviation as well as the deviator's anticipation of those reactions.

[15] This satisfies a property called "no-signaling-what-you-don't-know" by Fudenberg and Tirole (1991) and is standard in the sequential search literature following Wolinsky (1986).

Conditional on b^*, all firms offer the same wage that is equal to the worker's reservation wage. There is no benefit to offering a higher wage, and offering a lower wage would result in the worker rejecting. Under Assumption 1, the following result obtains:

Lemma 3: *In any equilibrium with a degenerate distribution of benefits, the wage distribution is also degenerate.*

(a) If $b^* = \underline{b}$, then $w^* = -\underline{b}$.

(b) If $b^* = \bar{b}$, then $w^* = -\bar{b} + \Delta$ for some $\Delta \in [0, \bar{\Delta}]$.

The proof is provided in Appendix D. In equilibrium with uniformly low benefits, the wage offer has to make the worker indifferent between accepting employment and exiting the labor market, which implies $w^* = -\underline{b}$. The underlying logic is reminiscent of the well-known Diamond paradox (Diamond, 1971).

In equilibrium with uniformly high benefits, however, a wage exceeding $-\bar{b}$ might arise. That is because due to her pessimistic beliefs, a sophisticated worker interprets an infinitesimal reduction in the wage offer as a signal of low benefits, which can break the logic underlying the Diamond paradox. Thus, while the wage rate is not uniquely determined in an equilibrium with high benefits, the monopsony wage $w^* = -\bar{b}$ remains a valid equilibrium candidate. As an equilibrium selection criterion, for the remainder of the paper I restrict attention to (firm-optimal) equilibria with lowest wages that ensure labor market participation of both worker types.

Observe that an equilibrium with uniformly low benefits always exists. Deviating to $\bar{b}$ would necessarily make a firm strictly worse off because with benefits being unobservable prior to acceptance the deviator still has to offer a wage of at least $w^* = -\underline{b}$ in order to attract any workers, sophisticated or naïve. Thus, deviating to high benefits must raise the firm's labor cost without affecting the worker's decision to accept or to stay in employment. In this equilibrium, firms make strictly positive profits:

$$2\big(y - w^* - (1-\tau)\underline{b}\big) \;=\; 2\big(y + \tau\underline{b}\big) \;>\; 0.$$

Next, under what conditions can the efficient outcome with high benefits being offered to all workers be supported in equilibrium? Suppose that $b^* = \bar{b}$ and $w^* = -\bar{b}$. A firm that deviates to $b = \underline{b}$ cannot offer a wage lower than w^*, as this would result in both worker types rejecting and choosing the outside option instead. Consider first a deviation whereby the firm does not adjust its wage offer and simply combines w^* with low benefits $\underline{b}$. Despite discovering that the firm provides low benefits, a hired worker remains in employment in period 2 if $c_2 \geq (\bar{b} - \underline{b})$, thus making the deviation profitable. Otherwise, the worker leaves, but the deviator might still be better off if the profit from hiring a worker for a single period exceeds the profit from retaining the worker with high benefits:

$$y - w^* - (1-\tau)\underline{b} \;>\; 2(y - w^* - (1-\tau)\bar{b}) \iff$$
$$y + (\bar{b} - \underline{b}) + \tau\underline{b} \;>\; 2\big(y + \tau\bar{b}\big) \iff (1-\tau)(\bar{b} - \underline{b}) \;>\; \big(y + \tau\bar{b}\big),$$

which has the intuitive interpretation that the gain from deceiving the worker should exceed the surplus from the match lost in period 2.[16]

Second, if $c_2 < (\bar{b} - \underline{b})$, the deviating firm might combine $\underline{b}$ with some $\tilde{w}$ that prevents the worker from leaving in period 2. The lowest such wage that would be accepted by both types in period 1 is $w^*(\underline{b}) = -\underline{b}$, which implies that this kind of deviation can never be profitable as $2(y + \tau\underline{b}) < 2(y + \tau\bar{b})$. Alternatively, the deviator could offer $\tilde{w}$ that satisfies:

$$\tilde{w} + \underline{b} \;=\; \underbrace{w^* + \bar{b}}_{=0} - c_2 \iff \tilde{w} \;=\; -\underline{b} - c_2.$$

As long as:

$$\tilde{w} \;\geq\; -\bar{b} \iff c_2 \;\leq\; (\bar{b} - \underline{b}),$$

the offer is accepted by a naïve worker in period 1 and makes the firm better off when:

$$2\lambda(y - \tilde{w} - (1-\tau)\underline{b}) \;>\; 2(y - w^* - (1-\tau)\bar{b}) \iff$$
$$\lambda(y + c_2 + \tau\underline{b}) \;>\; (y + \tau\bar{b}),$$

which requires that the deviator's gain from hiring a naïve worker on worse terms exceeds the surplus generated by a high-benefit match. For an equilibrium with $b^* = \bar{b}$ and $w^* = -\bar{b}$ to exist, none of the deviations considered above can be profitable.

In sum, the following result obtains:

Proposition 1: *An equilibrium in which all firms offer low benefits ($b^* = \underline{b}$) and a monopsony wage ($w^* = -\underline{b}$) always exists. An equilibrium in which all firms offer high benefits ($b^* = \bar{b}$) and a monopsony wage ($w^* = -\bar{b}$) exists only if $c_2 < (\bar{b} - \underline{b})$, $(y + \tau\bar{b}) \geq (1-\tau)(\bar{b} - \underline{b})$, and $y + \tau\bar{b} \geq \lambda(y + c_2 + \tau\underline{b})$ hold simultaneously.*

Proposition 1 captures the effects of modeling workplace benefits as experience goods on the incentives of firms to provide high benefits. While an equilibrium with uniformly low benefits always exists, an (efficient) equilibrium with uniformly high benefits exists only if several parametric conditions hold simultaneously, ruling out various profitable deviations. In that

[16] It is assumed that aside from forgone productivity there are no other, direct turnover costs. Introducing positive turnover costs would make it less profitable for the firms to offer compensation packages that induce a worker to leave.

sense, the high-benefits equilibrium is more difficult to sustain. Note that in the present setting, the worker's option to search in period 2 provides the only incentive for the firms to ever offer high benefits.[17]

In addition, a larger share of naïve workers λ increases the profitability of the final deviation from the high-benefits equilibrium whereby a firm reduces benefits and raises its wage by a smaller amount. Thus, the presence of naïve workers exacerbates the difficulty of sustaining the efficient outcome in equilibrium.

Since the surplus and therefore profits are strictly higher in the high-benefits equilibrium, unobservability of benefits by the workers can be detrimental to the firms. In either equilibrium with a degenerate distribution of benefits and monopsony wages, both worker types receive the value of their outside option.

3.2 Equilibria with differentiation in benefits

Because sophisticated and naïve workers may adopt different search strategies when the distribution of benefits is non-degenerate, there exist equilibria with differentiated compensation packages, even though firms' costs and workers' preferences are homogeneous. Following the logic of Albrecht and Axell (1984), differentiated offers may coexist in equilibrium as long as they generate equivalent expected profits. The relevant indifference condition will reflect the firm's trade-off between profits per worker and the number of hires.

Before discussing their properties, I outline an algorithm for constructing such equilibria with differentiated compensation packages. The steps in the following derivations are:

1. Given some unknown $w_L < w_H$, and $p \equiv \mathbb{P}[w = w_H]$, assume $b(w_L) \neq b(w_H)$ and derive workers' beliefs about the distribution of offers.

2. Assuming a cost ranking of the two compensation packages, suppose that the more expensive package attracts more workers and/or allows the firm to retain a higher proportion of hired workers.

3. Rule out profitable deviations by the firms in terms of benefits.

4. Derive the workers' incentive compatibility conditions that lead to search rules in periods 1 and 2 that are consistent with the above.

[17] The finding that unobservability of benefits prior to acceptance distorts firms' incentives toward low benefit provision is consistent with the general theme in the literature on experience goods (Nelson, 1970) and has been confirmed in a variety of market settings, including those with reputation building (Shapiro, 1982, 1983) and repeated purchases (Riordan, 1986).

5. Use the workers' participation constraints and the firms' indifference condition to pin down w_L, w_H, and p.

6. Rule out profitable deviations by the firms in terms of wages.

7. Collect the equilibrium conditions and verify their mutual consistency.

3.2.1 Compensating equilibria

Following the terminology in Sorkin (2018), consider the case of a *compensating equilibrium* in pure strategies, where $b(w_H) = \underline{b}$ and $b(w_L) = \bar{b}$. That is, jobs that provide low benefits compensate by paying higher wages, and vice versa. The following derivations will illustrate that the resulting wage differential $(w_H - w_L)$ is not necessarily equal to, and in fact may be strictly smaller than, the difference in workers' valuation of benefits $(\bar{b} - \underline{b})$.

Suppose that fraction $p \in (0, 1)$ of firms offer the high-wage, low-benefit package, while fraction $(1 - p)$ offer the low-wage, high-benefit package. Sophisticated workers correctly infer what benefits are associated with a given wage offer, while naïve workers mistakenly believe that any wage offer comes with high benefits with probability $(1 - p)$ and low benefits with probability p. Thus, the two types have the following perceived option value of leaving employment and searching again in period 2:

$$(v_2^S - c_2) = \max\{p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2\,;\,(w_H + \underline{b}) - c_2/p\,;\,(w_L + \bar{b}) - c_2/(1-p)\},$$
$$(v_2^N - c_2) = \max\{p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2\,;\,w_H + p\underline{b} + (1-p)\bar{b} - c_2/p\},$$

depending on the adopted search strategy. Notice that in addition to accepting the first sampled offer, a sophisticated worker may search specifically for either $(w_H, \underline{b})$ or $(w_L, \bar{b})$, depending on which compensation package yields higher overall utility. A naïve worker, by contrast, would only either accept the first offer or search specifically for w_H.

In constructing these equilibria, it will be useful to distinguish between two cases, based on the cost ranking of the two compensation packages. Consider the following possibilities in turn.

Case 1: $w_H + (1-\tau)\underline{b} > w_L + (1-\tau)\bar{b}$.
Since the high-wage, low-benefit package is more costly to provide, the equal profits condition requires that firms offering it attract a larger mass of workers, retain a higher proportion of the workers they hired, or both. Note, however, that a compensation package including low benefits cannot lead to strictly lower turnover than a compensation package with high benefits, as this would contradict the optimality of the workers' search strategies in period 1 (Corollary 1). Thus, in order to satisfy the equal profits condition, the firms offering

$(w_H, \underline{b})$ must attract a larger mass of workers. Crucially, this requires that naïve workers search specifically for a high-wage offer in period 1, because in any compensating equilibrium their perceived valuation of sampling w_H is higher than that of sophisticated workers.

Below, I construct an equilibrium in which sophisticated workers accept the first offer sampled in period 1, naïve workers search specifically for w_H, and there is no turnover on the equilibrium path. For a construction of an alternative equilibrium with positive turnover of naïve workers, see Appendix E.

Expecting to remain in employment, a sophisticated worker accepts the first offer sampled in period 1 if:

$$u_1^S(w_H) - u_1^S(w_L) \;=\; 2((w_H - w_L) - (\bar{b} - \underline{b})) \;\leq\; c_1/p$$

and:

$$u_1^S(w_L) - u_1^S(w_H) \;=\; 2((\bar{b} - \underline{b}) - (w_H - w_L)) \;\leq\; c_1/(1-p).$$

For $(w_L, \bar{b})$ to be offered in equilibrium, workers should leave upon discovering $(w_L, \underline{b})$. In this case, the relevant condition applies to sophisticated workers, who are the only type accepting w_L in period 1. They would leave low-wage, low-benefit jobs in period 2 if:

$$w_L + \underline{b} \;<\; p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2 \;\iff\; c_2 \;<\; p(w_H - w_L) + (1-p)(\bar{b} - \underline{b}),$$

since accepting the first sampled offer in period 1 implies that the same strategy must be optimal in period 2. Notice that when $(v_2^S - c_2) = p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2$, this implies $(v_2^N - c_2) \geq (v_2^S - c_2)$. Consequently, naïve workers also expect to leave jobs offering $(w_L, \underline{b})$. They search for w_H in period 1 as long as:

$$u_1^N(w_H) - u_1^N(w_L) \;=\; (w_H - w_L) + p\big(\max\{(w_H + \underline{b})\,;\,(v_2^N - c_2)\} - (v_2^N - c_2)\big) + (1 - p)\big((w_H + \bar{b}) - \max\{(w_L + \bar{b})\,;\,(v_2^N - c_2)\}\big) \;>\; c_1/p.$$

What is $(v_2^N - c_2)$ in this case? Suppose that when searching in period 2, naïve workers also search for w_H, which requires:

$$(w_H - w_L) \;>\; c_2/p.$$

Then, $(v_2^N - c_2) = w_H + p\underline{b} + (1-p)\bar{b} - c_2/p$. Both types stay employed in high-wage, low-benefit jobs provided that:

$$w_H + \underline{b} \;\geq\; (v_2^N - c_2) \;\iff\; c_2/p \;\geq\; (1-p)(\bar{b} - \underline{b}),$$

since $(v_2^N - c_2) > (v_2^S - c_2)$. Even though naïve workers never accept low-wage offers in period 1, we need to specify their expectations regarding low-wage, high-benefit jobs. To isolate the case that does not result in a contradiction, suppose that naïve workers expect to stay in such jobs:

$$w_L + \bar{b} \;\geq\; (v_2^N - c_2) \;\;\Longleftrightarrow\;\; c_2/p \;\geq\; (w_H - w_L) - p(\bar{b} - \underline{b}).$$

Then, naïve workers indeed search for w_H in period 1 as long as:

$$u_1^N(w_H) - u_1^N(w_L) \;=\; (w_H - w_L) + p\big((w_H + \underline{b}) - w_H - p\underline{b} - (1-p)\bar{b} + c_2/p\big) + (1 - p)\big((w_H + \bar{b}) - (w_L + \bar{b})\big) \;>\; c_1/p \;\;\Longleftrightarrow$$
$$c_1/p \;<\; (2-p)(w_H - w_L) - p(1-p)(\bar{b} - \underline{b}) + c_2.$$

Next, we take into account the workers' participation constraints (PC). Keeping in mind that the first search is free, the relevant PC of a worker that guarantees their full participation in the labor market is the one associated with sampling their (weakly) less preferred offer first. These *ex-post PC's* of a sophisticated worker, who samples exactly once, are:

$$u_1^S(w_L) \;=\; 2(w_L + \bar{b}) \;\geq\; 0$$

and:

$$u_1^S(w_H) \;=\; 2(w_H + \underline{b}) \;\geq\; 0.$$

While that of a naïve worker, who searches specifically for a high-wage offer, is:

$$u_1^N(w_H) - c_1/p \;=\; 2(w_H + p\underline{b} + (1-p)\bar{b}) - c_1/p \;\geq\; 0.$$

A binding ex-post PC of a sophisticated worker who samples a low-wage offer implies $w_L = -\bar{b}$, but which of the two remaining PC's should determine w_H? Observe that it cannot be the case that the ex-post PC of a naïve worker is binding ($u_1^N(w_H) - c_1/p = 0$), while that of a sophisticated worker is slack ($u_1^S(w_H) > 0$). If that was the case, a firm could marginally reduce its wage offer below w_H without deterring any workers: sophisticated workers with pessimistic beliefs would still accept as long as $u_1^S(w_H - \epsilon) \geq 0$, and so would naïve workers with passive beliefs as long as $u_1^N(w_H - \epsilon) = u_1^N(w_H) - 2\epsilon \geq u_1^N(w_H) - c_1/p$.

Thus, in equilibrium, a binding ex-post PC of a sophisticated worker also determines w_H, implying $w_H = -\underline{b}$ and therefore $(w_H - w_L) = (\bar{b} - \underline{b})$. Then, a marginal reduction of a wage offer below w_H would result in all sophisticated workers rejecting, leading to a discrete loss in profits. Moreover, $u_1^N(w_H) - c_1/p \geq 0$ for $w_H = -\underline{b}$ as long as:

$$c_1/(2p) \;\leq\; (1-p)(\bar{b} - \underline{b}).$$

As for the firms, the equal profits condition boils down to:

$$\pi(w_H) = 2((1-\lambda)+\tfrac{\lambda}{p})(y-w_H-(1-\tau)\underline{b}) \;=\; 2(1-\lambda)(y-w_L-(1-\tau)\bar{b}) = \pi(w_L) \iff$$
$$(1-\lambda)\,/\,((1-\lambda)+\tfrac{\lambda}{p}) \;=\; \tfrac{(y-w_H-(1-\tau)\underline{b})}{(y-w_L-(1-\tau)\bar{b})}.$$

While a deviation from $(w_H, \underline{b})$ to $(w_H, \bar{b})$ is strictly unprofitable, a deviation from $(w_L, \bar{b})$ to $(w_L, \underline{b})$ is not profitable as long as:

$$2(1-\lambda)(y-w_L-(1-\tau)\bar{b}) \;\geq\; (1-\lambda)(y-w_L-(1-\tau)\underline{b}) \iff$$
$$2 \;\geq\; \tfrac{(y-w_L-(1-\tau)\underline{b})}{(y-w_L-(1-\tau)\bar{b})}.$$

The fact that the ex-post PC of a sophisticated worker associated with sampling w_H is binding implies that a deviation to any wage offer $w \in (w_L, w_H)$ would attract only naïve workers. Let

$$\tilde{w} \;=\; \max\{w_H + (1-p)(\bar{b}-\underline{b}) - c_2/p;\, w_H - c_1/(2p)\} \tag{3}$$

be the lowest wage offer that allows the firm to hire and retain a naïve worker, when combined with low benefits. The first threshold in the curly brackets guarantees retention in period 2, and the second guarantees that a naïve worker prefers to accept $\tilde{w}$ in period 1 rather than search further. Then, the deviating firm is not better off as long as:

$$2\tfrac{\lambda}{p}(y-\tilde{w}-(1-\tau)\underline{b}) \;\leq\; 2((1-\lambda)+\tfrac{\lambda}{p})(y-w_H-(1-\tau)\underline{b}).$$

Finally, a deviation to some intermediate wage combined with high benefits, which attracts only naïve workers, is dominated by $(\tilde{w}, \underline{b})$ if the reduction in the wage offer relative to $\tilde{w}$ (so that a naïve worker stays upon discovering high benefits but would leave otherwise) is smaller than the increase in the cost of benefits.

A naïve worker stays upon discovering $\bar{b}$ if $w \geq w_H - p(\bar{b}-\underline{b}) - c_2/p$ but leaves upon discovering $\underline{b}$ if $w < w_H + (1-p)(\bar{b}-\underline{b}) - c_2/p \leq \tilde{w}$. Then, she accepts the deviator's offer in period 1 if:

$$w + p(\underline{b} + (v_2^N - c_2)) + (1-p)(w+2\bar{b}) \;\geq\; 2(w_H + p\underline{b} + (1-p)\bar{b}) - c_1/p \iff$$
$$w \;\geq\; w_H - c_1/(p(2-p)) + c_2/(2-p) - (\bar{b}-\underline{b})p(1-p)/(2-p).$$

Thus, letting

$$\hat{w} \;\equiv\; \max\{w_H - p(\bar{b}-\underline{b}) - c_2/p;\, w_H - c_1/(p(2-p)) + c_2/(2-p) - (\bar{b}-\underline{b})p(1-p)/(2-p)\}, \tag{4}$$

if

$$\hat{w} + (1-\tau)\bar{b} \;\geq\; \tilde{w} + (1-\tau)\underline{b},$$

then such a deviation is not profitable.

In sum, for such a compensating equilibrium to exist, all of the following must hold:

(i) $c_1/p \;\geq\; 2((w_H - w_L) - (\bar{b} - \underline{b})),$

(ii) $c_1/(1-p) \;\geq\; 2((\bar{b} - \underline{b}) - (w_H - w_L)),$

(iii) $c_2 \;<\; p(w_H - w_L) + (1-p)(\bar{b} - \underline{b}),$

(iv) $c_2/p \;<\; (w_H - w_L),$

(v) $c_2/p \;\geq\; (1-p)(\bar{b} - \underline{b}),$

(vi) $c_2/p \;\geq\; (w_H - w_L) - p(\bar{b} - \underline{b}),$

(vii) $c_1/p \;<\; (2-p)(w_H - w_L) - p(1-p)(\bar{b} - \underline{b}) + c_2,$

(viii) $(1-\lambda) \,/\, ((1-\lambda) + \frac{\lambda}{p}) \;=\; \frac{(y - w_H - (1-\tau)\underline{b})}{(y - w_L - (1-\tau)\bar{b})},$

(ix) $2 \;\geq\; \frac{(y - w_L - (1-\tau)\underline{b})}{(y - w_L - (1-\tau)\bar{b})},$

(x) $c_1/(2p) \;\leq\; (1-p)(\bar{b} - \underline{b}),$

(xi) $w_H \;=\; -\underline{b},$

(xii) $w_L \;=\; -\bar{b},$

(xiii) $\frac{\lambda}{p}(y - \tilde{w} - (1-\tau)\underline{b}) \;\leq\; ((1-\lambda) + \frac{\lambda}{p})(y - w_H - (1-\tau)\underline{b}),$ $\tilde{w}$ given by (3),

(xiv) $\hat{w} + (1-\tau)\bar{b} \;\geq\; \tilde{w} + (1-\tau)\underline{b},$ $\hat{w}$ given by (4).

Plugging in $w_L = -\bar{b}$ and $w_H = -\underline{b}$, the above simplifies to:

(iv') $c_2/p \;<\; (\bar{b} - \underline{b}),$

(v') $c_2/p \;\geq\; (1-p)(\bar{b} - \underline{b}),$

(vii') $c_1/p \;<\; (2 - 2p + p^2)(\bar{b} - \underline{b}) + c_2,$

(viii') $(1-\lambda) \,/\, ((1-\lambda) + \frac{\lambda}{p}) \;=\; \frac{(y + \tau\underline{b})}{(y + \tau\bar{b})},$

(ix') $2 \;\geq\; \frac{(y + (\bar{b} - \underline{b}) + \tau\underline{b})}{(y + \tau\bar{b})},$

(x') $c_1/(2p) \;\leq\; (1-p)(\bar{b} - \underline{b}),$

(xiii') $\frac{\lambda}{p}(y - \tilde{w} - (1-\tau)\underline{b}) \leq ((1-\lambda) + \frac{\lambda}{p})(y + \tau\underline{b})$, $\tilde{w}$ given by (3),

(xiv') $\hat{w} + (1-\tau)\bar{b} \geq \tilde{w} + (1-\tau)\underline{b}$, $\hat{w}$ given by (4).

Indeed, these conditions can hold simultaneously.[18] In such an equilibrium, naïve workers bear higher expected search costs as they search too intensely for high-wage offers in period 1. Positive turnover of naïve workers is not necessary to support it, but can feature in such equilibria, see Appendix E.

Case 2: $w_H + (1-\tau)\underline{b} \leq w_L + (1-\tau)\bar{b}$**.**
When the high-wage, low-benefit package is less costly to provide, it must deliver strictly lower utility to the workers. The equal profits condition, in turn, requires that the firms offering the more expensive low-wage, high-benefit package attract more workers, retain a higher proportion of the workers they hired, or both.

Here, I construct an equilibrium in which both worker types remain in initially accepted employment, but low-wage, high-benefit jobs attract more workers. Given the corresponding utility ranking, this requires that sophisticated workers search specifically for $(w_L, \bar{b})$ in period 1. In any other case, the naïve workers' higher perceived valuation of sampling a high-wage offer would imply that high-wage firms attract a weakly greater mass of workers, contradicting the equal profits condition.

Sophisticated workers indeed search specifically for $(w_L, \bar{b})$ if:

$$u_1^S(w_L) - u_1^S(w_H) = 2((w_L + \bar{b}) - (w_H + \underline{b})) > c_1/(1-p).$$

Suppose for concreteness that when searching in period 2 a sophisticated worker would accept the first sampled offer, which requires:

$$(w_L + \bar{b}) - (w_H + \underline{b}) \leq c_2/(1-p).$$

Then, $(v_2^S - c_2) = p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2$. Clearly, a sophisticated worker has no incentive to leave a low-wage, high-benefit job which provides the highest overall utility. In case of a deviation to low benefits, she would leave a job providing $(w_L, \underline{b})$ if:

$$w_L + \underline{b} < p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2 \iff c_2 < p(w_H - w_L) + (1-p)(\bar{b} - \underline{b}),$$

but stays with $(w_H, \underline{b})$ if:

$$w_H + \underline{b} \geq p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2 \iff c_2 \geq (1-p)((w_L + \bar{b}) - (w_H + \underline{b})),$$

[18] For example for $\lambda = 0.086$, $\bar{b} = 41.409$, $\underline{b} = 11.080$, $c_1 = 5.508$, $c_2 = 12.345$, $\tau = 0.350$, and $y = 54.999$, which solve for $p = 0.521$, $w_H = -11.080$, $w_L = -41.409$, $\tilde{w} = -16.364$, and $\hat{w} = -14.996$.

which coincides with the condition for accepting the first sampled offer in period 2.

In turn, a naïve worker also accepts the first sampled offer in period 2 if:

$$c_2/p \;\geq\; (w_H - w_L),$$

which implies $(v_2^N - c_2) = (v_2^S - c_2)$. Then, upon discovering the value of benefits, the behavior of a naïve worker coincides with that of her sophisticated counterpart. Suppose that in contrast to a sophisticated worker, a naïve worker searches specifically for a high-wage offer in period 1:[19]

$$u_1^N(w_H) - u_1^N(w_L) \;=\; (w_H - w_L) + p\big((w_H + \underline{b}) - p(w_H + \underline{b}) - (1-p)(w_L + \bar{b}) + c_2\big) + (1 - p)\big((w_H + \bar{b}) - (w_L + \bar{b})\big) \;=$$
$$(2 - p^2)(w_H - w_L) - p(1-p)(\bar{b} - \underline{b}) + pc_2 \;>\; c_1/p.$$

Next, the binding ex-post PC of a sophisticated worker pins down w_L:

$$u_1^S(w_L) - c_1/(1-p) \;=\; 2(w_L + \bar{b}) - c_1/(1-p) \;=\; 0,$$

while that of a naïve worker guaranteeing full labor market participation would require:

$$u_1^N(w_H) - c_1/p \;\geq\; 0.$$

Note, however, a subtle point that the latter cannot hold in equilibrium where only naïve workers accept high-wage offers, even when binding. Suppose that $u_1^N(w_H) - c_1/p = 0$. Then, marginally reducing the wage offer below w_H would necessarily make a firm better off: while sophisticated workers would continue searching for a low-wage, high-benefit offer, all naïve workers who sample the firm's offer would accept as long as $u_1^N(w_H - \epsilon) = u_1^N(w_H) - 2\epsilon \geq u_1^N(w_H) - c_1/p$. Thus, in equilibrium, only the ex-post PC of a naïve worker associated with sampling their preferred offer first can hold as a binding constraint:

$$u_1^N(w_H) \;=\; 2(w_H + p\underline{b} + (1-p)\bar{b}) \;=\; 0.$$

This means that only those of naïve workers who sample a high-wage offer immediately end up participating in the labor market. In contrast, w_L compensates sophisticated workers for their search costs, and therefore ensures full labor market participation, because a marginal reduction of the wage offer below w_L would be interpreted as a signal of low benefits and thus rejected in favor of further search.

[19] When naïve workers are the only type accepting high-wage offers in equilibrium, their PC pins down w_H. A binding PC of the form $u_1^N(w_H) = 0$ (see below) implies $u_1^N(w_L) < 0$, thus ruling out the optimality of accepting a low-wage offer.

The above conditions highlight why one should expect a "compressed" wage differential to arise in this type of equilibria, i.e. a wage differential that does not fully compensate the workers for the difference in benefits across the compensation packages. In compensating equilibria, naïve workers have a stronger (perceived) incentive to search for high-wage offers. When they are in fact the only type accepting high-wage, low-benefit jobs, this distorts w_H downward, relative to $-\underline{b}$, because naïve workers mistakenly believe that w_H is associated with high benefits with some strictly positive probability. In turn, when sophisticated workers are searching specifically for $(w_L, \bar{b})$, their search costs distort w_L upward, relative to $-\bar{b}$. Both effects lead to a compressed wage differential $(w_H - w_L) < (\bar{b} - \underline{b})$.

As for the firms, the equal profits condition adjusted for partial participation of naïve workers in the labor market is:

$$\pi(w_H) \;=\; 2\,\tfrac{\lambda\times p}{p}\,(y - w_H - (1-\tau)\underline{b}) \;=\; 2\,\tfrac{(1-\lambda)}{(1-p)}\,(y - w_L - (1-\tau)\bar{b}) \;=\; \pi(w_L) \;\iff$$

$$\tfrac{\lambda\times(1-p)}{(1-\lambda)} \;=\; \tfrac{(y-w_L-(1-\tau)\bar{b})}{(y-w_H-(1-\tau)\underline{b})}.$$

A deviation from $(w_H, \underline{b})$ to $(w_H, \bar{b})$ is never profitable. A deviation from $(w_L, \bar{b})$ to $(w_L, \underline{b})$ is not profitable as long as:

$$2\,\tfrac{(1-\lambda)}{(1-p)}\,(y - w_L - (1-\tau)\bar{b}) \;\geq\; \tfrac{(1-\lambda)}{(1-p)}\,(y - w_L - (1-\tau)\underline{b}).$$

Finally, any deviation to $w \in (w_L, w_H)$ would attract no workers: sophisticated workers would continue searching for $(w_L, \bar{b})$ and naïve workers would reject in favor of the outside option.

Taken together, for such an equilibrium to exist, all of the following must hold:

(i) $c_1/(1-p) \;<\; 2((w_L + \bar{b}) - (w_H + \underline{b}))$,

(ii) $c_2/(1-p) \;\geq\; (w_L + \bar{b}) - (w_H + \underline{b})$,

(iii) $c_2 \;<\; p(w_H - w_L) + (1-p)(\bar{b} - \underline{b})$,

(iv) $c_2/p \;\geq\; (w_H - w_L)$,

(v) $c_1/p \;<\; (2 - p^2)(w_H - w_L) - p(1-p)(\bar{b} - \underline{b}) + pc_2$,

(vi) $\tfrac{\lambda\times(1-p)}{(1-\lambda)} \;=\; \tfrac{(y-w_L-(1-\tau)\bar{b})}{(y-w_H-(1-\tau)\underline{b})}$,

(vii) $2(y - w_L - (1-\tau)\bar{b}) \;\geq\; (y - w_L - (1-\tau)\underline{b})$,

(viii) $w_L \;=\; -\bar{b} + c_1/(2(1-p))$,

(ix) $w_H \;=\; -p\underline{b} - (1-p)\bar{b}$.

Indeed, these conditions can hold simultaneously.[20] In contrast to the previous case, in such an equilibrium naïve workers bear lower expected search costs than sophisticated ones, as they search suboptimally little for low-wage offers in period 1. Positive turnover of naïve workers is again not necessary, but can feature in such equilibria, see Appendix E.

To summarize, the following result obtains:

Proposition 2: *Compensating equilibria, in which high wages are paired with low benefits, and vice versa, exist for certain parameter constellations. If the high-wage, low-benefit package is strictly more costly to provide, then in any such equilibrium naïve workers search too intensely for a high-wage offer in period 1. Conversely, if the high-wage, low-benefit package is weakly less costly to provide, then in any such equilibrium naïve workers search too little for a low-wage offer in period 1. Positive turnover of naïve workers is not necessary, but can arise in such equilibria.*

Despite the multiplicity of compensating equilibria and their qualitatively different predictions regarding, for example, the search behavior of naïve and sophisticated workers, the model delivers an unambiguous prediction regarding the size of the wage differential that arises under imperfect information about workplace benefits. In particular, the wage differential must be (weakly) smaller than the benefit differential:

Corollary 2: *In any compensating equilibrium, the wage differential is bounded from above by the benefit differential, i.e.* $(w_H - w_L) \leq (\bar{b} - \underline{b})$.

The proof, which is provided in Appendix F, rests on the observation that a larger wage differential, i.e. $(w_H - w_L) > (\bar{b} - \underline{b})$, would necessarily leave room for a profitable deviation for firms offering high wages and low benefits. Corollary 2 therefore establishes that the presence of naïve workers can compress the equilibrium wage differential across firms providing low and high benefits, relative to the "appropriate" compensating differential $(w_H - w_L) = (\bar{b} - \underline{b})$.

Furthermore, the equal profits condition implies that the share of naïve workers λ determines the equilibrium proportion of high-wage, low-benefit offers p. Specifically, the larger the mass of naïve workers searching for high wages, the larger the equilibrium share of such offers. In the first case above, when naïve workers search specifically for w_H in period 1 while sophisticated workers accept the first sampled offer, this has a positive effect on total welfare: although both compensation packages deliver identical utility to workers, naïve workers bear lower expected search costs. In the second case, when naïve workers search specifically for

[20] For example for $\lambda = 0.665$, $\bar{b} = 90.070$, $\underline{b} = 75.531$, $c_1 = 2.531$, $c_2 = 5.741$, $\tau = 0.217$, and $y = 142.426$, which solve for $p = 0.515$, $w_H = -82.583$, and $w_L = -87.460$.

w_H in period 1 while sophisticated workers search specifically for w_L, this has a negative effect on total welfare: a larger share of naïve workers end up participating in the labor market where their high-wage, low-benefit package provides utility strictly below the outside option, and sophisticated workers bear higher expected search costs in order to avoid it. Thus, the welfare effects of λ depend on the particular equilibrium.

In the seminal model of Rosen (1974, 1986), workers have heterogeneous preferences over benefits and firms have heterogeneous costs of providing benefits. Under perfect information, workers and firms would sort into compensation packages with either high or low benefits, with the resulting wage differential determined by the indifference condition of a marginal worker and a marginal firm. With homogeneous preferences and costs as assumed above, this framework yields a unique prediction: all firms should offer (efficiently) high benefits.

This first-best benchmark does not generally obtain under the kind of information frictions introduced here. Even when the wage differential corresponds exactly to the workers' valuation of benefits, i.e. $(w_H - w_L) = (\bar{b} - \underline{b})$, such compensating equilibria remain inefficient because there coexist compensation packages that provide identical utility to workers while generating different provision costs for firms.

Finally, the above spurious differentiation in compensation packages, arising from the unobservability of benefits and despite the homogeneity of workers' preferences and firms' costs, is qualitatively different from *pure wage dispersion* predicted by classical search models such as Burdett and Judd (1983), Stahl (1989), or Burdett and Mortensen (1998). In those models, specific search frictions (e.g., non-sequential search or random arrival of outside offers) generate equilibrium dispersion along the wage dimension. However, they would still predict that all firms should offer high benefits provided that these are observable to workers. That is because all firms would find it optimal to offer the uniquely efficient (high) benefit level, irrespective of any randomization over the wage.

3.2.2 Augmenting equilibria

Do there also exist *augmenting equilibria* in which $b(w_H) = \bar{b}$ and $b(w_L) = \underline{b}$? This would mean the coexistence of offers that are strictly dominant and strictly dominated along both dimensions of a compensation package. In Appendix G, I show that any such equilibrium must have two key properties (Lemma A1). First, only naïve workers accept low-wage offers in period 1, while sophisticated workers search specifically for a high-wage, high-benefit offer. Second, the dominated compensation packages induce strictly higher turnover, meaning that all naïve workers who initially accept a low-wage offer leave in period 2 upon discovering low benefits. The resulting equilibrium conditions indicate that low-wage, low-benefit jobs are necessarily exploitative in the sense that they deliver utility strictly below the worker's outside option.

While these properties are intuitive, I was only able to construct numerical examples in which all but one of the equilibrium conditions are satisfied simultaneously, with the remaining condition failing by a vanishingly small margin. These examples point to a limiting case in which firms offer the same wage but differentiate their benefits. Since correlation neglect plays no role under such a distribution of offers, the search rules of naïve and sophisticated workers coincide. In particular, workers accept the first offer sampled in period 1 and leave in period 2 if the realized benefits are low. Firms, in turn, are indifferent between retaining workers with high benefits and losing them with low benefits. In such equilibria, strictly dominant and dominated compensation packages still coexist, but dominance is restricted to the benefit dimension.

3.3 Discussion

Discrete choice of benefits. How would the results be affected if instead of a discrete choice of benefits $b \in \{\underline{b}, \bar{b}\}$, the firms could select any benefit level from an interval $b \in [\underline{b}, \bar{b}]$? Consider a putative equilibrium with a degenerate distribution of benefits and a monopsony wage $w(b) = -b$ (Lemma 3). Suppose that a firm deviates by reducing its benefits by some $\epsilon > 0$. The deviator does not lose the worker who accepted, and is therefore strictly better off, as long as:

$$w(b) + b - c_2 \leq w(b) + b - \epsilon \iff \epsilon \leq c_2.$$

Since $c_2 > 0$, one can always find a small enough increment ϵ such that the above holds. Thus, for any $b > \underline{b}$, there exists a profitable deviation in terms of benefits and a result that is effectively an extreme version of Proposition 1 obtains. Namely, when workers cannot observe the benefits prior to acceptance and firms' choice of benefits is continuous, there do not exist equilibria with uniform benefits exceeding $\underline{b}$.

Analogously, equilibria with differentiated offers (Proposition 2) no longer exist if the firms offering high benefits are able to reduce those by an arbitrarily small amount. Thus, in the current setting, the assumption of discrete choice of benefits by the firms is essential for supporting equilibria with benefits exceeding $\underline{b}$.

Note, however, that this strong caveat applies here because all workers are assumed to have identical, positive search costs in period 2. In order to restore equilibria with benefits exceeding $\underline{b}$ when the choice of benefits is continuous, one could follow the literature and introduce a fraction of "shoppers" with zero search costs (Stahl, 1989) or uncertainty about the outside option (Burdett and Mortensen, 1998), for instance.

Off-path beliefs. What role does the assumption of sophisticated workers' pessimistic beliefs play in sustaining the above equilibria and what might be a viable alternative? Because firms have homogeneous provision costs, the wage offer alone cannot credibly signal the benefit level. Thus, to discipline off-path beliefs one must invoke an intuitive-criterion-type argument, whereby a sophisticated worker reasons through the firms' incentives to deviate to a non-equilibrium wage offer.

Such reasoning would not refine pessimistic beliefs in two classes of equilibria. First, in an equilibrium where all firms offer high benefits and a monopsony wage, no firm has a profitable deviation to high benefits paired with any other acceptable wage. Second, in compensating equilibria, no deviation in terms of the wage can be specific to a high-benefit firm: low-wage firms already combine high benefits with the lowest wage consistent with workers' acceptance, and high-wage, low-benefit firms are indifferent. Since there are no deviations that would be strictly profitable for a high-benefit firm but weakly unprofitable for a low-benefit firm, the intuitive criterion does not refine pessimistic beliefs in these equilibria.

The only kind of equilibrium for which this refinement may matter is the least efficient equilibrium in which all firms offer low benefits and a monopsony wage $w^* = -\underline{b}$. Consider a deviation to high benefits and some lower wage $\hat{w} < w^*$. If a sophisticated worker believed that the deviating firm provides high benefits, then $\hat{w} = -\bar{b}$ would make her indifferent between accepting and taking up the outside option. Moreover, when:

$$\hat{w} + \underline{b} \; < \; w^* + \underline{b} - c_2 \quad \Longleftrightarrow \quad c_2 \; < \; (\bar{b} - \underline{b})$$

and:

$$2(y + \bar{b} - (1-\tau)\bar{b}) \; \geq \; (y + \bar{b} - (1-\tau)\underline{b}) \quad \Longleftrightarrow \quad 2(y + \tau\bar{b}) \; \geq \; (y + (\bar{b} - \underline{b}) + \tau\underline{b}),$$

then the deviating firm indeed prefers to pair $\hat{w}$ with high benefits and retain the worker, rather than to pair it with low benefits and lose the worker in period 2. If, in addition:

$$(y + (\bar{b} - \underline{b}) + \tau\underline{b}) \; \leq \; 2(y + \tau\underline{b}),$$

then only a high-benefit firm would benefit from a deviation to $\hat{w}$ and the intuitive criterion would require sophisticated workers to adjust their beliefs accordingly. Thus, the only qualitative difference from the above predictions derived under Assumption 1 would be an additional parametric condition for the existence of a low-benefits equilibrium.

Worker's utility function. The predictions of the model would carry over, in a qualitative sense, to specifications with a more general additive utility function, such as:

$$u_{ik} \; = \; w_k + q(b_k),$$

where $q(\cdot)$ is strictly increasing and weakly concave. Moreover, the predicted outcomes would have the same efficiency properties as discussed above provided that $q'(\bar{b}) > (1 - \tau)$.

Probabilistic revelation. The above formulation of the model assumes that after acceptance of an offer the worker observes the value of the associated benefits with certainty. In reality, this process is more likely probabilistic, reflecting factors such as information provided exogenously by one's union representatives (Gustman and Steinmeier, 2005) or peers (Duflo and Saez, 2002, 2003). Introducing probabilistic revelation would not qualitatively affect the model's predictions, although it would further reduce the firm's incentives to provide high benefits.

Commitment vs. renegotiation. Firms are assumed to commit to the same compensation package for both periods. This is consistent with standard models of wage posting (Rogerson et al., 2005), but also crucial for studying the firms' incentives to provide high benefits despite their unobservability during job search, with the advantage of high benefits being lower turnover experienced in period 2. Introducing the possibility to renegotiate, or reset, the worker's wage in period 2 would eliminate any incentive to offer high benefits: a firm could simply set its period-2 wage to the minimal level required to retain the worker, irrespective of b.

Two periods. Relatedly, firms would have no incentive to offer high benefits in a one-shot version of the model that does not account for the realization of benefits and resulting turnover. Furthermore, despite the assumption of finite horizon, the above results are not driven by end-of-game effects precisely because of the firms' commitment to the initially offered compensation package. As the derivations suggest, however, extending the finite horizon to $T > 2$ periods would substantially increase complexity without necessarily generating additional insights. Since the fundamental trade-off between profits per worker and the number of hires and retentions would carry over to an infinite-horizon version of the model, so would the above qualitative predictions.

Behavioral types. What role do specific assumptions about workers' beliefs (on the equilibrium path) play in generating the above predictions? In any pure-strategy equilibrium, sophisticated workers associate the correct benefit level with each wage offer. Still, the assumption of unobservability prior to acceptance determines the profitability of potential deviations by the firms: it makes deviations to low benefits more profitable, and vice versa, relative to the case of observable benefits.

Naïve workers, in turn, may appear to search as if they focused exclusively on the observable wage component. Nevertheless, the fact that they anticipate future realization of benefits is reflected in their search strategy in the form of the option value of searching again in period 2, which distinguishes them from the myopic agents typically considered in the behavioral industrial organization literature.

To illustrate, recall that Gabaix and Laibson (2006) consider a model of add-on pricing with two types of consumers. In equilibrium where the add-on prices are shrouded, akin to the benefits here, "sophisticated consumers" apply Bayesian updating to form their beliefs and "myopic consumers" completely ignore the add-on price. Although the ranking of offers would be identical across myopic consumers and naïve workers, their perceived benefit of further search may still differ. Specifically, myopic consumers would always search (weakly) more intensely for high-wage offers than naïve workers:

$$\begin{aligned} u_1^N(w_H) - u_1^N(w_L) &= (w_H - w_L) + p\Big(\underbrace{(w_H + \bar{b}) - \max\{(w_L + \bar{b})\,;\,(v_2^N - c_2)\}}_{\leq (w_H - w_L)}\Big) \\ &\quad + (1-p)\Big(\underbrace{\max\{(w_H + \underline{b})\,;\,(v_2^N - c_2)\} - \max\{(w_L + \underline{b})\,;\,(v_2^N - c_2)\}}_{\leq (w_H - w_L)}\Big) \\ &\leq 2(w_H - w_L) \;=\; u_1^{Myop}(w_H) - u_1^{Myop}(w_L), \end{aligned}$$

where $p \equiv \mathbb{P}[b = \bar{b}]$. This implies that for any non-degenerate distribution of offers, myopic consumers would adopt a (weakly) higher reservation wage than naïve workers, or $R_1^{Myop} \geq R_1^N$. Consequently, myopia only exacerbates the kind of mistake that naïve workers make in compensating equilibria (Proposition 2). At the same time, invoking myopia instead of correlation neglect to model the strategic error of naïve workers is conceptually less appealing in the present setting with non-negative workplace benefits. That is because myopic workers would require strictly higher wages to satisfy their PC than sophisticated workers and would never leave a previously accepted job upon realizing the benefits, resulting in a trivial lack of turnover.

By contrast, comparison with Schumacher (2024) is less straightforward. Applying the concept of personal equilibrium (Kőszegi, 2010; Spiegler, 2016), Schumacher (2024) considers a model with a boundedly rational consumer who correctly anticipates the add-on price of a product she purchases on the equilibrium path but potentially misperceives it for an alternative she never selects, believing that the product choice has no impact on the probability of being charged a positive add-on price. This kind of belief formation would imply that an agent has correct beliefs about the benefits associated with the wage offers she accepts in equilibrium and expects the same benefit to materialize even if she adopted a different reservation wage. In case of a compensating equilibrium, the resulting decision rule is therefore not directly comparable to that of a naïve worker in the present model, who always obtains new information on the equilibrium path.

4 Empirical relevance

In this section I relate the model's predictions to several empirical puzzles in the literature on workplace benefits, thus highlighting the importance of accounting for information frictions. While the model admits multiple equilibria with different qualitative properties, the discussion focuses on general predictions that reflect the model's core mechanics and underlie all equilibria characterized above. I then consider some policy implications.

Inefficiently low benefits. The model identifies a specific obstacle to delegating benefit provision to employers. Namely, when benefits are unobservable to workers prior to acceptance, this distorts firms' incentives toward providing inefficiently low benefits (e.g., Proposition 1).

This mechanism provides a novel rationale for the findings of Cole and Taska (2023). Combining data on job transitions with an online experiment, they find that a majority of workers have a high willingness to pay (in terms of their total compensation) for access to a workplace pension plan and for employer contributions. Nevertheless, a calibrated model of on-the-job search implies that 80% of firms could improve their hiring probability by shifting some of the total compensation toward higher benefits.

Cole and Taska (2023) attribute this finding to regulatory constraints on designing worker-specific compensation packages, in particular non-discrimination rules for workplace pensions. While this can explain why the distribution of benefits would be relatively compressed within firms, it is not clear why it should imply specifically low benefit provision. The fact that most workers in their sample indeed place a high value on workplace benefits suggests that information frictions may well play an important complementary role.

Benefits and turnover. The model predicts that due to workers discovering the value of benefits on the job, firms providing low benefits experience (weakly) higher turnover (Corollary 1). This resonates with empirical evidence indicating that lower benefits are indeed associated with higher worker turnover (Bennett et al., 1993; Lee et al., 2006; Ouimet and Tate, 2023). This raises a question: why do workers accept such jobs in the first place, or why don't firms combine an increase in benefits with a wage reduction to minimize turnover? The model offers an explanation based on information frictions on the worker side, specifically the difficulty of evaluating benefits at the time of job search.[21]

[21]When discussing potential mechanisms that can explain their findings, Ouimet and Tate (2023) write: "To the degree that these factors [regulatory and administrative constraints] induce unwanted convergence of total compensation across workers, the firm could conceivably undo the effects by adjusting wages. However,

Large between-firm dispersion in benefits. Recent work by Ouimet and Tate (2023) shows that the observed benefits inequality (Pierce, 2001; Kristal et al., 2020) is driven by large between-firm dispersion in benefits, rather than by within-firm dispersion. While alternative mechanisms can rationalize the compression of benefits within a firm (most importantly, various administrative and regulatory constraints), they do not explain why the dispersion across firms should be so large, or why it should exceed the wage dispersion.

The model sheds light on these patterns by showing how information frictions can lead to spurious differentiation in compensation packages across firms, even in the absence of heterogeneity in preferences, costs, or productivity. In such equilibria with differentiation, the wage differential turns out to be bounded from above by the benefit differential (Corollary 2). Intuitively, the presence of naïve workers compresses the equilibrium wage differential because they fail to fully take into account the difference in benefits provided by different firms.

Estimating compensating differentials. Despite its conceptual appeal, the theory of compensating differentials has enjoyed only limited empirical success, see Schiller and Weiss (1980) and Montgomery and Shaw (1997) for early studies on workplace pension benefits. Despite access to richer data and more sophisticated estimation techniques, rationalizing empirical estimates of compensating differentials remains difficult. One possible reason, typically absent from standard frameworks, is the presence of information frictions (Lavetti, 2023).[22]

The model offers a formal explanation for these empirical challenges: workers' imperfect understanding of workplace benefits at the time of job search can lead to equilibria with spurious differentiation in compensation packages, in which the wage differential does not reflect the preferences of a marginal worker or the costs of a marginal firm (Proposition 2, Corollary 2). In particular, when $(w_H - w_L) < (\bar{b} - \underline{b})$ due to workers' imperfect information, the observed wage differential undercompensates for the underlying benefit differential. As a result, a researcher may rationalize the allocation of workers with identical preferences into different compensation packages by attributing it to unobserved "residual amenities" – overestimating amenities at high-wage, low-benefit firms and underestimating them at low-

there are significant constraints on firms' ability to do so, given minimum wage laws, downward wage rigidities, and differences in the salience and individual relative valuations of wages and benefits. [...] Limited information can constrain the ability of new hires to value nonwage benefits prior to joining the firm, thus making wages relatively more salient when comparing multiple job offers."

[22]See Sorkin (2018) and Lamadon et al. (2022) for approaches that go beyond hedonic regressions and incorporate various labor market frictions, but nevertheless assume that workers are perfectly informed about pecuniary and non-pecuniary amenities. Regarding workers' beliefs, Belot et al. (2022) and Sockin (2024) find that job seekers tend to perceive high-wage firms as providing better amenities.

wage, high-benefit firms – even though no such differences exist.

4.1 Policy implications

The model captures several sources of inefficiency arising from the unobservability of workplace benefits prior to joining a firm. First, firms' incentives are distorted toward providing inefficiently low benefits. Second, if worker types differ in their beliefs about unobservable benefits, this can lead to spurious differentiation in compensation packages and hence to excessive search and turnover. Under the assumption that each worker values benefits more than it costs each firm to provide them, the first-best efficient outcome would entail all firms providing high benefits and workers incurring zero search costs.

An immediate policy prescription is to mandate that firms provide at least some minimum level of workplace benefits, thus raising $\underline{b}$.[23] This simple solution, however, relies on the assumption of homogeneity in workers' preferences and firms' costs. In more realistic settings with heterogeneity along these dimensions, a uniform mandate is unlikely to yield a Pareto improvement.

A potentially more relevant policy instrument concerns the (firms') tax advantages to providing workplace benefits, denoted by τ. Recalling Proposition 1, increasing τ makes the first-best efficient outcome unambiguously easier to sustain in equilibrium, holding other parameters fixed. Similarly, in any equilibrium with differentiated offers, increasing τ reduces the profitability of deviations from high to low benefits. In the limit, as $\tau \to 1$, it becomes a weakly dominant strategy for the firms to offer high benefits. While the question of social costs and benefits of public subsidies to workplace pension plans remains outside the scope of this paper, these observations suggest that policymakers should consider targeting financial incentives at firms rather than workers, who are empirically less informed and less responsive to such incentives (Chetty et al., 2014; Fadlon et al., 2016). Unlike a minimum benefit mandate, the policy prescription of raising τ is more robust to introducing heterogeneity since it affects the firms' marginal incentives, rather than imposing a blanket constraint.

5 Conclusion

This paper analyzes the implications of workers' imperfect understanding of pecuniary benefits during job search, and the subsequent realization of benefits on the job, for the design of multidimensional compensation packages. Methodologically, it contributes to the literature in behavioral industrial organization by studying equilibrium effects of the agent observing the hidden attribute and possibly re-contracting.

[23] Such regulations are in place in the UK, Australia, and New Zealand, for example (OECD, 2023).

I show that unobservability of benefits prior to acceptance distorts firms' incentives toward providing inefficiently low benefits, despite the fact that lower benefits induce higher worker turnover (Corollary 1). Consequently, the efficient outcome where all firms provide high benefits is more difficult to sustain in equilibrium than an outcome where all firms provide inefficiently low benefits (Proposition 1). This result highlights a limitation of delegating the provision of benefits to firms when workers are imperfectly informed.

Furthermore, the coexistence of sophisticated and naïve workers generates additional equilibria with spurious differentiation in compensation packages (Proposition 2). Although qualitatively similar to the prediction of the classical compensating differentials theory (Rosen, 1974, 1986), the logic behind these equilibria and their efficiency properties are starkly different. Importantly, the model predicts that in such equilibria, the wage differential is bounded from above by the benefit differential (Corollary 2).

Taken together, the model's predictions offer a unified explanation for several puzzling patterns reported in the empirical literature on workplace benefits, namely inefficiently low benefit provision, the link between benefits and turnover, large between-firm dispersion in benefits, and the challenges in interpreting compensating differential estimates. This highlights the relevance of accounting for information frictions on the workers' side in the context of pecuniary workplace benefits. In terms of policy implications, the model indicates that targeting financial incentives at the (better informed) firms may effectively promote the provision of workplace benefits.

There are several promising directions for future research. Extending the model to incorporate heterogeneous worker productivity and screening or multi-worker firms could shed further light on how information frictions interact with classical economic rationales for the provision of workplace benefits.

References

ABELER, J., D. B. HUFFMANN, AND C. RAYMOND (2023): "Incentive Complexity, Bounded Rationality and Effort Provision," *IZA Discussion Paper No. 16284*.

AGNEW, J. R., L. R. SZYKMAN, S. P. UTKUS, AND J. A. YOUNG (2012): "Trust, Plan Knowledge and 401(k) Savings Behavior," *Journal of Pension Economics & Finance*, 11, 1–20.

ALBRECHT, J. W. AND B. AXELL (1984): "An Equilibrium Model of Search Unemployment," *Journal of Political Economy*, 92, 824–840.

AYRES, I. AND Q. CURTIS (2015): "Beyond Diversification: The Pervasive Problem of Excessive Fees and Dominated Funds in 401(k) Plans," *The Yale Law Journal*, 124, 1476.

BELOT, M., P. KIRCHER, AND P. MULLER (2022): "How Wage Announcements Affect Job Search – A Field Experiment," *American Economic Journal: Macroeconomics*, 14, 1–67.

BENNETT, N., T. C. BLUM, R. G. LONG, AND P. M. ROMAN (1993): "A Firm-Level Analysis of Employee Attrition," *Group & Organization Management*, 18, 482–499.

BESHEARS, J., J. J. CHOI, D. LAIBSON, AND B. C. MADRIAN (2018): "Behavioral Household Finance," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 1, 177–276.

BHARGAVA, S., G. LOEWENSTEIN, AND J. SYDNOR (2017): "Choose to Lose: Health Plan Choices from a Menu with Dominated Options," *The Quarterly Journal of Economics*, 132, 1319–1372.

BUBB, R. AND P. L. WARREN (2020): "An Equilibrium Theory of Retirement Plan Design," *American Economic Journal: Economic Policy*, 12, 22–45.

BURDETT, K. (1978): "A Theory of Employee Job Search and Quit Rates," *American Economic Review*, 68, 212–220.

BURDETT, K. AND K. L. JUDD (1983): "Equilibrium Price Dispersion," *Econometrica*, 955–969.

BURDETT, K. AND D. T. MORTENSEN (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 257–273.

CALIENDO, M., D. A. COBB-CLARK, AND A. UHLENDORFF (2015): "Locus of Control and Job Search Strategies," *Review of Economics and Statistics*, 97, 88–103.

CHETTY, R., J. N. FRIEDMAN, S. LETH-PETERSEN, T. H. NIELSEN, AND T. OLSEN (2014): "Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark," *The Quarterly Journal of Economics*, 129, 1141–1219.

CHOI, J. J., D. LAIBSON, AND B. C. MADRIAN (2010): "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds," *The Review of Financial Studies*, 23, 1405–1432.

——— (2011): "$100 Bills on the Sidewalk: Suboptimal Investment in 401(k) Plans," *Review of Economics and Statistics*, 93, 748–763.

COLE, A. AND B. TASKA (2023): "Worker Valuation of Retirement Benefits," *Working Paper*.

DELLAVIGNA, S. AND M. D. PASERMAN (2005): "Job Search and Impatience," *Journal of Labor Economics*, 23, 527–588.

DIAMOND, P. A. (1971): "A Model of Price Adjustment," *Journal of Economic Theory*, 3, 156–168.

DOHMEN, T. (2014): "Behavioral Labor Economics: Advances and Future Directions," *Labour Economics*, 30, 71–85.

DUFLO, E. AND E. SAEZ (2002): "Participation and Investment Decisions in a Retirement Plan: The Influence of Colleagues' Choices," *Journal of Public Economics*, 85, 121–148.

——— (2003): "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *The Quarterly Journal of Economics*, 118, 815–842.

ENGLMAIER, F., M. FAHN, U. GLOGOWSKY, AND M. A. SCHWARZ (2023a): "When Protection Becomes Exploitation: The Impactof Firing Costs on Present-Biased Employees," *Working Paper*.

ENGLMAIER, F., M. FAHN, AND M. A. SCHWARZ (2023b): "Long-Term Employment Relations When Agents are Present Biased," *Working Paper*.

ENKE, B. AND F. ZIMMERMANN (2019): "Correlation Neglect in Belief Formation," *The Review of Economic Studies*, 86, 313–332.

EYSTER, E. (2019): "Errors in Strategic Reasoning," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 2, 187–259.

EYSTER, E. AND M. RABIN (2005): "Cursed Equilibrium," *Econometrica*, 73, 1623–1672.

FADLON, I., J. LAIRD, AND T. H. NIELSEN (2016): "Do Employer Pension Contributions Reflect Employee Preferences? Evidence from a Retirement Savings Reform in Denmark," *American Economic Journal: Applied Economics*, 8, 196–216.

FAHN, M. AND R. SEIBEL (2022): "Present Bias in the Labor Market – When It Pays to Be Naive," *Games and Economic Behavior*, 135, 144–167.

FUDENBERG, D. AND J. TIROLE (1991): "Perfect Bayesian Equilibrium and Sequential Equilibrium," *Journal of Economic Theory*, 53, 236–260.

GABAIX, X. AND D. LAIBSON (2006): "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *The Quarterly Journal of Economics*, 121, 505–540.

GAMP, T. AND D. KRÄHMER (2022): "Competition in Search Markets with Naive Consumers," *The RAND Journal of Economics*, 53, 356–385.

——— (2023): "Biased Beliefs in Search Markets," *American Economic Journal: Microeconomics*, 15, 414–464.

GRUBER, J. AND R. MCKNIGHT (2016): "Controlling Health Care Costs Through Limited Network Insurance Plans: Evidence from Massachusetts State Employees," *American Economic Journal: Economic Policy*, 8, 219–250.

GUSTMAN, A. L. AND T. L. STEINMEIER (2005): "Imperfect Knowledge of Social Security and Pensions," *Industrial Relations: A Journal of Economy and Society*, 44, 373–397.

HEIDHUES, P., J. JOHNEN, AND B. KŐSZEGI (2021): "Browsing versus Studying: A Pro-market Case for Regulation," *The Review of Economic Studies*, 88, 708–729.

HEIDHUES, P., B. KŐSZEGI, AND T. MUROOKA (2023): "Procrastination Markets," *Working Paper*.

HEIDHUES, P. AND B. KŐSZEGI (2018): "Behavioral Industrial Organization," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 1, 517–612.

HEIDHUES, P., B. KŐSZEGI, AND T. MUROOKA (2016): "Exploitative Innovation," *American Economic Journal: Microeconomics*, 8, 1–23.

JÄGER, S., C. ROTH, N. ROUSSILLE, AND B. SCHOEFER (2024): "Worker Beliefs About Outside Options," *The Quarterly Journal of Economics*, qjae001.

JEHIEL, P. (2005): "Analogy-Based Expectation Equilibrium," *Journal of Economic Theory*, 123, 81–104.

JOHNEN, J. (2019): "Automatic-renewal contracts with heterogeneous consumer inertia," *Journal of Economics & Management Strategy*, 28, 765–786.

——— (2020): "Dynamic Dompetition in Deceptive Markets," *The RAND Journal of Economics*, 51, 375–401.

KARLE, H., H. SCHUMACHER, AND R. VØLUND (2023): "Consumer Loss Aversion and Scale-Dependent Psychological Switching Costs," *Games and Economic Behavior*, 138, 214–237.

KŐSZEGI, B. (2010): "Utility from Anticipation and Personal Equilibrium," *Economic Theory*, 44, 415–444.

KRISTAL, T., Y. COHEN, AND E. NAVOT (2020): "Workplace Compensation Practices and the Rise in Benefit Inequality," *American Sociological Review*, 85, 271–297.

LAMADON, T., M. MOGSTAD, AND B. SETZLER (2022): "Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market," *American Economic Review*, 112, 169–212.

LAVETTI, K. (2023): "Compensating Wage Differentials in Labor Markets: Empirical Challenges and Applications," *Journal of Economic Perspectives*, 37, 189–212.

LEE, C.-H., M.-L. HSU, AND N.-H. LIEN (2006): "The Impacts of Benefit Plans on Employee Turnover: A Firm-Level Analysis Approach on Taiwanese Manufacturing Industry," *The International Journal of Human Resource Management*, 17, 1951–1975.

LUSARDI, A. AND O. S. MITCHELL (2014): "The Economic Importance of Financial Literacy: Theory and Evidence," *Journal of Economic Literature*, 52, 5–44.

MCAFEE, R. P. AND M. SCHWARTZ (1994): "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity," *American Economic Review*, 84, 210–230.

MCCALL, J. J. (1970): "Economics of Information and Job Search," *The Quarterly Journal of Economics*, 113–126.

MITCHELL, O. S. (1988): "Worker Knowledge of Pension Provisions," *Journal of Labor Economics*, 6, 21–39.

MONTGOMERY, E. AND K. SHAW (1997): "Pensions and Wage Premia," *Economic Inquiry*, 35, 510–522.

MUROOKA, T. AND M. A. SCHWARZ (2019): "Consumer Exploitation and Notice Periods," *Economics Letters*, 174, 89–92.

NELSON, P. (1970): "Information and Consumer Behavior," *Journal of Political Economy*, 78, 311–329.

OECD (2020): *Financial Incentives for Funded Private Pension Plans 2020*, OECD Publishing, Paris.

——— (2023): *Pensions at a Glance 2023*, OECD Publishing, Paris.

OUIMET, P. AND G. A. TATE (2023): "Firms with Benefits? Nonwage Compensation and Implications for Firms and Labor Markets," *NBER Working Paper No. 31463*.

PETRONGOLO, B. AND C. A. PISSARIDES (2001): "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39, 390–431.

PIERCE, B. (2001): "Compensation Inequality," *The Quarterly Journal of Economics*, 116, 1493–1525.

REA, S. A. (1981): "Workmen's Compensation and Occupational Safety Under Imperfect Information," *American Economic Review*, 71, 80–93.

RIORDAN, M. H. (1986): "Monopolistic Competition with Experience Goods," *The Quarterly Journal of Economics*, 101, 265–279.

ROGERSON, R., R. SHIMER, AND R. WRIGHT (2005): "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, 43, 959–988.

ROSEN, S. (1974): "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82, 34–55.

——— (1986): "The Theory of Equalizing Differences," *Handbook of Labor Economics*, 1, 641–692.

SCHILLER, B. R. AND R. D. WEISS (1980): "Pensions and Wages: A Test for Equalizing Differences," *The Review of Economics and Statistics*, 529–538.

SCHUMACHER, H. (2024): "Competitive Markets, Add-On Prices, and Bounded Rationality," *Working Paper*.

SCHUMACHER, H. AND H. C. THYSEN (2022): "Equilibrium Contracts and Boundedly Rational Expectations," *Theoretical Economics*, 17, 371–414.

SHAPIRO, C. (1982): "Consumer Information, Product Quality, and Seller Reputation," *The Bell Journal of Economics*, 20–35.

——— (1983): "Premiums for High Quality Products as Returns to Reputations," *The Quarterly Journal of Economics*, 98, 659–679.

SOCKIN, J. (2024): "Show Me the Amenity: Are Higher-Paying Firms Better All Around?" *Working Paper*.

SORKIN, I. (2018): "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*, 133, 1331–1393.

SPIEGLER, R. (2016): "Bayesian Networks and Boundedly Rational Expectations," *The Quarterly Journal of Economics*, 131, 1243–1290.

SPINNEWIJN, J. (2015): "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs," *Journal of the European Economic Association*, 13, 130–167.

STAHL, D. O. (1989): "Oligopolistic Pricing with Sequential Consumer Search," *American Economic Review*, 79, 700–712.

STIGLITZ, J. E. (1979): "Equilibrium in Product Markets with Imperfect Information," *American Economic Review*, 69, 339–345.

VISCUSI, W. K. (1980): "Imperfect Job Risk Information and Optimal Workmen's Compensation Benefits," *Journal of Public Economics*, 14, 319–337.

WOLINSKY, A. (1986): "True Monopolistic Competition as a Result of Imperfect Information," *The Quarterly Journal of Economics*, 101, 493–511.

WU, L. (2024): "Partially Directed Search in the Labor Market," *Journal of Economic Theory*, 220, 105858.

A Proof of Lemma 1

Suppose that $-1 < d\,\mathbb{E}[b \mid w] \,/\, d\,w < 0$ for all w and assume for the sake of contradiction that $R_2^S \geq R_2^N$. Then, the optimality condition of a sophisticated worker implies:

$$\begin{aligned} c_2 &= \int_{w \geq R_2^S} \left(w + \mathbb{E}[b \mid w] - R_2^S - \mathbb{E}[b \mid R_2^S]\right)\phi(w)dw \\ &< \int_{w \geq R_2^S} (w - R_2^S)\phi(w)dw, \end{aligned}$$

where the inequality follows from $\mathbb{E}[b \mid w] - \mathbb{E}[b \mid R_2^S] < 0$ for all $w > R_2^S$.

In turn, the optimality condition of a naïve worker implies:

$$\begin{aligned} c_2 &= \int_{w \geq R_2^N} \left(w + \mathbb{E}[b] - R_2^N - \mathbb{E}[b]\right)\phi(w)dw \\ &= \int_{w \geq R_2^N} (w - R_2^N)\phi(w)dw \\ &\geq \int_{w \geq R_2^S} (w - R_2^S)\phi(w)dw, \end{aligned}$$

where the inequality follows from the assumption that $R_2^S \geq R_2^N$. Taken together, the two inequalities clearly lead to a contradiction. Thus, for $d\,\mathbb{E}[b \mid w] \,/\, d\,w < 0$, $R_2^S < R_2^N$ must obtain.

The proof for the case when $d\,\mathbb{E}[b \mid w] \,/\, d\,w > 0$ proceeds in the same way.

□

B Illustration of Lemma 2

To illustrate the indeterminacy result of Lemma 2, consider the following parametric example. Suppose that wage offers are distributed uniformly over the interval $[0, \bar{w}]$ for some $\bar{w} > 0$. Workplace benefits can take one of two values, either $\underline{b} = 0$ or $\bar{b} > 0$ for some $\bar{b} < \bar{w}$. Further, suppose for concreteness that the expected benefit level is decreasing in the wage offer. Specifically, let $\mathbb{P}[b = \bar{b} \mid w] = (1 - \frac{w}{\bar{w}})$. Then, $\mathbb{P}[b = \bar{b} \mid w]$ is strictly decreasing in w, with $\mathbb{P}[b = \bar{b} \mid 0] = 1$ and $\mathbb{P}[b = \bar{b} \mid \bar{w}] = 0$. Consequently, $\mathbb{E}[b \mid w] = (1 - \frac{w}{\bar{w}}) \times \bar{b}$ is also strictly decreasing in w.[1] From the perspective of a naïve worker, who is subject to correlation neglect but takes into account the correct marginal distribution of benefits, $\mathbb{P}[b = \bar{b}] = \mathbb{E}_w[\mathbb{P}[b = \bar{b} \mid w]] = 0.5$ and $\mathbb{E}[b] = 0.5 \times \bar{b}$.

[1] Imposing $\bar{w} > \bar{b}$ guarantees that $-1 < d\,\mathbb{E}[b \mid w] \,/\, d\,w < 0$, as assumed in the main text.

Solving for the perceived-optimal reservation wages in period 2, R_2^S solves:

$$
\begin{aligned}
c_2 &= \int_{w=R_2^S}^{\bar{w}} \left(w + \mathbb{E}[b \mid w] - R_2^S - \mathbb{E}[b \mid R_2^S]\right) \frac{1}{\bar{w}} dw \\
&= \int_{w=R_2^S}^{\bar{w}} \left(w + (1 - \frac{w}{\bar{w}})\bar{b} - R_2^S - (1 - \frac{R_2^S}{\bar{w}})\bar{b}\right) \frac{1}{\bar{w}} dw \\
&= \int_{w=R_2^S}^{\bar{w}} \left((1 - \frac{\bar{b}}{\bar{w}})(w - R_2^S)\right) \frac{1}{\bar{w}} dw \\
&= \frac{(\bar{w} - \bar{b})(\bar{w} - R_2^S)^2}{2\bar{w}^2},
\end{aligned}
$$

which yields $R_2^S = \bar{w}(1 - \sqrt{2c_2/(\bar{w} - \bar{b})})$. In turn, R_2^N solves:

$$
\begin{aligned}
c_2 &= \int_{w=R_2^N}^{\bar{w}} \left(w + \mathbb{E}[b] - R_2^N - \mathbb{E}[b]\right) \frac{1}{\bar{w}} dw \\
&= \int_{w=R_2^N}^{\bar{w}} \left(w - R_2^N\right) \frac{1}{\bar{w}} dw \\
&= \frac{(\bar{w} - R_2^N)^2}{2\bar{w}},
\end{aligned}
$$

which yields $R_2^N = \bar{w}(1 - \sqrt{2c_2/\bar{w}})$.

Observe that $R_2^N > R_2^S$ for any $\bar{w} > \bar{b} > 0$ and $c_2 > 0$. This is an implication of Lemma 1: when $\mathbb{E}[b \mid w]$ is strictly decreasing in w, sophisticated workers have a weaker incentive to search for high-wage offers *in a one-shot problem*. Consequently, they adopt a strictly lower reservation wage in period 2, relative to naïve workers who ignore the correlation between the wage and expected benefits.

Turning to search behavior in period 1, the sophisticated worker's valuation of sampling wage offer w is:

$$
u_1^S(w) = \left(w + (1 - \tfrac{w}{\bar{w}})\bar{b}\right) \ + \ (1 - \tfrac{w}{\bar{w}}) \cdot \left(\max\{w + \bar{b};\, v_2^S - c_2\}\right) + \ \tfrac{w}{\bar{w}} \cdot \left(\max\{w;\, v_2^S - c_2\}\right),
$$

where $v_2^S - c_2 = R_2^S + \mathbb{E}[b \mid R_2^S]$. The optimal cutoff R_1^S solves:

$$
c_1 = \textstyle\int_{w=R_1^S}^{\bar{w}} \left(u_1^S(w) - u_1^S(R_1^S)\right) \frac{1}{\bar{w}} dw.
$$

The above does not have a closed-form solution but can be solved numerically for concrete values of $\bar{w}, \bar{b}, c_1$, and c_2. Analogously, the naïve worker's valuation is given by:

$$
u_1^N(w) = \left(w + 0.5\, \bar{b}\right) \ + \ 0.5 \cdot \left(\max\{w + \bar{b};\, v_2^N - c_2\}\right) + \ 0.5 \cdot \left(\max\{w;\, v_2^N - c_2\}\right),
$$

where $v_2^N - c_2 = R_2^N + 0.5\,\bar{b}$. The optimal cutoff R_1^N solves:

$$c_1 = \int_{w=R_1^N}^{\bar{w}} \left(u_1^N(w) - u_1^N(R_1^N)\right) \tfrac{1}{\bar{w}} dw.$$

To demonstrate that the ordering between R_1^S and R_1^N depends on a specific parametrization, and thus is not pinned down by the sign of $d\,\mathbb{E}[b \mid w] \,/\, d\,w$, consider the following examples:

(a) $\bar{w} = 84.98697$, $\bar{b} = 10.85944$, $c_1 = 0.25898$, $c_2 = 0.31666$
Under this parametrization, $R_2^S = 77.13146$ and $R_2^N = 77.65049$ ($R_2^S < R_2^N$), but $R_1^S = 79.96359$ and $R_1^N = 79.68276$ ($R_1^S > R_1^N$).

(b) $\bar{w} = 65.45503$, $\bar{b} = 2.10638$, $c_1 = 11.76293$, $c_2 = 15.88642$
Under this parametrization, $R_2^S = 19.09941$ and $R_2^N = 19.85139$ ($R_2^S < R_2^N$), and $R_1^S = 37.24964$ and $R_1^N = 37.70718$ ($R_1^S < R_1^N$).

This confirms the claim of Lemma 2.

□

C Proof of Corollary 1

A worker accepts a wage offer $\hat{w}$ if:

$$u_1^j(\hat{w}) \;\geq\; v_1^j - c_1.$$

Given $\hat{w} + \bar{b} < v_2^j - c_2$,

$$u_1^j(\hat{w}) \;=\; (\hat{w} + \mathbb{E}[b \mid \hat{w}]) + (v_2^j - c_2) \;<\; 2(v_2^j - c_2).$$

Taken together, this would imply:

$$2(v_2^j - c_2) \;>\; v_1^j - c_1 \quad\Longleftrightarrow\quad c_1 - c_2 \;>\; (v_1^j + c_2) - 2v_2^j.$$

This is a contradiction, since $c_1 - c_2 \leq 0$ and $(v_1^j + c_2) - 2v_2^j \geq 0$. The latter holds because $(v_1^j + c_2) = 2v_2^j$ would require $c_1 = c_2$ and that the agent must unavoidably exit employment after one period (at a cost of c_2). In that case, the two value functions represent the expected utility from participating in a spot labor market twice. Otherwise, $c_1 < c_2$ and/or the option value of staying in the same employment in period 2 upon discovering b imply $(v_1^j + c_2) \geq 2v_2^j$. Thus, a strategy of accepting an offer in period 1 that is surely going to induce turnover in period 2 can never be optimal for either worker type.

□

D Proof of Lemma 3

First, consider the case when $b^* = \underline{b}$ and suppose that $w^* = -\underline{b} + \Delta$ for some $\Delta > 0$. Then, there exists a profitable deviation whereby a firm lowers its wage by some $0 < \epsilon \leq \min\{c_1/2; \Delta\}$ and still hires and retains any worker who encounters it, because:

$$2(w^* - \epsilon + \underline{b}) = 2(\Delta - \epsilon) \geq \max\{2\underbrace{(w^* + \underline{b})}_{=\Delta} - c_1; 0\}.$$

Since such a profitable deviation exists for any $\Delta > 0$, $w^* = -\underline{b}$ must obtain in equilibrium.

Second, consider the case when $b^* = \bar{b}$ and suppose that $w^* = -\bar{b} + \Delta$ for some $\Delta > 0$. Is a marginal wage reduction necessarily profitable? Similarly to the above, a firm deviating to $w^* - \epsilon$, where $0 < \epsilon \leq \min\{c_1/2; \Delta\}$, can hire and retain a naïve worker more profitably. However, a sophisticated worker with pessimistic beliefs accepts the deviator's offer and remains in employment for two periods only if:

$$\begin{aligned} 2(w^* - \epsilon + \underline{b}) \;&\geq\; \max\{2(w^* + \bar{b}) - c_1; 0\} \iff \\ \epsilon \;&\leq\; \min\{c_1/2 - (\bar{b} - \underline{b}); \Delta - (\bar{b} - \underline{b})\}. \end{aligned}$$

Thus, there is room for profitable deviation, and $w^* = -\bar{b} + \Delta$ cannot arise in equilibrium, when both $c_1/2 > (\bar{b} - \underline{b})$ and $\Delta > (\bar{b} - \underline{b})$ hold.

If either one of these conditions fails, an infinitesimal wage reduction results in a sophisticated worker rejecting the offer and the deviating firm may only attract naïve workers. Such a deviation is nonetheless profitable when:

$$\begin{aligned} \lambda 2(y - w^* + \bar{\epsilon} - (1-\tau)\bar{b}) \;>\; 2(y - w^* - (1-\tau)\bar{b}) \;&\iff\; \lambda(y - \Delta + \bar{\epsilon} + \tau\bar{b}) \;>\; (y - \Delta + \tau\bar{b}) \\ &\iff\; \Delta \;>\; y + \tau\bar{b} - \tfrac{\lambda}{1-\lambda}\bar{\epsilon}, \end{aligned}$$

where $\bar{\epsilon} = \min\{c_1/2; \Delta\}$ captures the most profitable feasible deviation.

Otherwise, no profitable deviation exists and one can construct an equilibrium with uniformly high benefits and a wage strictly exceeding the monopsony wage by some small enough margin $0 \leq \Delta \leq \bar{\Delta}$, where:

$$\bar{\Delta} \;=\; \max\{y + \tau\bar{b} - \tfrac{\lambda}{1-\lambda} \times \tfrac{c_1}{2}; (1-\lambda)(y + \tau\bar{b})\}.$$

$\square$

E Remaining Derivations Underlying Proposition 2

Case 1: $w_H + (1-\tau)\underline{b} > w_L + (1-\tau)\bar{b}$

Equilibrium with turnover on the equilibrium path. To demonstrate that a compensating equilibrium may feature turnover of naïve workers on the equilibrium path, here I construct an equilibrium in which naïve workers leave jobs offering $(w_H, \underline{b})$.

A sophisticated worker accepts the first offer sampled in period 1 if:

$$u_1^S(w_H) - u_1^S(w_L) \;=\; 2((w_H - w_L) - (\bar{b} - \underline{b})) \;\leq\; c_1/p$$

and:

$$u_1^S(w_L) - u_1^S(w_H) \;=\; 2((\bar{b} - \underline{b}) - (w_H - w_L)) \;\leq\; c_1/(1-p).$$

The above also imply that a sophisticated worker accepts the first offer sampled in period 2 and consequently $(v_2^N - c_2) \geq (v_2^S - c_2)$.

For $(w_L, \bar{b})$ to be offered in equilibrium, sophisticated workers should leave upon discovering $(w_L, \underline{b})$:

$$w_L + \underline{b} \;<\; p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2 \quad\Longleftrightarrow\quad c_2 \;<\; p(w_H - w_L) + (1-p)(\bar{b} - \underline{b}).$$

They nevertheless remain in employment offering $(w_H, \underline{b})$ when:

$$w_H + \underline{b} \;\geq\; p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2 \quad\Longleftrightarrow$$
$$c_2 \;\geq\; (1-p)((\bar{b} - \underline{b}) - (w_H - w_L)).$$

Turning to naïve workers, they search for w_H in period 1 as long as:

$$u_1^N(w_H) - u_1^N(w_L) \;=\; (w_H - w_L) + p\big(\max\{(w_H + \underline{b})\,;\,(v_2^N - c_2)\} - (v_2^N - c_2)\big) + (1 - p)\big((w_H + \bar{b}) - \max\{(w_L + \bar{b})\,;\,(v_2^N - c_2)\}\big) \;>\; c_1/p.$$

Suppose that when searching in period 2, naïve workers also search for w_H, which requires:

$$(w_H - w_L) \;>\; c_2/p.$$

Then, $(v_2^N - c_2) \;=\; w_H + p\underline{b} + (1-p)\bar{b} - c_2/p$ and a naïve worker leaves upon discovering $(w_H, \underline{b})$ when:

$$w_H + \underline{b} \;<\; (v_2^N - c_2) \quad\Longleftrightarrow\quad c_2/p \;<\; (1-p)(\bar{b} - \underline{b}).$$

A naïve worker also expects to leave $(w_L, \bar{b})$ if:

$$w_L + \bar{b} \;<\; (v_2^N - c_2) \quad\Longleftrightarrow\quad c_2/p \;<\; (w_H - w_L) - p(\bar{b} - \underline{b}).$$

Then, naïve workers indeed search for w_H in period 1 as long as:

$$\begin{aligned} u_1^N(w_H) - u_1^N(w_L) \;=\; &(w_H - w_L) + p\big((v_2^N - c_2) - (v_2^N - c_2)\big) + (1-p)\big((w_H + \bar{b}) - w_H - \\ &p\underline{b} - (1-p)\bar{b} + c_2/p\big) \;>\; c_1/p \;\Longleftrightarrow \\ c_1/p \;<\; &(w_H - w_L) + (1-p)p(\bar{b} - \underline{b}) + (1-p)c_2/p. \end{aligned}$$

Next, we take into account the workers' ex-post PC's. Those of a sophisticated worker require:

$$\begin{aligned} u_1^S(w_L) &= 2(w_L + \bar{b}) \;\geq\; 0 \\ u_1^S(w_H) &= 2(w_H + \underline{b}) \;\geq\; 0. \end{aligned}$$

The ex-post PC of a naïve worker is:

$$\begin{gathered} u_1^N(w_H) - c_1/p \;= \\ w_H + p(\underline{b} + w_H + p\underline{b} + (1-p)\bar{b} - c_2/p) + (1-p)(\bar{b} + w_H + \bar{b}) - c_1/p \;\geq\; 0. \end{gathered}$$

It can be shown that when the PC of a sophisticated worker is binding, yielding $w_H = -\underline{b}$, then that of a naïve worker must be slack. Therefore, the PC's of a sophisticated worker again determine both w_L and w_H, implying $w_L = -\bar{b}$ and $w_H = -\underline{b}$.[2]

As for the firms, the equal profits condition is unaffected by the churning of naïve workers across high-wage firms:

$$\begin{gathered} \pi(w_H) = 2((1-\lambda) + \tfrac{\lambda}{p})(y - w_H - (1-\tau)\underline{b}) \;=\; 2(1-\lambda)(y - w_L - (1-\tau)\bar{b}) = \pi(w_L) \quad\Longleftrightarrow \\ (1-\lambda)\;/\;((1-\lambda) + \tfrac{\lambda}{p}) \;=\; \tfrac{(y - w_H - (1-\tau)\underline{b})}{(y - w_L - (1-\tau)\bar{b})}. \end{gathered}$$

Because of that, a deviation from $(w_H, \underline{b})$ to $(w_H, \bar{b})$ is strictly unprofitable. In turn, a deviation from $(w_L, \bar{b})$ to $(w_L, \underline{b})$ is unprofitable as long as:

[2] Plugging in $w_H = -\underline{b}$ yields:

$$u_1^N(w_H) - c_1/p \;=\; (2+p)(1-p)(\bar{b} - \underline{b}) - c_2 - c_1/p.$$

But:

$$c_2 + c_1/p \;\leq\; c_2(1+p)/p \;<\; (1+p)(1-p)(\bar{b} - \underline{b}) \;<\; (2+p)(1-p)(\bar{b} - \underline{b})$$

by $c_2/p < (1-p)(\bar{b} - \underline{b})$. Thus, $u_1^N(w_H) - c_1/p > 0$.

$$2(1-\lambda)(y-w_L-(1-\tau)\bar{b}) \;\geq\; (1-\lambda)(y-w_L-(1-\tau)\underline{b}) \iff$$
$$2 \;\geq\; \frac{(y-w_L-(1-\tau)\underline{b})}{(y-w_L-(1-\tau)\bar{b})}.$$

Consider a deviation to some intermediate wage offer $w \in (w_L, w_H)$ combined with low benefits, which can attract only naïve workers. Let

$$\tilde{w} \;=\; \max\{w_H - p(\bar{b}-\underline{b}) - c_2/p;\; w_H - c_2/p;\; w_H - c_1/((2-p)p)\} \tag{A1}$$

be the lowest wage offer that is accepted by naïve workers in both periods 1 and 2. The first condition inside the curly brackets ensures that a naïve worker anticipates to stay in a job offering $(\tilde{w}, \bar{b})$ rather than search again in period 2. The second condition ensures that $\tilde{w}$ is accepted by a naïve worker searching in period 2, and the third that $\tilde{w}$ is accepted by a naïve worker searching in period 1. Then, taking into account the churning of naïve workers on the equilibrium path, the deviating firm is not better off as long as:

$$2\tfrac{\lambda}{p}(y-\tilde{w}-(1-\tau)\underline{b}) \;\leq\; 2((1-\lambda)+\tfrac{\lambda}{p})(y-w_H-(1-\tau)\underline{b}).$$

Finally, a deviation to some intermediate wage combined with high benefits also attracts only naïve workers and is therefore dominated by $(\tilde{w}, \underline{b})$.

In sum, such equilibrium with turnover of naïve workers exists as long as all of the following hold:[3]

(i) $c_1/p \;\geq\; 2((w_H-w_L)-(\bar{b}-\underline{b}))$,

(ii) $c_1/(1-p) \;\geq\; 2((\bar{b}-\underline{b})-(w_H-w_L))$,

(iii) $c_1/p \;<\; (w_H-w_L)+(1-p)p(\bar{b}-\underline{b})+(1-p)c_2/p$,

(iv) $c_2 \;\geq\; (1-p)((\bar{b}-\underline{b})-(w_H-w_L))$,

(v) $c_2/p \;<\; (1-p)(\bar{b}-\underline{b})$,

(vi) $c_2/p \;<\; (w_H-w_L)-p(\bar{b}-\underline{b})$,

(vii) $(1-\lambda)\,/\,((1-\lambda)+\frac{\lambda}{p}) \;=\; \frac{(y-w_H-(1-\tau)\underline{b})}{(y-w_L-(1-\tau)\bar{b})}$,

(viii) $2 \;\geq\; \frac{(y-w_L-(1-\tau)\underline{b})}{(y-w_L-(1-\tau)\bar{b})}$,

(ix) $w_H+\underline{b} \;=\; 0$,

(x) $w_L+\bar{b} \;=\; 0$,

[3]Note that $c_2/p < (w_H-w_L)-p(\bar{b}-\underline{b})$ implies both $c_2/p < (w_H-w_L)$ and $c_2 < p(w_H-w_L)+(1-p)(\bar{b}-\underline{b})$.

(xi) $\frac{\lambda}{p}(y - \tilde{w} - (1-\tau)\underline{b}) \;\leq\; ((1-\lambda) + \frac{\lambda}{p})(y - w_H - (1-\tau)\underline{b})$, $\tilde{w}$ given by (A1).

Plugging in $w_H = -\underline{b}$ and $w_L = -\bar{b}$, the above boils down to:

(iii') $c_1/p \;<\; (1 + p - p^2)(\bar{b} - \underline{b}) + (1-p)c_2/p$,

(v') $c_2/p \;<\; (1-p)(\bar{b} - \underline{b})$,

(vii') $(1-\lambda) \,/\, ((1-\lambda) + \frac{\lambda}{p}) \;=\; \frac{(y+\tau\underline{b})}{(y+\tau\bar{b})}$,

(viii') $2 \;\geq\; \frac{(y+(\bar{b}-\underline{b})+\tau\underline{b})}{(y+\tau\bar{b})}$,

(xi') $\frac{\lambda}{p}(y - \tilde{w} - (1-\tau)\underline{b}) \;\leq\; ((1-\lambda) + \frac{\lambda}{p})(y + \tau\underline{b})$, $\tilde{w}$ given by (A1).

Indeed, these conditions can hold simultaneously.[4]

Case 2: $w_H + (1-\tau)\underline{b} \;\leq\; w_L + (1-\tau)\bar{b}$

Equilibrium with turnover on the equilibrium path. Consider a modification of the equilibrium characterized in the main text that features positive turnover of naïve workers on the equilibrium path. Here, I construct an equilibrium in which sophisticated workers search specifically for $(w_L, \bar{b})$ in period 1 and remain in the initially accepted employment, while naïve workers search specifically for w_H in period 1 and subsequently leave upon discovering low benefits.

Sophisticated workers search specifically for $(w_L, \bar{b})$ if:

$$u_1^S(w_L) - u_1^S(w_H) \;=\; 2((w_L + \bar{b}) - (w_H + \underline{b})) \;>\; c_1/(1-p).$$

Suppose that when searching in period 2, a sophisticated worker would accept the first sampled offer, which requires:

$$(w_L + \bar{b}) - (w_H + \underline{b}) \;\leq\; c_2/(1-p).$$

Then, $(v_2^S - c_2) = p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2$. A sophisticated worker has no incentive to leave a low-wage, high-benefit job which provides the highest overall utility. In case of a deviation to low benefits, she would leave a job providing $(w_L, \underline{b})$ if:

$$w_L + \underline{b} \;<\; p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2 \;\iff\; c_2 \;<\; p(w_H - w_L) + (1-p)(\bar{b} - \underline{b}),$$

but stays with $(w_H, \underline{b})$ if:

[4] For example for $\lambda = 0.227$, $\bar{b} = 116.157$, $\underline{b} = 24.083$, $c_1 = 9.525$, $c_2 = 17.071$, $\tau = 0.391$, and $y = 69.905$, which solve for $p = 0.648$, $w_H = -24.083$, $w_L = -116.157$, and $\tilde{w} = -34.951$.

$$w_H + \underline{b} \ \geq \ p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2 \iff c_2 \ \geq \ (1-p)((w_L + \bar{b}) - (w_H + \underline{b})),$$

which coincides with the condition for accepting the first offer sampled in period 2.

In contrast to the case considered in the main text, consider parametrizations such that a naïve worker searches specifically for w_H also in period 2, which requires:

$$c_2/p \ < \ (w_H - w_L).$$

This implies $(v_2^N - c_2) > (v_2^S - c_2)$, which is necessary for supporting a decision rule under which a naïve worker leaves $(w_H, \underline{b})$ while a sophisticated worker would not.

Then, $(v_2^N - c_2) = w_H + p\underline{b} + (1-p)\bar{b} - c_2/p$ and a naïve worker indeed leaves $(w_H, \underline{b})$ if:

$$w_H + \underline{b} \ < \ w_H + p\underline{b} + (1-p)\bar{b} - c_2/p \iff c_2/p \ < \ (1-p)(\bar{b} - \underline{b}).$$

This implies that a naïve worker would also leave $(w_L, \underline{b})$. Suppose that a naïve worker would nevertheless stay in a job offering $(w_L, \bar{b})$:

$$w_L + \bar{b} \ \geq \ w_H + p\underline{b} + (1-p)\bar{b} - c_2/p \iff c_2/p \ \geq \ (w_H - w_L) - p(\bar{b} - \underline{b}).$$

Then:

$$\begin{aligned} u_1^N(w_H) &= w_H + p(\underline{b} + (v_2^N - c_2)) + (1-p)(\bar{b} + w_H + \bar{b}), \\ u_1^N(w_L) &= w_L + p(\underline{b} + (v_2^N - c_2)) + (1-p)(\bar{b} + w_L + \bar{b}) \end{aligned}$$

and a naïve worker indeed searches specifically for w_H in period 1 as long as:

$$u_1^N(w_H) - u_1^N(w_L) \ = \ (2-p)(w_H - w_L) \ > \ c_1/p.$$

Next, the binding ex-post PC of a sophisticated worker pins down w_L:

$$u_1^S(w_L) - c_1/(1-p) \ = \ 2(w_L + \bar{b}) - c_1/(1-p) \ = \ 0,$$

while that of a naïve worker associated with sampling a high-wage offer immediately pins down w_H:

$$u_1^N(w_H) \ = \ w_H + p(\underline{b} + w_H + p\underline{b} + (1-p)\bar{b} - c_2/p) + (1-p)(\bar{b} + w_H + \bar{b}) \ = \ 0.$$

As for the firms, the equal profits condition, adjusted for partial participation of naïve workers in the labor market and their churning between high-wage firms, is:

$$\pi(w_H) \;=\; 2\,\tfrac{\lambda\times p}{p}\,(y - w_H - (1-\tau)\underline{b}) \;=\; 2\,\tfrac{(1-\lambda)}{(1-p)}\,(y - w_L - (1-\tau)\bar{b}) \;=\; \pi(w_L) \;\iff$$
$$\tfrac{\lambda\times(1-p)}{(1-\lambda)} \;=\; \tfrac{(y-w_L-(1-\tau)\bar{b})}{(y-w_H-(1-\tau)\underline{b})}.$$

A deviation from $(w_H, \underline{b})$ to $(w_H, \bar{b})$ is never profitable. A deviation from $(w_L, \bar{b})$ to $(w_L, \underline{b})$ is not profitable as long as:

$$2\,\tfrac{(1-\lambda)}{(1-p)}\,(y - w_L - (1-\tau)\bar{b}) \;\geq\; \tfrac{(1-\lambda)}{(1-p)}\,(y - w_L - (1-\tau)\underline{b}).$$

Finally, any deviation to $w \in (w_L, w_H)$ would attract no workers: sophisticated workers would continue searching for $(w_L, \bar{b})$ and naïve workers would reject in favor of the outside option.

Taken together, for such an equilibrium to exist, all of the following must hold:[5]

(i) $c_1/(1-p) \;<\; 2((w_L + \bar{b}) - (w_H + \underline{b}))$,

(ii) $c_2/(1-p) \;\geq\; (w_L + \bar{b}) - (w_H + \underline{b})$,

(iii) $c_2/p \;<\; (w_H - w_L)$,

(iv) $c_2/p \;<\; (1-p)(\bar{b} - \underline{b})$,

(v) $c_2/p \;\geq\; (w_H - w_L) - p(\bar{b} - \underline{b})$,

(vi) $c_1/p \;<\; (2-p)(w_H - w_L)$,

(vii) $\tfrac{\lambda\times(1-p)}{(1-\lambda)} \;=\; \tfrac{(y-w_L-(1-\tau)\bar{b})}{(y-w_H-(1-\tau)\underline{b})}$,

(viii) $2(y - w_L - (1-\tau)\bar{b}) \;\geq\; (y - w_L - (1-\tau)\underline{b})$,

(ix) $w_L \;=\; -\bar{b} + c_1/(2(1-p))$,

(x) $w_H \;=\; -(2-p-p^2)/2 \times \bar{b} - p(1+p)/2 \times \underline{b} + c_2/2$.

Indeed, these conditions can hold simultaneously.[6]

[5] Note that $c_2/p < (w_H - w_L)$ implies $c_2 < p(w_H - w_L) + (1-p)(\bar{b} - \underline{b})$.

[6] For example for $\lambda = 0.774$, $\bar{b} = 57.071$, $\underline{b} = 36.487$, $c_1 = 2.848$, $c_2 = 3.688$, $\tau = 0.221$, and $y = 181.155$, which solve for $p = 0.718$, $w_H = -42.527$, and $w_L = -52.017$.

F Proof of Corollary 2

Consider a putative compensating equilibrium in which the wage differential strictly exceeds the benefit differential, i.e. $(w_H - w_L) > (\bar{b} - \underline{b})$. This implies that the high-wage, low-benefit package is strictly more costly for the firms to provide:

$$(w_H - w_L) > (\bar{b} - \underline{b}) \implies w_H + (1-\tau)\underline{b} > w_L + (1-\tau)\bar{b}.$$

Thus, in order for the equal profits condition to hold, firms offering $(w_H, \underline{b})$ must attract a strictly larger mass of workers, as a compensation package with low benefits can never induce strictly lower turnover than a compensation package with high benefits.

Which worker type should search specifically for w_H? It cannot be the case that sophisticated workers accept only high-wage offers in period 1. If that was the case, so would all naïve workers because in any pure-strategy compensating equilibrium:

$$u_1^N(w_H) - u_1^N(w_L) \;>\; u_1^S(w_H) - u_1^S(w_L).$$

Therefore, the only way in which a compensation package with high wage and low benefits can attract a strictly larger mass of workers is for sophisticated workers to accept the first sampled offer in period 1 while naïve workers search specifically for a high-wage offer.[7]

The relevant participation constraints in this case are:

(i) $u_1^S(w_L) \;=\; 2(w_L + \bar{b}) \;\geq\; 0,$

(ii) $u_1^S(w_H) \;=\; 2(w_H + \underline{b}) \;\geq\; 0,$

(iii) $u_1^N(w_H) - c_1/p \;=\; w_H + p(\underline{b} + \max\{w_H + \underline{b};\, (v_2^N - c_2)\}) + (1-p)(\bar{b} + w_H + \bar{b}) - c_1/p \;\geq\; 0.$

Constraint (i) is always binding and implies $w_L = -\bar{b}$. Now, which of the other two constraints (ii) and (iii) should determine w_H? If (ii) was binding, we would have $(w_H - w_L) = (\bar{b} - \underline{b})$. Thus, for $(w_H - w_L) > (\bar{b} - \underline{b})$ to hold, it is necessary that (iii) is binding while (ii) is slack.

Formalizing an intuitive argument presented already in the main text, consider w_H such that (iii) is binding and (ii) is slack. Then, there always exists a profitable deviation to $(w_H - \epsilon, \underline{b})$ for some $\epsilon > 0$, such that workers' acceptance decisions in period 1 and their behavior in period 2 remain unchanged relative to $(w_H, \underline{b})$. First, since $w_H + \underline{b} > 0$, sophisticated workers with pessimistic beliefs accept the offer and remain in employment for two periods for small enough $\epsilon > 0$.

[7] If instead sophisticated workers were to search specifically for w_L in period 1, then $(w_H - w_L) < (\bar{b} - \underline{b})$ would obtain in equilibrium, see the derivation of case 2 in the main text.

Second, naïve workers also accept in period 1 as long as $u_1^N(w_H) - u_1^N(w_H - \epsilon) = 2\epsilon \leq c_1/p$. Their behavior in period 2 also coincides with that following the acceptance of $(w_H, \underline{b})$, except in the knife-edge case when $w_H + \underline{b} = (v_2^N - c_2)$. Recall that $(v_2^N - c_2) = \max\{p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2;\, w_H + p\underline{b} + (1-p)\bar{b} - c_2/p\}$. In the first case, the knife-edge cannot materialize as $w_H + \underline{b} > p(w_H + \underline{b}) + (1-p)(w_L + \bar{b}) - c_2$ for $w_H + \underline{b} > w_L + \bar{b}$. In the second case, the knife-edge would require that:

$$w_H + \underline{b} \;=\; w_H + p\underline{b} + (1-p)\bar{b} - c_2/p \iff c_2/p \;=\; (1-p)(\bar{b} - \underline{b}).$$

But, on the other hand, (ii) being slack while (iii) is binding would require:

$$-\underline{b} \;<\; c_1/(2p) - p\underline{b} - (1-p)\bar{b} \iff c_1/(2p) \;>\; (1-p)(\bar{b} - \underline{b})$$

and the two conditions contradict $c_2 \geq c_1$. Thus, a putative compensating equilibrium in which $(w_H - w_L) > (\bar{b} - \underline{b})$ would necessarily leave room for a profitable deviation for the firms offering $(w_H, \underline{b})$. Consequently, in any compensating equilibrium $(w_H - w_L) \leq (\bar{b} - \underline{b})$.

□

G Augmenting Equilibria

Consider the case of an *augmenting equilibrium* in which $b(w_H) = \bar{b}$ and $b(w_L) = \underline{b}$. That is, jobs that pay high wages also provide high benefits, and vice versa. Suppose that fraction $p \in (0,1)$ of firms offer the high-wage, high-benefit package, while fraction $(1-p)$ offer the low-wage, low-benefit package.

Sophisticated workers correctly infer what benefits are associated with a given wage offer, while naïve workers mistakenly believe that any wage offer comes with high benefits with probability p and low benefits with probability $(1-p)$. Consequently, the two types have the following option value of searching in period 2:

$$\begin{aligned}(v_2^S - c_2) \;&=\; \max\{p(w_H + \bar{b}) + (1-p)(w_L + \underline{b}) - c_2\,;\, (w_H + \bar{b}) - c_2/p\},\\ (v_2^N - c_2) \;&=\; \max\{p(w_H + \bar{b}) + (1-p)(w_L + \underline{b}) - c_2\,;\, w_H + p\bar{b} + (1-p)\underline{b} - c_2/p\},\end{aligned}$$

depending on whether they accept the first sampled offer or keep searching until an offer with w_H is found. Notice that $(v_2^S - c_2) \geq (v_2^N - c_2)$, which implies that sophisticated workers are more inclined to search again in period 2 and, if they do, to search for a high-wage offer. Intuitively, the fact that high wages are paired with high benefits increases the sophisticated worker's valuation of sampling w_H, relative to her naïve counterpart (Lemma 1).

Further, since $c_2 \geq c_1$, sophisticated workers never accept an offer in period 1 with the intention to search again in period 2, but naïve workers might anticipate searching again for low realizations of b and still accept. Clearly, neither type has an incentive to search further if the accepted offer turns out to provide $(w_H, \bar{b})$. Therefore:

$$u_1^S(w_H) = 2(w_H + \bar{b}), \qquad u_1^S(w_L) = 2(w_L + \underline{b}),$$
$$u_1^N(w_H) = (w_H + p\bar{b} + (1-p)\underline{b}) + p(w_H + \bar{b}) + (1-p)\max\{(w_H + \underline{b})\,;\,(v_2^N - c_2)\},$$
$$u_1^N(w_L) =$$
$$(w_L + p\bar{b} + (1-p)\underline{b}) + p\max\{(w_L + \bar{b})\,;\,(v_2^N - c_2)\} + (1-p)\max\{(w_L + \underline{b})\,;\,(v_2^N - c_2)\},$$

which implies $u_1^S(w_H) - u_1^S(w_L) > u_1^N(w_H) - u_1^N(w_L)$. Thus, if the two types search differently in period 1, it must be the case that sophisticated workers are searching specifically for $(w_H, \bar{b})$, while naïve workers accept the first sampled offer.

From the firms' perspective, the compensation package with high wage and high benefits is strictly more costly to provide. Thus, for the equal profits condition to hold, the firms offering $(w_H, \bar{b})$ must either attract a larger mass of workers, retain a higher proportion of hired workers, or both. It can be shown that the first two cases necessarily result in a contradiction:

Lemma A1: *There do not exist augmenting equilibria in which the differentiated offers induce the same turnover rate or attract the same mass of workers.*

Proof: First, consider a putative augmenting equilibrium in which firms offering the high-wage, high-benefit package attract the same mass of workers as firms offering the low-wage, low-benefit package, but only high-benefit jobs retain all workers that they attracted in period 1.

Conditional on workers entering the labor market, the former is true only when both worker types accept the first sampled offer. Since sophisticated workers have a higher perceived benefit from further search upon sampling w_L, both types accept the first sampled offer as long as sophisticated workers do so:

$$u_1^S(w_H) - u_1^S(w_L) = 2(w_H + \bar{b}) - (w_L + \underline{b}) - \max\{(w_L + \underline{b})\,;\,(v_2^S - c_2)\} \leq c_1/p. \quad \text{(A2)}$$

Moreover, sophisticated workers have a higher perceived valuation of searching again in period 2. Thus, for any type to leave a low-wage, low-benefit job, the sophisticated workers must do so. Since $c_2 \geq c_1$, the sophisticated workers who search in period 2 accept the first sampled offer, as they do in period 1. These workers leave $(w_L, \underline{b})$ if:

$$w_L + \underline{b} < (v_2^S - c_2) = p(w_H + \bar{b}) + (1-p)(w_L + \underline{b}) - c_2 \iff$$
$$c_2 < p\big((w_H + \bar{b}) - (w_L + \underline{b})\big). \quad \text{(A3)}$$

Then, combining (A2) with (A3) yields:

$$c_1 \;\geq\; p\big((w_H + \bar{b}) - (w_L + \underline{b})\big) \;+\; p\big(\underbrace{(w_H + \bar{b}) - (v_2^S - c_2)}_{>0}\big) \;>\; c_2,$$

which contradicts $c_2 \geq c_1$. Therefore, there does not exist an augmenting equilibrium in which both job offers are accepted by both types in period 1, but at least one type leaves a low-wage, low-benefit job in period 2. Intuitively, that is because this kind of equilibrium would require that sophisticated workers accept low-wage jobs in period 1, despite anticipating that they will subsequently leave. This cannot be optimal if the search costs are (weakly) increasing (Corollary 1).

Second, consider a putative augmenting equilibrium in which high-wage, high-benefit jobs attract more workers than low-wage, low-benefit jobs, but both job types retain the same proportion of workers. The former requires that sophisticated workers search until they sample w_H in period 1, while naïve workers accept the first sampled offer. The latter requires that no worker searches in period 2.

Anticipating to stay, sophisticated workers search for w_H in period 1 as long as:

$$u_1^S(w_H) - u_1^S(w_L) \;=\; 2((w_H - w_L) + (\bar{b} - \underline{b})) \;>\; c_1/p,$$

while naïve workers accept also w_L if:

$$u_1^N(w_H) - u_1^N(w_L) \;=\; (w_H - w_L) + p\big((w_H + \bar{b}) - \max\{(w_L + \bar{b})\,;\,(v_2^N - c_2)\}\big) + (1 - p)\big(\max\{(w_H + \underline{b})\,;\,(v_2^N - c_2)\} - \max\{(w_L + \underline{b})\,;\,(v_2^N - c_2)\}\big) \;\leq\; c_1/p.$$

Naïve workers remain employed in low-wage, low-benefit jobs if $(w_L + \underline{b}) \geq (v_2^N - c_2)$, under which the above simplifies to:

$$u_1^N(w_H) - u_1^N(w_L) \;=\; 2(w_H - w_L) \;\leq\; c_1/p.$$

Then, naïve workers accept the first sampled offer also in period 2, which implies:

$$(v_2^N - c_2) \;=\; p(w_H + \bar{b}) + (1 - p)(w_L + \underline{b}) - c_2 \;\leq\; w_L + \underline{b} \;\iff$$

$$c_2 \;\geq\; p\big((w_H + \bar{b}) - (w_L + \underline{b})\big). \tag{A4}$$

For high-wage firms to not deviate from providing high benefits, at least sophisticated workers should leave employment upon discovering low benefits. When (A4) holds, sophisticated workers searching in period 2 would accept the first sampled offer, which implies:

$$w_H + \underline{b} \;<\; (v_2^S - c_2) \;=\; p(w_H + \bar{b}) + (1 - p)(w_L + \underline{b}) - c_2 \;\iff$$

$$c_2 \;<\; (1 - p)(\underbrace{w_L - w_H}_{<0}) + p(\bar{b} - \underline{b}),$$

contradicting (A4). Intuitively, no augmenting equilibrium can sustain the workers' decision rule under which naïve workers remain in employment upon discovering $(w_L, \underline{b})$ and sophisticated workers leave $(w_H, \underline{b})$ despite $(v_2^S - c_2) = (v_2^N - c_2)$.

□

Thus, in any augmenting equilibrium, if it exists, the low-wage, low-benefit jobs must attract a strictly lower mass of workers as well as induce a strictly higher turnover. The former requires that only naïve workers accept the low-wage offers in period 1. The latter requires that all naïve workers who have accepted a low-wage job in period 1 leave upon discovering low benefits.

In period 1, sophisticated workers search for $(w_H, \bar{b})$ if:

$$u_1^S(w_H) - u_1^S(w_L) \;=\; 2((w_H - w_L) + (\bar{b} - \underline{b})) \;>\; c_1/p,$$

while naïve workers accept also w_L if:

$$u_1^N(w_H) - u_1^N(w_L) \;=\; (w_H - w_L) + p\big((w_H + \bar{b}) - \max\{(w_L + \bar{b})\,;\,(v_2^N - c_2)\}\big) + (1 - p)\big(\max\{(w_H + \underline{b})\,;\,(v_2^N - c_2)\} - \max\{(w_L + \underline{b})\,;\,(v_2^N - c_2)\}\big) \;\leq\; c_1/p.$$

A naïve worker leaves a job offering $(w_L, \underline{b})$ if $w_L + \underline{b} \;<\; (v_2^N - c_2)$. Since they accept the first sampled offer in period 1, $c_2 \geq c_1$ implies that the same must be optimal in period 2 and therefore $(v_2^N - c_2) = p(w_H + \bar{b}) + (1 - p)(w_L + \underline{b}) - c_2$. Thus, a naïve worker leaves $(w_L, \underline{b})$ if:

$$c_2 \;<\; p\big((w_H - w_L) + (\bar{b} - \underline{b})\big).$$

Firms are indifferent between offering either one of the two compensation packages if the following equal profits condition accounting for worker flows holds:

$$\begin{aligned}
\pi(w_H) \;&=\; \Big(2\,\frac{1-\lambda}{p} + 2\lambda + \lambda(1-p)\Big) \times (y - w_H - (1-\tau)\bar{b}) \;= \\
\pi(w_L) \;&=\; (\lambda + \lambda(1-p)) \times (y - w_L - (1-\tau)\underline{b}) \;\;\Longleftrightarrow \\
&\Big(2\,\tfrac{1-\lambda}{p} + (3-p)\lambda\Big)/((2-p)\lambda) \;=\; \tfrac{(y - w_L - (1-\tau)\underline{b})}{(y - w_H - (1-\tau)\bar{b})}.
\end{aligned}$$

To understand the above expression note the following. First, in period 1 each firm offering low wages is contacted by a naïve worker with probability λ. Any such worker accepts the firm's offer, but leaves after one period of employment. In period 2, the mass of naïve workers who are searching again is therefore $\lambda(1 - p)$, which is spread evenly across all firms. Thus, every low-wage firm expects to hire workers for a total number of $\lambda + \lambda(1 - p)$ periods. Second, in period 1 the mass $(1 - \lambda)$ of sophisticated workers is spread evenly across the mass

p of firms offering high wages. Once hired, sophisticated workers stay in the same employment for two periods. In addition, a firm offering high wages is contacted by a naïve worker with probability λ. Any such worker accepts the offer and remains in employment for two periods. Finally, in period 2 a high-wage firm expects to hire $\lambda(1-p)$ of naïve workers who are searching again, thus yielding $2\,\frac{1-\lambda}{p} + 2\lambda + \lambda(1-p)$ of expected periods of employment.

While high-wage firms bear a strictly higher cost of employing each worker, they also attract a strictly larger mass of workers, which makes it possible for the two compensation packages to yield equivalent expected profits.

Next, one needs to rule out profitable deviations by the firms. For $(w_H, \bar{b})$ to be offered in equilibrium, at least sophisticated workers must leave employment upon discovering w_H paired with $\underline{b}$, which should lower the deviating firm's profit. Since $c_2/p < (w_H + \bar{b}) - (w_L + \underline{b})$, $(v_2^S - c_2) = w_H + \bar{b} - c_2/p$ and thus the sophisticated worker hired in period 1 searches again upon discovering low benefits if:

$$(v_2^S - c_2) \;>\; w_H + \underline{b} \iff c_2/p \;<\; (\bar{b} - \underline{b}).$$

To focus attention, consider a case when a naïve worker would also leave $(w_H, \underline{b})$:[8]

$$c_2 \;<\; -(1-p)(w_H - w_L) + p(\bar{b} - \underline{b}).$$

Then, the deviating firm is not better off as long as:

$$\begin{gathered}(2\,\tfrac{1-\lambda}{p} + 2\lambda + \lambda(1-p)) \times (y - w_H - (1-\tau)\bar{b}) \;\geq \\ (\tfrac{1-\lambda}{p} + \lambda + \lambda(1-p)) \times (y - w_H - (1-\tau)\underline{b}) \iff \\ (2\,\tfrac{1-\lambda}{p} + \lambda(3-p))/(\tfrac{1-\lambda}{p} + \lambda(2-p)) \;\geq\; \tfrac{(y - w_H - (1-\tau)\underline{b})}{(y - w_H - (1-\tau)\bar{b})}.\end{gathered}$$

For $(w_L, \underline{b})$ to be offered in equilibrium, the deviation to high benefits should be unprofitable. Such a deviation allows the firm to retain the naïve workers that it hires in period 1 when:

$$c_2 \;\geq\; p(w_H - w_L) - (1-p)(\bar{b} - \underline{b}),$$

but is nonetheless unprofitable if:

$$\begin{gathered}(\lambda + \lambda(1-p)) \times (y - w_L - (1-\tau)\underline{b}) \;\geq\; (2\lambda + \lambda(1-p)) \times (y - w_L - (1-\tau)\bar{b}) \iff \\ (3-p)/(2-p) \;\leq\; \tfrac{(y - w_L - (1-\tau)\underline{b})}{(y - w_L - (1-\tau)\bar{b})}.\end{gathered}$$

Given their beliefs, the condition for naïve workers to accept w_L in period 1 becomes:

[8]If naïve workers remained in employment upon discovering $(w_H, \underline{b})$, this would increase the high-wage firms' incentive to deviate to low benefits, making an augmenting equilibrium only more difficult to sustain.

$$(1+p)(w_H - w_L) \;\leq\; c_1/p.$$

The ex-post PC of a naïve worker determining w_L is:

$$u_1^N(w_L) \;=\; w_L + p(w_L + 2\bar{b}) + (1-p)(\underline{b} + p(w_H + \bar{b}) + (1-p)(w_L + \underline{b}) - c_2) \;\geq\; 0.$$

In turn, the ex-post PC of a sophisticated worker searching for w_H in period 1 that would guarantee full labor market participation is:

$$u_1^S(w_H) - c_1/p \;=\; 2(w_H + \bar{b}) - c_1/p \;\geq\; 0.$$

Thus, in equilibrium, the sophisticated worker's binding PC pins down w_H, while the naïve worker's binding PC determines w_L. Note that as long as the resulting wages satisfy $w_H > w_L$, then $u_1^N(w_L) = 0$ implies $u_1^N(w_H) > 0$.

A firm marginally reducing its wage offer below w_L attracts no workers. In turn, a firm marginally reducing its wage offer below w_H satisfying $u_1^S(w_H) - c_1/p = 0$ experiences a discrete loss in profits if and only if $w_H - \epsilon$ is rejected by sophisticated workers with pessimistic beliefs in favor of further search for $(w_H, \bar{b})$. This is indeed the case when:

$$c_1/p \;<\; 2(\bar{b} - \underline{b}).$$

Otherwise, only a version of the sophisticated worker's ex-post PC associated with partial labor market participation (i.e., $u_1^S(w_H) = 0$) could hold in equilibrium. However, the implied $w_H = -\bar{b}$ would necessarily contradict $w_H > w_L$, where w_L is derived from the naïve worker's PC, and thus we impose $c_1/p < 2(\bar{b} - \underline{b})$.

What about deviations to some wage offer different from w_L or w_H? Deviating to any $w < w_L$ or $w > w_H$ can never be profitable.[9] If there exists an offer $w \in (w_L, w_H)$ which combined with $\bar{b}$ allows the firm to attract and retain both worker types, then this offer must be dominated by $(w_H, \bar{b})$ due to the sophisticated workers' pessimistic beliefs. Otherwise, consider a wage offer $w \in (w_L, w_H)$ that attracts only naïve workers. When paired with low benefits, any such offer is dominated by $(w_L, \underline{b})$, because even $(w_H, \underline{b})$ does not allow the firm to retain naïve workers. Denote by $\tilde{w} < w_H$ the wage such that:

$$\tilde{w} + \bar{b} \;=\; (v_2^N - c_2) \;\iff\; \tilde{w} \;=\; pw_H + (1-p)w_L - (1-p)(\bar{b} - \underline{b}) - c_2.$$

[9]Offers with $w < w_L$ attract no workers, because a sophisticated worker would search further and a naïve worker would take up their outside option. In turn, deviating to some $w > w_H$ which is not accepted by sophisticated workers due to their pessimistic beliefs, combined with either $\underline{b}$ or $\bar{b}$, is dominated by $(w_L, \bar{b})$, which is itself sub-optimal. Deviating to $w > w_H$ high enough so that it attracts (and retains) both sophisticated and naïve workers is dominated by $(w_H, \bar{b})$.

That is, $\tilde{w}$ is the lowest wage that allows the firm to retain a naïve worker when paired with high benefits. Notice that if $\tilde{w} \geq w_L$, then the deviation to $(\tilde{w}, \bar{b})$ is dominated by $(w_L, \bar{b})$, which is itself sub-optimal. If $\tilde{w} < w_L$, then such an offer is rejected by a naïve worker. Thus, there is no room for a profitable deviation in terms of the wage offer.

To sum up, for such an augmenting equilibrium to exist, the following must hold simultaneously:[10]

(i) $c_1/p \; < \; 2(\bar{b} - \underline{b})$,

(ii) $c_1/p \; \geq \; (1+p)(w_H - w_L)$,

(iii) $(2\,\frac{1-\lambda}{p} + (3-p)\lambda)/((2-p)\lambda) \; = \; \frac{(y - w_L - (1-\tau)\underline{b})}{(y - w_H - (1-\tau)\bar{b})}$,

(iv) $c_2 \; < \; p(\bar{b} - \underline{b}) - (1-p)(w_H - w_L)$,

(v) $(2\,\frac{1-\lambda}{p} + (3-p)\lambda)/(\frac{1-\lambda}{p} + (2-p)\lambda) \; \geq \; \frac{(y - w_H - (1-\tau)\underline{b})}{(y - w_H - (1-\tau)\bar{b})}$,

(vi) $c_2 \; \geq \; p(w_H - w_L) - (1-p)(\bar{b} - \underline{b})$,

(vii) $(3-p)/(2-p) \; \leq \; \frac{(y - w_L - (1-\tau)\underline{b})}{(y - w_L - (1-\tau)\bar{b})}$,

(viii) $w_L + p(w_L + 2\bar{b}) + (1-p)(\underline{b} + p(w_H + \bar{b}) + (1-p)(w_L + \underline{b}) - c_2) \; = \; 0$,

(ix) $2(w_H + \bar{b}) - c_1/p \; = \; 0$.

Note that the above imply some interesting properties, in addition to those outlined in Lemma A1. First, combining conditions (i) and (ix) implies that $(w_L + \underline{b}) < 0$, i.e. the low-wage, low-benefit package must be exploitative in the sense that it delivers utility lower than the worker's outside option. Second, combining conditions (iv) and (vi) implies that $(\bar{b} - \underline{b}) > (w_H - w_L)$, i.e. the wage differential that arises in such an equilibrium must be "small".

When searching for parameter constellations under which this kind of equilibrium would exist, I was only able to generate examples in which a single condition, typically condition (v), fails by an arbitrarily small margin while all the other conditions hold. That is likely due to the tension between conditions (v) and (vii) which capture that it is optimal for the firms to pair w_H, but not w_L, with high benefits. Although these two conditions are not inherently contradictory, in all the numerical examples I generated, only one of them could be satisfied at a time, given the remaining conditions. In limiting cases, however, the margin by which condition (v) fails is vanishingly small.[11]

[10] Note that $c_1/p < 2(\bar{b} - \underline{b})$ implies $c_1/p < 2((w_H - w_L) + (\bar{b} - \underline{b}))$, while $c_2 < -(1-p)(w_H - w_L) + p(\bar{b} - \underline{b})$ implies both $c_2 < p(\bar{b} - \underline{b})$ and $c_2 < p\big((w_H - w_L) + (\bar{b} - \underline{b})\big)$.

[11] For example, consider $\lambda = 0.999$, $\bar{b} = 131.615$, $\underline{b} = 0$, $c_1 = 47.270$, $c_2 = 85.427$, $\tau = 0.628$, and $y = 10.102$, which solve for $p = 0.748$, $w_H = -100.003$, and $w_L = -100.043$. Then, all conditions hold except for (v), which fails by a margin of less than 0.05%.

Based on these examples, notice that in the limit as $\lambda \to 1$, the left-hand sides of conditions (iii), (v), and (vii) coincide. Then, to avoid contradiction between (iii) and (vii), $w_H = w_L$ must hold. Consider a putative equilibrium where firms offer the same wage but differentiate their benefits, with p denoting the fraction of firms offering high benefits. Then, correlation neglect plays no role and the beliefs, as well as search rules, of naïve and sophisticated workers coincide. The above system of conditions simplifies considerably to:

(i*) $c_2 \;<\; p(\bar{b} - \underline{b})$,

(ii*) $(3-p)/(2-p) \;=\; \frac{(y-w-(1-\tau)\underline{b})}{(y-w-(1-\tau)\bar{b})}$,

(iii*) $w + p(w + 2\bar{b}) + (1-p)(\underline{b} + p(w+\bar{b}) + (1-p)(w+\underline{b}) - c_2) \;=\; 0$,

which describe an equilibrium in which the workers of either type accept the first sampled offer in period 1 and search again in period 2 if the realized benefits are low. The firms are indifferent between retaining all workers they attracted in period 1 with $\bar{b}$, and losing them with $\underline{b}$. There are no profitable deviations to either higher or lower wages. In this equilibrium, strictly dominant and dominated compensation packages coexist, but relative to an augmenting equilibrium with $w_H > w_L$, the dominance is restricted to the benefit dimension of an offer.

G.1 Remaining Derivations

Recall that naïve workers searching again in period 2 upon discovering $(w_L, \underline{b})$ is a prerequisite for an augmenting equilibrium to exist (Lemma A1), while no worker has an incentive to search further upon discovering $(w_H, \bar{b})$. Here, I consider alternative assumptions about the expected behavior of naïve workers off the equilibrium path, i.e. upon discovering $(w_L, \bar{b})$ and $(w_H, \underline{b})$, which can be invoked to construct an augmenting equilibrium.

Naïve workers leave upon observing $(w_L, \bar{b})$. A decision rule under which naïve workers leave low-paying jobs in period 2 irrespective of the realization of benefits must be inconsistent with them accepting low-wage offers in period 1. This follows from Corollary 1.

Naïve workers stay upon observing both $(w_L, \bar{b})$ and $(w_H, \underline{b})$. If only sophisticated workers were to leave a job providing $(w_H, \underline{b})$, the high-wage firms would have a strictly greater incentive to deviate to low benefits, relative to the case considered above. More precisely, naïve workers stay upon discovering $(w_H, \underline{b})$ if:

$$c_2 \;\geq\; -(1-p)(w_H - w_L) + p(\bar{b} - \underline{b}).$$

Then, the deviation to low benefits is not profitable for a high-wage firm as long as:

$$(2\,\tfrac{1-\lambda}{p} + (3-p)\lambda)/(\tfrac{1-\lambda}{p} + (3-p)\lambda) \;\geq\; \tfrac{(y-w_H-(1-\tau)\underline{b})}{(y-w_H-(1-\tau)\bar{b})}.$$

Under alternative beliefs regarding off-path realizations, naïve workers accept the first sampled offer in period 1 when:

$$u_1^N(w_H) - u_1^N(w_L) \;=$$
$$(w_H - w_L) + p\big((w_H + \bar{b}) - (w_L + \bar{b})\big) + (1-p)\big((w_H + \underline{b}) - (v_2^N - c_2)\big) \;=$$
$$(2 + p^2 - p)(w_H - w_L) - p(1-p)(\bar{b} - \underline{b}) + (1-p)c_2 \;\leq\; c_1/p,$$

and their PC is binding when:

$$u_1^N(w_L) \;=\; w_L + p(w_L + 2\bar{b}) + (1-p)(\underline{b} + p(w_H + \bar{b}) + (1-p)(w_L + \underline{b}) - c_2) \;=\; 0.$$

The remaining conditions underlying such augmenting equilibrium coincide with those outlined above and thus it exists as long as the following hold simultaneously:

(i) $c_1/p \;<\; 2(\bar{b} - \underline{b})$,

(ii) $c_1/p \;\geq\; (2 + p^2 - p)(w_H - w_L) - p(1-p)(\bar{b} - \underline{b}) + (1-p)c_2$,

(iii) $(2\,\frac{1-\lambda}{p} + (3-p)\lambda)/((2-p)\lambda) \;=\; \frac{(y-w_L-(1-\tau)\underline{b})}{(y-w_H-(1-\tau)\bar{b})}$,

(iv) $c_2 \;\geq\; p(\bar{b} - \underline{b}) - (1-p)(w_H - w_L)$,

(v) $c_2 \;\geq\; p(w_H - w_L) - (1-p)(\bar{b} - \underline{b})$,

(vi) $c_2 \;<\; p(\bar{b} - \underline{b})$,

(vii) $(2\,\frac{1-\lambda}{p} + (3-p)\lambda)/(\frac{1-\lambda}{p} + (3-p)\lambda) \;\geq\; \frac{(y-w_H-(1-\tau)\underline{b})}{(y-w_H-(1-\tau)\bar{b})}$,

(viii) $(3-p)/(2-p) \;\leq\; \frac{(y-w_L-(1-\tau)\underline{b})}{(y-w_L-(1-\tau)\bar{b})}$,

(ix) $w_L + p(w_L + 2\bar{b}) + (1-p)(\underline{b} + p(w_H + \bar{b}) + (1-p)(w_L + \underline{b}) - c_2) \;=\; 0$,

(x) $2(w_H + \bar{b}) - c_1/p \;=\; 0$.