Delegated Recruitment with Workers of Uncertain Productivity

Jacob Kohlhepp* and Stepan Aleksenko

March 15, 2021

Abstract

In this paper, we analyze a model of delegated recruitment. We assume that after search but before hire job applicants have uncertain productivity, characterized by a mean and variance. We constrain the contract space to binary refund contracts. These contracts, where the recruiter receives an upfront fee and issues a refund if the searched employee falls below an exogenous threshold, are commonly observed. We find that the contract induces risk averse behavior in a risk-neutral recruiter: the applicants hired through delegated search have lower productivity variance. In a parametric version of the model, we show that the efficiency of delegated search hinges crucially on the level of interim variance heterogeneity.

*Department of Economics, UCLA. Contact: jkohlhepp@ucla.edu. We are grateful for the guidance of Simon Board, Moritz Meyer-ter-Vehn, and Tara Sinclair. We have benefited greatly from comments by participants in the UCLA Theory Proseminar.

1 Introduction

An influential literature in economics has explored worker-driven job search. While this type of search remains important, the labor market landscape is changing. According to Black, Hasan, and Koning (2020) around 18 percent of US workers are now hired through firm-driven recruiting which are undertaken by a recruiter on behalf of the firm. Interviews with practitioners reveal most firms utilize a simple refund contract to compensate recruiters, wherein the recruiter keeps their fee if the worker remains employed for an initial period. At the same time, theoretical and empirical work in the search and matching literature has highlighted that the productivity of a worker-firm match is uncertain prior to hire, and learned over time.

We bring these delegation, contracting and uncertain match productivity aspects together in a model of delegated recruitment. In the model, a firm must hire a recruiter to search over workers with uncertain productivity. After search but prior to hire, a worker's productivity is a random variable with known mean and variance. Using the model, we analyze the impact of delegation on the types of applicants a firm hires. We find the contract results in a special type of moral hazard which we term induced risk-aversion: the recruiter wastes search effort constructing a candidate pool which has less productivity variance than the first-best benchmark.

We also explore how worker productivity heterogeneity impacts social welfare. In general, the contract restriction prevents the firm from achieving first-best social welfare. In a parametric specification of the model, the social welfare loss from delegation hinges crucially on the Pareto-index of the productivity variance distribution. When there are a small fraction of workers with very high productivity variance, delegation results in large agency loss for the firm, driving down social welfare. When productivity variance is similar across workers, delegation results in less agency loss for the firm, increasing social welfare. Thus, our model highlights that it is not the level of productivity variance which matters, but the relative distribution of variance across workers.

Our modeling of productivity uncertainty is motivated by the idea that upon finding a worker, a recruiter forms some sort of belief about how the worker will perform in the role. We do not formally specify the information structure or beliefs. Rather, we specify the joint distribution of the mean and variance. Intuitively, this joint distribution is a reduced-form object which involves the true underlying productivity distribution and the information available to the recruiter prior to hire.

To illustrate this framework, consider two candidates for a data science position at an established corporation. Candidate A is traditional: they went to a four year college, got a degree in statistics, and interned at another large corporation. Candidate B is nontraditional: they only have a high school degree and are self-taught, but they won a freelance machine-learning competition (something like a Kaggle competition).

Such competitions often feature thousands of competitors, and winning requires a blend of ingenuity and luck. Thus, candidate B might have a higher expected productivity than candidate A. However, there is a large element of chance in the competition. It is hard to say how Candidate B will perform in a structured work environment. Also, a recruiter probably has little experience placing such a non-traditional applicant. As a result, candidate B may have high productivity variance. In comparison, A has a traditional background, and the recruiter has placed this type of candidate before. Thus, candidate A may have a lower productivity variance.

In this paper we restrict attention to a certain type of contract. We interviewed three recruiters, and all reported the typical contract is a fixed percentage of applicant salary conditional on the employee being hired. They also stated that most recruiters offer some sort of guarantee: if the employee leaves for any reason prior to some agreed upon number of days, the recruiter will either refund the fee or place a new candidate for free. Among the three we interviewed, two specifically mentioned the period was 90 days. This is consistent with a survey conducted by Top Echelon[1], which found 96% of recruiters offer a guarantee and 86% have a guarantee period between 30 and 90 days. Among those, 61% provided a replacement but not money back if the candidate failed to stay, while 26% offered a partial or full refund.

Returning to Applicants A and B, one can imagine situations where a firm would be willing to take a risk on B. That is, the firm would select B for the job if the firm directly searched. But because the recruiter is the one building the applicant pool, and the contract warps incentives, B may never make it to a hiring manager's desk. This is the sort of situation we are interested in understanding. Therefore, the implications of our results for human capital formation, income mobility, and inequality are large. To the extent precise signals of productivity are costly, the findings in this paper suggest current contracts may be causing recruiters to distribute opportunity unevenly, by being unwilling to take risks on less traditional applicants. Jobs form the main method through which individuals escape poverty and form human capital, so the human costs of this uneven opportunity can add up quickly.

Although the focus of this paper is recruitment, the framework can be used to think about any economic context where a principal delegates search over uncertain objects to an agent. Examples include loan intermediaries searching for lenders, financial advisers searching for stocks, matchmakers searching for husbands/wives, and lawyers searching for trial strategies. What is important is that search is delegated and the common contract takes a binary refund form.

The paper is organized as follows. In the next section we review other work and explain how this research fits within the broader literature. Next, we introduce the model. Following that, we present results

1. Top Echelon is a company that makes software for recruiters.

assuming only independence of the productivity variance and productivity mean. We then derive additional comparative statics assuming a Pareto distribution for both productivity characteristics. Finally, we discuss the results and their implications for policy and further research.

2 Literature

Although our model is driven primarily by observations in the recruiting industry, it draws deeply from three strands of the theoretical literature. The first strand is rather recent and niche, and we term it the "delegated search" literature. The most relevant paper in this literature is Ulbricht (2016), which explores a general delegated sequential search problem. The strongest result, and one very relevant for this paper, is that a principal can achieve the first-best even when the full set of searched applicants is unobserved. This is true regardless of the risk attitudes of the agent. This is in stark contract to the typical moral hazard problem, where the first-best cannot be achieved. Our model features moral hazard, in the sense that the firm cannot observe the workers which are searched but not chosen. Unlike Ulbricht (2016) we restrain the contract shape to be of a refund form, and consider the case where productivity is ex-ante uncertain.

The second strand is the more general delegated choice literature. Papers in this literature generally consider a principal who can choose the amount of discretion to give an agent. The principal must trade-off the comparative advantage of the agent (usually the agent has better information) with the bias of the agent. Within this strand, our work is most closely related to Armstrong and Vickers (2010) and Alexander Frankel (2014), which consider more abstract delegation problems. We contribute to this literature by specifying how particular contracting limitations can generate misaligned preferences. Additionally, we use the idea of aligned delegation introduced in Alexander Frankel (2014) to derive and understand our results.

The third strand consists of theoretical papers exploring delegation in the hiring process. Alex Frankel (2020) is the most relevant to this paper, and it focuses on the delegation of hiring decisions to a biased hiring manager. We extend this literature by considering the problem of delegating search to a recruiter who is biased due to contracting restrictions. In this way, our paper helps paint a fuller picture of the principal-agent problems that are a critical component of the hiring process.

In addition to the theoretical strands, our model is motivated by the large literature on labor search and matching. Early search and matching models incorporated homogeneous workers, but eventually transitioned to include heterogeneity in productivity prior to hire. Two papers which represent this stage of the literature well are Shimer and Smith (2000) and Postel–Vinay and Robin (2002). Another literature, starting with Jovanovic (1979), highlights how worker-firm match quality is learned over the course of the employment relationship. More recently, this literature has led to examining how things like informal referrals and networks and experience play a role in hiring (Galenianos 2013, Dustmann et al. 2016, Burks et al. 2015). Our paper considers how intermediation and the distribution of ex-ante heterogeneity interact to impact the matching process. By focusing on a previously overlooked part of the matching process (recruiters), we help build a more comprehensive picture of the aggregate labor market.

3 Model

Players and Actions: There is a single risk neutral firm which desires to fill a single job opening. To fill the opening, it hires a recruiter to search. The recruiter is risk neutral and operates a sequential search technology for applicants.

Applicants and Information: Applicants are fully described by a single attribute, a, which can be interpreted as output produced net of some fixed market wage.[2] a is realized at the end of the game and it is not observed by either the firm or the recruiter prior to hire. Instead, when the recruiter samples an applicant, it observes two characteristics: (μ, σ). These attributes are distributed in the population with joint CDF $G(\mu, \sigma)$. Conditional on observing these two characteristics, the distribution of a for that applicant is $N(\mu, \sigma^2)$. μ can be interpreted as the interim expected match productivity of worker i with the firm. Throughout the paper we refer to it as the interim mean. σ^2 can be interpreted as the interim variance of match productivity. We call them interim because they characterize the distribution of a before hire but after search.

Search Technology: The recruiter can take i.i.d. draws of applicants at unit cost c. There is no limit to the number of searches. After drawing an applicant, the recruiter views their attributes (μ, σ) and then may either suggest the applicant for hire or continue search. I assume there is no recall (this is without loss). Search takes place in a single period, so there is no discounting. The firm does not observe any of the applicants the firm searches but does not select.

Payoffs and Contracts: The firm is restricted to contracts of the form: $t(a) = \alpha - \beta \mathbb{I}\{a < 0\}$. That is a refund contract where α is the recruiter's payment if the search is successful and β is the refund if the employee does not fit the job (the recruiter must return it to the firm if the realized productivity a is less than 0). If the recruiter rejects the contract, both the firm and the recruiter get 0. If the recruiter accepts the contract she runs a search over applicants pool and offers one of them to the firm. The firm receives the realized net output a less the payments to the recruiter $t(a)$ and the recruiter receives the transfer $t(a)$ according to the contract less the search costs. We restrict attention to cases where recruitment is profitable.

2. Another interpretation is that a is productivity relative to some break-even type of worker, where break-even is a worker who produces exactly the market wage.

That is, $E[a] = E[\mu] > c$. [3]

Timing:

- First the firm commits to a contract consisting of (α, β).
- Then the recruiter accepts or rejects the contract. If she rejects, she receives her outside option 0 (the firms receives 0 too).
- If she accepts, the recruiter sequentially searches for an applicant by taking i.i.d. draws from G.
- The recruiter suggests one applicant to the firm, and the firm hires the applicant.
- Finally, a realizes, the firm receives a. The contract realizes.

3.1 Model Comments

We can interpret a as the residual benefit of hiring an employee above the market wage per period. Then total profit is exactly equal to a because after the first period, a is realized. If $a > 0$, the employee can request a wage increase so that in subsequent periods $a' = 0$. If $a < 0$ and a is only observed by the firm and the employee, the firm cannot lower the wage because the employee can still obtain the market wage elsewhere. Thus the firm fires the employee.

In our model, firing is not endogenous: the firm terminates the employee if $a < 0$. We think this is reasonable for three reasons. First, there is a sense in which productivity is unobservable and not easily contractible. Indeed, this is why we limit contracts to be binary. Firms can probably prove a worker was not meeting some average standard, but it is unlikely they can prove an employee is below an arbitrary threshold. Second, firms can sometimes be sued for discrimination or wrongful termination if their firing decision is not based on "just cause." There is a sense in which it would be unfair if the firm had a different termination rule for recruited and directly hired applicants. Third, there is a sense in which the firm cannot commit to a firing rule ex-ante, due to the hard-to-observe nature of a.

The two applicant dimensions, μ and σ, can be thought of as quantifying the point estimate of applicant productivity and the uncertainty surrounding the estimate. As an example, consider again Applicants A and B from the introduction. Recall Applicant A is traditional, while Applicant B is nontraditional. It is possible a recruiter sees these two as having the same expected performance, but the second might be perceived as having greater variance, because there is a large chance they will either outperform or under-perform the market wage.

3. There is always a contract inducing the recruiter to choose the first sampled applicant ($\alpha = c$, $\beta = 0$). Under the assumption $E[a] > c$, this contract is profitable for the firm, and therefore the outside option of not hiring anyone (which would yield profit of 0) is never a part of any equilibrium.

Importantly, two different recruiters could have different beliefs over the same applicant due to either true information asymmetries or biases. In this sense, σ also captures how familiar an applicant is to a recruiter: how frequently the recruiter has placed similar candidates. It also may be related to issues of homophily and statistical discrimination: recruiters may better understand the background of similar race applicants. In this way, the primitive $G(\mu, \sigma)$ can be thought of as a reduced form of the true distribution of productivity combined with an information structure.

3.2 Other Examples

The working example throughout will be the firm/recruiter relationship. But we believe our framework extends to many circumstances where search is over objects of uncertain quality. In the context of mortgage brokers, a would represent the return to the lender including the possibility of default. The broker then earns a fixed fee after closing the deal, and likely only holds liability for the return if the loan turns out to be unsuitable, which is determined by a set of legal requirements. In the case of venture capitalists, a represents the profit of the startup. When venture capital fund managers collect a fixed fee, they may be concerned with whether or not the return exceeds a specific threshold rather than with the exact realization of a. These examples highlight the key ingredients which make our model a good approximation:

1. Search is delegated.
2. The value of the searched object is uncertain prior to consumption.
3. The principal wishes to maximize expected value.
4. The agent wishes to maximize the probability the value exceeds a certain threshold.
5. The probability of exceeding a threshold is well approximated by the ratio of characteristics (in the normal case, mean and variance).
6. The firm does not have much control of this threshold.

4 Non-Parametric Analysis

In this section, we analyze the first-best benchmark and the actual equilibrium without imposing additional assumptions on G. The following object will be helpful.

Definition 1 *An **acceptance region**, denoted $\mathcal{D}_i$, is the set of applicant types (μ, σ) which are accepted.*

This object describes the search strategy of a player, and is similar in spirit to the reservation rule when search is over one-dimensional objects.

4.1 First-Best Benchmark

For the first-best benchmark, I consider the case when the firm can operate the search technology directly.[4] The firm is risk neutral, so it seeks to maximize expected profit. After searching an applicant, expected a is: $E[a|\mu,\sigma]=\mu$. As a result, the firm cares only about μ and the problem reduces to one-dimensional search.

Lemma 1 *In the first-best benchmark, where the firm operates the search technology directly, the acceptance region is given by:*

$$\mathcal{D}_F=\{\mu,\sigma|\mu\geq\mu^*\}$$

where μ^* *solves:*

$$c=\int_{\mu\geq\mu^*}(1-G_\mu(\mu))d\mu$$

Or equivalently

$$(\mathbb{E}[\mu|\mu>\mu^*]-\mu^*)\cdot\Pr(\mu>\mu^*)=c$$

The proof is provided in the Appendix. This lemma formalizes the idea that the firm does not intrinsically care about the uncertainty (σ) dimension of an applicant. Symmetry of the distribution of a combined with risk-neutrality implies that the firm is indifferent between the upside potential and downside risk. Lemma 1 also shows that the optimal threshold μ^* is positive since it is equal to the firm's expected profit, which is also positive due to the assumption that $\mathbb{E}[a]>c$. The last expression emphasizes that the firm selects μ^* by equating the marginal gain of an additional search (the left side) with the cost of an additional search (the right side).

4.2 Firm-Optimal Contract

We now consider the firm-proposing optimal contract, where the firm must delegate search to the recruiter. The full problem is characterized below.

$$\max_{\alpha,\beta,\mathcal{D}_R} E[a|(\mu,\sigma)\in\mathcal{D}_R]-\alpha+\beta E[\mathbb{I}\{a<0\}|(\mu,\sigma)\in\mathcal{D}_R] \tag{OBJ}$$

4. Equivalently, when there is no contract restriction.

s.t.

$$\alpha - \beta + U \geq 0 \tag{IR}$$

$$\mathcal{D}_R = \{\mu, \sigma | \beta E_a[\mathbb{I}\{a > 0\} | (\mu, \sigma)] - U \geq 0\} \tag{IC}$$

where U is the value function of the recruiter (after α is sunk) during the sequential search problem, defined as:

$$U = -c + \int \max\{\beta E_a[\mathbb{I}\{a > 0\} | (\mu, \sigma)], U\} dG(\mu, \sigma) \tag{VAL}$$

Throughout this paper, we call solution to this problem the "second-best" or the "delegated search equilibrium." In this problem, the firm must choose a search strategy for the recruiter and a compensation scheme. Because search is not observed, the search strategy must be incentive compatible given the contract form restriction. Throughout the rest of this paper, we denote Φ to be the standard normal CDF. With this notation, we can write the utility the recruiter obtains from a candidate with characteristics μ, σ as:

$$\beta E[\mathbb{I}\{a > 0\} | (\mu, \sigma)] = \beta \Phi\left(\frac{\mu}{\sigma}\right)$$

This simplification shows that the recruiter ranks candidates according to the ratio μ/σ. We will call this ratio **standardized productivity** throughout this paper.

Definition 2 *(Standardized productivity) Standardized productivity, denoted $\tilde{\mu}$, of a candidate is the ratio of her expected productivity over her productivity uncertainty (in other words, how many standard deviations candidate's expected productivity is away from zero)*

$$\tilde{\mu} = \frac{\mu}{\sigma}$$

Comparing the delegated to direct search, we see the firm's indifference curves are horizontal lines while the recruiter's curves are sloped lines emanating from the origin. Higher indifference curves have steeper slopes. When μ is positive, the recruiter prefers lower applicant uncertainty. When μ is negative, she prefers higher σ. Intuitively, the recruiter gains or loses nothing from applicant upside or downside, and wants to pick someone with the highest probability of being acceptable ($a > 0$). This is shown graphically in Figure 1.

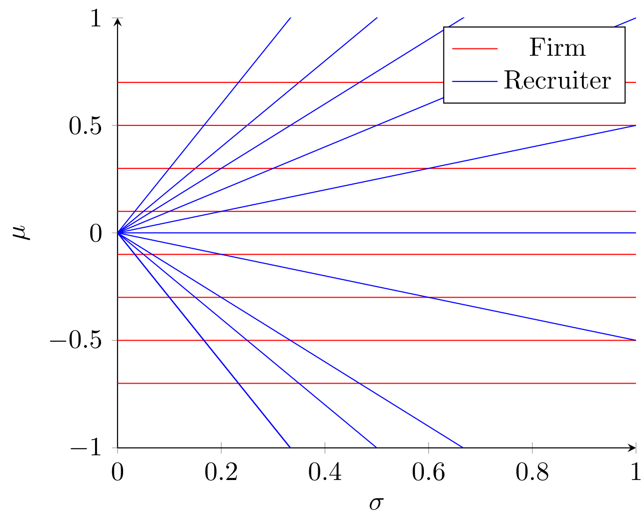


Figure 1: Indifference Curves

In order to focus on non-trivial cases of delegation, we introduce a weak assumption about the joint distribution G, which is expressed as follows.

Assumption 1 $\mathbb{E}[\mu|\tilde{\mu} = x]$ *is weakly increasing in* x.

This condition is often referred to in the statistics literature as positive quadrant dependence in expectation, which is slightly weaker than positive quadrant dependence and much weaker than positive affiliation. Intuitively, Assumption 1 means that larger standardized productivity implies larger expected productivity of the candidate. The assumption is quite natural given that $\tilde{\mu} = \mu/\sigma$. To break this assumption, the association between μ, σ most be positive and extremely strong. Consider the case of perfect dependence, where $\sigma = \gamma\mu$ for some constant γ. Even this is not sufficient, because this results in a flat conditional expectation. What is necessary is some form of dependence where the expectation grows faster than linearly or is not monotonic. For example, $\sigma = \gamma\mu^2$ will cause the conditional expectation to be decreasing.

For the rest of the analysis, this assumption is presumed to be satisfied. This assumption allows there to be a clear relationship between the firm's and the recruiter's preferences over $\tilde{\mu}$. This is important for further analysis as it eliminates situations where there is degenerate optimal search due to the diametrically opposed preferences over applicants.

From examining the shape of the recruiter's indifference curves we can see the shape of the acceptance regions. The recruiter's acceptance region will be a triangle, the area above one of the upward sloping blue lines. The firm's acceptance region will be a rectangle. Figure 2 illustrates this.

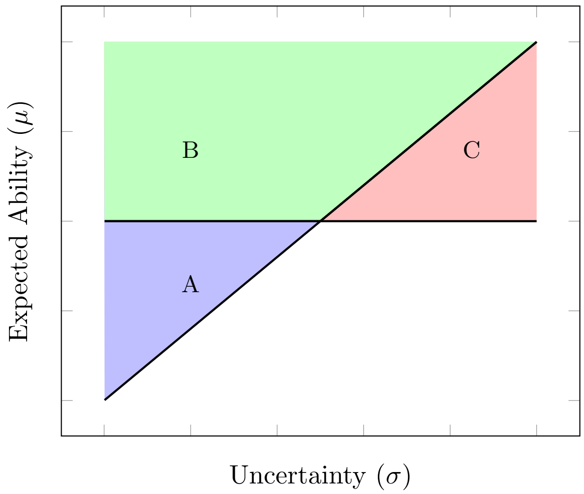


Figure 2: Recruiter vs. Firm Acceptance Regions Over Applicant Types

The green region represents applicants which both the firm (in the first-best benchmark) and the recruiter hire. The red region represents applicants which the firm would hire if it conducted search directly, but that are not selected by the recruiter in equilibrium. The blue region represents applicants which the firm would prefer not to hire, but that are selected anyway by the recruiter in equilibrium. The firm uses β to choose the slope of the diagonal line, trading-off the regions. We label these regions using the following formal definitions.

Remark 4.1 *When the support of expected productivity is positive, we can see from the graph that the recruiter is over-selecting "safe-bets." These low μ, low σ applicants are represented by the blue region. We call them "included inefficient" because they are selected but the firm would prefer they were excluded. The recruiter is also under-selecting "diamonds in the rough." These high μ, high σ applicants are represented by the red region. We call them "excluded efficient" because the firm would prefer they were included but they are excluded.*

We formalize our observations about the recruiter's problem in Lemma 2.

Lemma 2 *Given β, define $M(u)$ as the CDF of $u := \beta\Phi(\tilde{\mu})$. In any incentive compatible contract, the recruiter's acceptance region is given by:*

$$\mathcal{D}_R = \left\{\mu, \sigma | \tilde{\mu} \geq \Phi^{-1}\left(\frac{u^*}{\beta}\right)\right\}$$

where u^* *solves:*[5]

$$c = \int_{u \geq u^*} (1 - M(u))du$$

Lemma 2 is proved in the Appendix, but intuitively the proof is the same as Lemma 1, with the small additional step of defining the random variable u. By reducing the two-dimensional problem into one dimension, we can use the well-known result that the optimal strategy is a reservation rule, u^*. We next simplify the firm's problem using the previous lemmas.

Theorem 1 *The firm-optimal contract can be solved by solving the unconstrained problem:*

$$\max_x \mathbb{E}[\mu|\tilde{\mu} \geq x] - \frac{c}{Pr(\tilde{\mu} \geq x)}$$

which has the below F.O.C., uniquely defining $\tilde{\mu}^*$*:*

$$(\mathbb{E}[\mu|\tilde{\mu} \geq \tilde{\mu}^*] - \mathbb{E}[\mu|\tilde{\mu} = \tilde{\mu}^*]) \cdot \Pr(\tilde{\mu} \geq \tilde{\mu}^*) = c$$

Then β *can be obtained from the recruiter's IC constraint:*

$$\beta = \frac{c}{(E[\Phi(\tilde{\mu})|\tilde{\mu} \geq \tilde{\mu}^*] - E[\Phi(\tilde{\mu})|\tilde{\mu} = \tilde{\mu}^*]) \cdot \Pr(\tilde{\mu} > \tilde{\mu}^*)}$$

and α *from the recruiter's IR constraint:*

$$\alpha = \beta - \left(\beta \cdot E[\Phi(\tilde{\mu})|\tilde{\mu} \geq \tilde{\mu}^*]) - \frac{c}{\Pr(\tilde{\mu} > \tilde{\mu}^*}\right) = \beta \cdot (1 - E[\Phi(\tilde{\mu})|\tilde{\mu} = \tilde{\mu}^*])$$

Corollary 1.1 *Under the above assumption, the firm's profit is positive and equal to* $\mathbb{E}[\mu|\tilde{\mu} = \tilde{\mu}^*]$*. Then* $\tilde{\mu}^*$ *must positive too since the profit is positive under the assumption made in Section 3.*

Theorem 1 proves that the problem can be solved by unconstrained, single variable maximization. The firm uses α to extract all surplus from the recruiter, so that in the end the firm is maximizing social surplus subject to a constraint imposed by the contract form. The first term in the objective is the value of the employee to the firm, and the second term is the expected search cost to find an acceptable employee. We now derive our main nonparametric result.

Proposition 1 *If* μ *and* σ *are independent, the distribution of* σ *in the first-best acceptance region* $\mathcal{D}_F$ *first-order stochastically dominates the distribution of* σ *in the recruiter's acceptance region* $\mathcal{D}_R$*.*

5. This formulation is true for continuously distributed $\tilde{\mu}$ and a proper interior solution (non-degenerate search), but it can be easily generalized to a system of inequalities otherwise.

Proposition 1 demonstrates the contract form induces the recruiter to select less risky (lower σ) applicants. Part of the reason for this is that even if μ, σ are independent, $\sigma, \tilde{\mu}$ will often be negatively affiliated. This in turn implies that $\sigma | \tilde{\mu} > x$ will stochastically dominate the unconditional distribution of σ. Consider the following parametric example.

Example: $\mu \sim U[0,1], \sigma \sim exp(\lambda)$ where the two variables are independent. Then the joint pdf of $v := \mu/\sigma, \sigma$ is given by $f(v,\sigma) = \mathbb{I}\{v\sigma \leq 1\}\lambda exp(-\lambda\sigma)\sigma$. Two random variables are negatively affiliated if and only if their joint pdf is log submodular. Taking logs of the pdf gives: $log(f(v,\sigma)) = log(\mathbb{I}\{v\sigma \leq 1\}) - \lambda\sigma + log(\sigma)$. This function is submodular: if the indicator is 0 for at least one of two pairs of values, it will also be 0 for the pairwise minimum.

5 Parametric Analysis

In order to compare the first and second-best regions, and perform comparative statics, we impose a parametric assumption.

Assumption 2 ***Parametric Assumption.*** *μ, σ are distributed independently with marginal Pareto distributions. That is, their joint pdf is given by:*

$$g(\mu,\sigma) = \frac{\theta_\mu x_\mu^{\theta_\mu}}{\mu^{\theta_\mu+1}} \frac{\theta_\sigma x_\sigma^{\theta_\sigma}}{\sigma^{\theta_\sigma+1}} \mathbb{I}\{\mu \geq x_\mu\}\mathbb{I}\{\sigma \geq x_\sigma\}$$

where both variables have finite expectations ($\theta_\mu > 1, \theta_\sigma > 1$).

In this parameterization, x_μ, x_σ are the shift parameters, which give the lower bounds of the support. $\theta_\mu, \theta_\sigma$ are the shape parameters which intuitively control the amount of mass near the beginning of the distribution. Assumption 2 yields closed-form solutions to both the first-best and equilibrium problems. In the Appendix, we derive various intermediate joint and marginal distributions, including the distribution of $\tilde{\mu}$. Of particular interest is the following conditional expectation:

$$E[\mu | \tilde{\mu} = z] = \begin{cases} \frac{(\theta_\mu+\theta_\sigma)}{(\theta_\mu+\theta_\sigma-1)} x_\mu \text{ if } z \leq x_\mu / x_\sigma \\ \frac{(\theta_\mu+\theta_\sigma)}{(\theta_\mu+\theta_\sigma-1)} x_\sigma z \text{ else} \end{cases}$$

Note that this is weakly increasing in z, as required by Assumption 1. However, it is flat over an initial range. We know from Lemma 3 the firm will never select $\tilde{\mu}^*$ to be on the flat portion: it will always set $\tilde{\mu}^*$ to be the minimum of the support or it will set it above the flat portion. The first case is rather strange. it means that the search technology is not even being used, because the first applicant drawn is immediately hired. Thus, if the firm had the outside option of drawing a random applicant, then it would never use the recruiter. For this reason, and to focus on an interior solution for comparative statics, we make one additional assumption.

Assumption 3 ***Non-degenerate Search.*** *Search cost c and the parameters of the joint Pareto distribution satisfy:*

$$\frac{x_\mu \theta_\sigma}{(\theta_\mu + \theta_\sigma - 1)(\theta_\mu - 1)} \geq c$$

Then we can characterize interior closed-form solutions for the first-best and the second-best problems for the given distribution G.

Proposition 2 *If Assumption 2 is satisfied, the first-best benchmark has a unique solution characterized by:*

$$\mu^* := \left(\frac{x_\mu^{\theta_\mu}}{c(\theta_\mu - 1)}\right)^{\frac{1}{\theta_\mu - 1}}$$

If Assumption 3 is also satisfied, the firm-optimal contract has a unique solution characterized by:

$$\tilde{\mu}^* = \frac{1}{x_\sigma}\left(\frac{x_\mu^{\theta_\mu}\theta_\sigma}{c(\theta_\mu + \theta_\sigma - 1)(\theta_\mu - 1)}\right)^{\frac{1}{\theta_\mu - 1}}$$

The proof is provided in the Appendix because the steps are mainly algebraic with little economic insight. With these solutions in hand, we compare the second-best and the first-best optimal acceptance regions and the search intensity.

In the last set of results, we will focus on comparative statics in θ_σ. θ_σ represents the level of heterogeneity with respect to σ. As θ_σ rises, the population becomes more homogeneous. A greater number of workers have an interim productivity variance near the lower bound of the support given by x_σ. When it falls, the population becomes more heterogeneous. This can be visualized in Figure 3 which plots three different Pareto distributions with the same x_σ but different θ_σ.

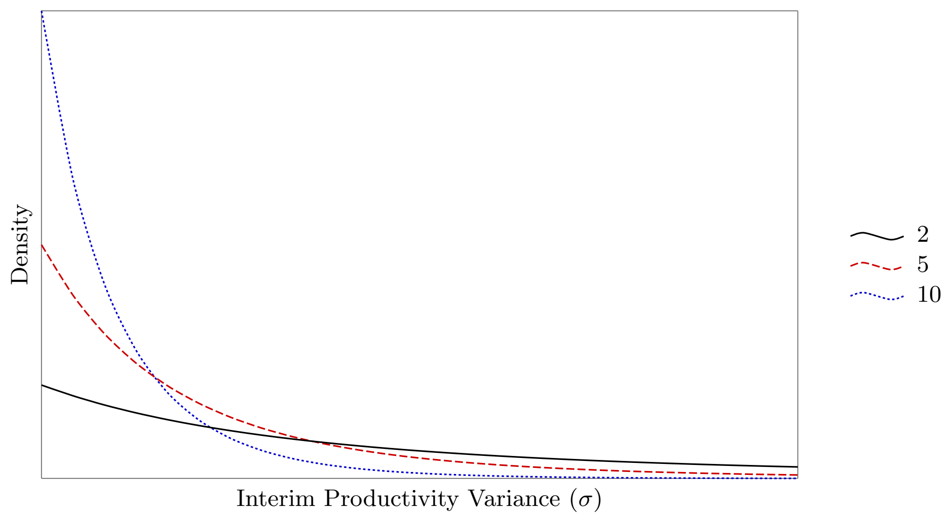


Figure 3: Densities of σ for Different Values of θ_σ

θ_σ is also a measure of inequality in a distribution, which is why it is often called the Pareto Index. In our case, θ_σ measures the level of inequality of interim productivity variance. When θ_σ is around 1.16 we can say 20 percent of workers possess 80 percent of interim productivity variance. When it increases to 1.59 the distribution becomes more equitable: around 33 percent of workers possess 67 percent of interim productivity variance. When $\theta_\sigma \to \infty$ we achieve perfect equality, where all workers have the same interim productivity variance. Figure 4 visualizes the inequality interpretation by plotting Lorenz curves of σ for different θ_σ values. The 45 degree line represents perfect equality ($\theta_\sigma \approx \infty$) and it is the Lorenz curve we converge to when $\theta_\sigma \to \infty$.

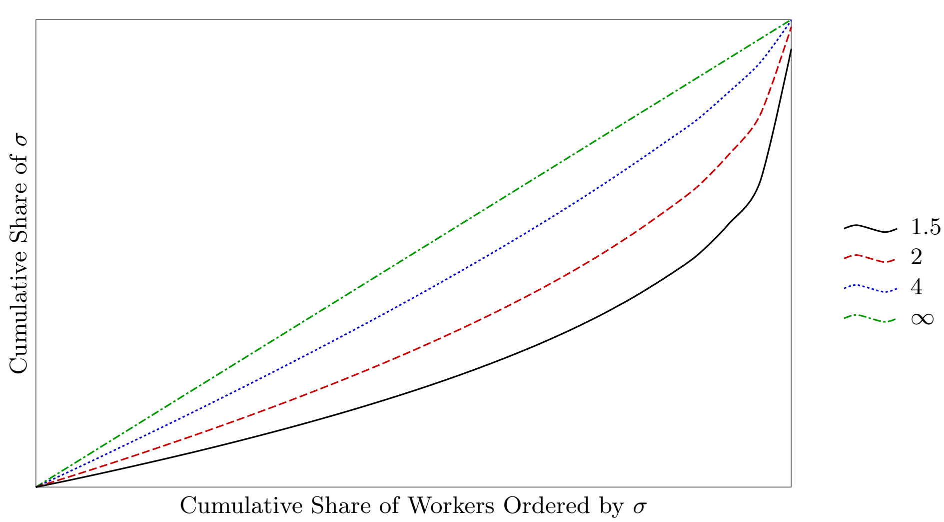


Figure 4: Lorenz Curves for Different Values of θ_σ

An interesting question is whether there is more or less search in the first-best as opposed to the second-best. We measure search intensity as the expected number of searches $Pr((/mu, \sigma) \in \mathcal{D}_i)^{-1}$.

Proposition 3 ***Search intensity.***

1. *Search intensity is always higher in the first-best than the second-best, that is:*

$$\frac{1}{Pr((\mu,\sigma) \in \mathcal{D}_{FB}} \geq \frac{1}{Pr((\mu,\sigma) \in \mathcal{D}_{FB}}$$

2. *Second-best search intensity is increasing in θ_σ and is equal to the first best search intensity in the limit $\theta_\sigma \to \infty$ (the acceptance regions are identical in the limit as well).*

Proposition 3 illustrates there is always less search in equilibrium than in the first-best. This is because the return to each additional search is lower when the firm has to delegate: the recruiter will sometimes pass over workers the firm wants it to select and will sometimes select someone the firm would prefer not to hire. As variance heterogeneity decreases (θ_σ rises) the agency loss declines and the marginal value of additional search improves. Eventually, when heterogeneity vanishes, search intensity in equilibrium achieves the first-best intensity.

This proposition means that the contract restriction in combination with variance heterogeneity induces a form of moral hazard, which is similar in flavor to canonical multi-tasking models (Holmstrom and Milgrom 1991) and the general literature on performance pay. We can think about our model as a multitasking model, where the two tasks are search along the μ dimension and search along the σ dimension. Like in classic multitasking models, the firm cannot separate the two dimensions, and can only encourage search over a separate measure. In our case, this separate measure is μ/σ. What is interesting is that if we take logs of this expression, we have:

$$log(\tilde{\mu}) = log(\mu) - log(\sigma)$$

What does this mean? Well, it implies the firm can only "buy" a 1 percent increase in μ if it is willing to "buy" a 1 percent decrease in the maximum σ. Individual rationality requires the firm compensate the recruiter for all search effort, so we are in a situation similar to the multitasking models where total effort is rewarded, but there is a wasteful task that cannot be distinguished properly from the productive task.

Proposition 4 ***Comparison of FB and SB acceptance regions.***

1. *The distribution of σ in the first-best acceptance region first-order stochastically dominates the distribution of σ in the recruiter's acceptance region.*

2. *The distribution of μ in the first-best acceptance region first-order stochastically dominates the distribution of μ in the recruiter's acceptance region.*

Proposition 4 first restates the general finding from the nonparametric results section. Under delegated search, accepted workers have lower interim productivity variance than in the first-best. The second part of the proposition states something new. Under the joint Pareto assumption, first-best accepted workers have higher interim expected productivity than second-best accepted workers in a first-order stochastic sense. Thus, in this particular parameterization, the recruiter is constructing a candidate pool which is safer but worse. There is a sense in which "safe bets" are being accepted at the expense of "diamonds in the rough," because the recruiter cannot be made to value upside potential.

Social welfare is another way to measure misalignment between the first and the second best solutions in this model. We wish to focus on the relative share of first-best social surplus which is lost due to delegation. To accomplish this, we introduce the concept of relative agency loss.

Definition 3 *Relative agency loss (RAL) is defined as the percentage of first-best surplus lost due to delegation. The social surplus is equal to the firm's profit in both FB and SB. Thus, we can compute relative agency loss in the following way.*

$$RAL = 1 - \frac{\Pi_{SB}}{\Pi_{FB}}$$

Comparative statics of relative agency loss are given in Theorem 2.

Theorem 2 *Relative agency loss has the following characteristics:*

1. *invariant to c, x_μ, x_σ*
2. *increasing in θ_σ*
3. $\lim_{\theta_\mu \to \infty} RAL(\theta_\mu) = 0$
4. $\lim_{\theta_\sigma \to \infty} RAL(\theta_\sigma) = 0$

The level of variances and means in the population do not impact efficiency. It is the amount of heterogeneity or inequality that matters. The shape parameter θ_σ captures the degree of inequality in the distribution of interim variance. When θ_σ is near 1, the distribution exhibits near perfect inequality: a small number of applicants have near infinite variance while the rest have low variance. The contract induces the recruiter to spend effort avoiding the high variance applicants, which is reflected in the agency loss. When θ_σ gets large, the distribution exhibits near perfect equality: all applicants essentially have the same variance. The recruiter becomes less tempted to search along the variance dimension, and we approach full efficiency.

We can think of comparative statics in x_σ, x_μ as uniform increases in the distributions of interim expected productivity and interim productivity variance. We have shown previously that these parameters do not effect efficiency because they merely change levels not the shape of the distribution.

One application of our findings about relative agency loss has to do with the decision to outsource recruiting. When firms decide whether to outsource the recruiting function, they must trade-off the agency loss that comes with outsourcing with the savings in terms of opportunity cost. We analyze how interim productivity variance heterogeneity impacts this decision in one final proposition.

Proposition 5 ***The Choice To Delegate.*** *Suppose the firm can either outsource recruiting to a recruiter who can perform search at unit cost c_R or it can perform search directly at unit cost c_F. Let c^* be the highest ratio of search costs c_R/c_F at which the firm prefers the delegation. Then c^* is increasing in θ_σ.*

Intuitively, the range of opportunity costs under which delegation is optimal increases if θ_σ increases or if workers become more homogeneous in terms of their interim productivity variance. Thus, we can see that the efficiency to delegate measured by the relative agency loss has a direct mapping to the choice to delegate if the firm can conduct its own search at higher cost. We can think of the ratio of c_R/c_F as measuring the comparative advantage of a recruiter conducting search compared to someone like the CEO. A recruiter is probably more efficient at searching for candidates, and a CEO probably has a higher opportunity cost.[6]

6 Discussion

Our model has several testable predictions about recruiters and the labor market. First, industries with large amounts of heterogeneity in productivity variance

Second, recruiters should suggest older experienced workers over younger inexperienced workers. Since Jovanovic (1979), studies have found empirical evidence of learning about match quality (lange and other one). In particular, Fredriksson, Hensvik, and Skans 2018 shows match quality is better among experienced workers than inexperienced. This is consistent with the existence of better information about experienced workers. In the language of our paper, older experienced workers should exhibit lower productivity variance, which in turn should make them preferred by recruiters.

Third, recruiters should be more likely to work with candidates of similar socioeconomic background, even when they have no intrinsic biases or prejudice. This is because similarity of backgrounds should improve communication and thus reduce the level of perceived interim productivity variance.

Third, the recruiter recommendations should reinforce path dependence.

6. For example, a CEO could use the time to plan firm strategy.

All the analysis in this paper highlights one key insight: the common bonus contract causes recruiters to waste effort finding applicants with low hiring risk rather than applicants with high expected productivity, despite the fact that the firm wants the recruiter to focus solely on expected productivity. This misaligned search strategy results in distorted applicant pools. Economy-wide, this means some applicants are receiving less attention from recruiters than is socially optimal.

Our findings are important for socioeconomic mobility. High quality signals of productivity are expensive. The cost of data science boot camps is often on the order of $2,000-$17,000 just for a small period of instruction (Williams 2020). Prestigious universities are usually either extremely expensive (a year's tuition can be in excess of the median annual salary) or extremely selective. Even with financial aid, individuals from disadvantaged backgrounds often do not have the resources to invest in the preparatory work needed to be admitted.[7]

Without credit constraints it wouldn't necessarily be an issue that high quality signals of productivity are expensive. But credit constraints and imperfect information are a real part of the labor market landscape. For example, currently only 71% of college applicants file the Free Application for Federal Student Aid (*How America Pays for College 2020*). As a result, job seekers will often need to pay for productivity signals using family support. Children of wealthy parents will tend to have lower productivity variance. This means that if we compared two workers with the same expected ability but different parental wealth, we would expect the child of wealthier parents to be approached more by recruiters. This will tend to reinforce existing socioeconomic inequality.

Our findings are also important for the literature on discrimination in hiring. Most past work on discrimination focuses on direct applications. We present another important pathway: firm-directed search. Our results imply that delegation of search changes the composition of the resulting candidate pool. As stated earlier, it could be that this delegation magnifies inequality of opportunity because recruiters prefer candidates which are similar to them (homophily). An interesting extension of our work would be to conduct a form of recruiter audit study. In such a study, artificial candidate profiles would be posted to a job search website like LinkedIn. The study would then measure the differences in how many recruiters reached out in a given time period.

A number of resume correspondence studies have shown disparities in callback rates across protected groups. Some scholars interpret this as evidence of racial and gender discrimination. A prominent critique of this

7. SAT preparation classes, tutoring, college admissions counseling, AP testing, etc.

6.1 The Contract Restriction

Crucially, it is restrictions on the space of contracts which generate inefficiency in our model. If a productivity is perfectly contractible, and a firm and recruiter could commit to any contract, it is straightforward to show that the first-best can be achieved and there is no agency loss. Indeed, simply allowing the firm to commit to a firing rule for the employee upfront will allow the firm to achieve first-best. Why then is the the binary refund contract (which motivated this entire paper) so common?

Before engaging in this discussion, we want to highlight this is a similar question discussed in Holmstrom and Milgrom (1991) under the heading ”Missing Incentive Clauses in Contracts.” As the authors point out, much of contracting theory points to the importance of explicit incentives in achieving efficiency (and this paper is no exception). Yet, real-world contracts often lack any explicit contingencies.Our discussion in this section is heavily inspired by

We propose two answers. First, measuring productivity of an individual person is hard, especially in non-manufacturing jobs with team-based work and productivity spillovers. Measuring whether someone still works somewhere is easy, and much less ambiguous. In addition, people are probably better at thinking in relative terms: is the engineering team better or worse since New Employee A joined? Is the company better or worse off if they fire New Employee A?

6.2 External vs. Internal Recruiters

The recruiting world is divided between two groups of recruiters. External recruiters who work for an agency or freelance but are not employees of a firm. And internal or retained recruiters who may either be paid a fixed retainer or a fixed salary. Two of the recruiters we interviewed were external recruiters: one employed by an agency, the other a sole proprietor. Internal recruiters on the other hand may play other roles in the firm beyond recruiting, and they can be paid fixed salaries. Of the three recruiters we interviewed only one was an internal recruiter. This person also happened to perform human resource roles beyond recruiting.

Our analysis so far applies most directly to external recruiters, as these individuals or the firms which employ them are paid solely based on placing an employee. However, it may also extend to internal recruiters. Multiple sources mention bonuses as a part of internal recruiter pay. If these bonuses are linked to number of hires or retention, the results of this paper extend. The one internal recruiter we spoke to mentioned that retention is often an explicit part of their annual performance bonus.

7 Conclusion

In the course of this paper, we outline a new theoretical framework of delegated recruitment with uncertain productivity. We characterize the general problem, and provide some results without specifying a distribution. We then impose parametric assumptions in order to compute comparative statics. Under the parametric assumptions, we show

The subject of intermediaries in the labor market has received little attention in both the theoretical and applied literature, and this paper is an attempt to build a cogent theory that captures the trade-offs of delegation in this environment. Moving up the job ladder is a process that frequently operates through recruiters, and it is key to improving socioeconomic outcomes. As a result, better understanding the incentives current contracts create for recruiters is more than just an academic exercise. It has wide-ranging implications for socioeconomic mobility, labor force composition, and government policy.

Our model represents a first step towards understanding the trade-offs firms face when deciding how to recruit. We show that agency loss from outsourcing or delegating the recruiting function depend largely on the degree of heterogeneity in productivity variance. How these findings fit into an equilibrium model of the labor market is an important question which we leave for further work. An insight which we hope will come from such future work will be how recruiters help generate the equilibrium distribution of productivity.

References

Armstrong, Mark, and John Vickers. 2010. "A model of delegated project choice." *Econometrica* 78 (1): 213–244.

Black, Ines, Sharique Hasan, and Rembrand Koning. 2020. "Hunting for talent: Firm-driven labor market search in America." *Available at SSRN.*

Burks, Stephen V, Bo Cowgill, Mitchell Hoffman, and Michael Housman. 2015. "The value of hiring through employee referrals." *The Quarterly Journal of Economics* 130 (2): 805–839.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker. 2016. "Referral-based job search networks." *The Review of Economic Studies* 83 (2): 514–546.

Frankel, Alex. 2020. "Selecting Applicants."

Frankel, Alexander. 2014. "Aligned delegation." *American Economic Review* 104 (1): 66–83.

Fredriksson, Peter, Lena Hensvik, and Oskar Nordström Skans. 2018. "Mismatch of talent: Evidence on match quality, entry wages, and job mobility." *American Economic Review* 108 (11): 3303–38.

Galenianos, Manolis. 2013. "Learning about match quality and the use of referrals." *Review of Economic Dynamics* 16 (4): 668–690.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design." *JL Econ. & Org.* 7:24.

Jovanovic, Boyan. 1979. "Job matching and the theory of turnover." *Journal of political economy* 87 (5, Part 1): 972–990.

Postel–Vinay, Fabien, and Jean–Marc Robin. 2002. "Equilibrium wage dispersion with worker and employer heterogeneity." *Econometrica* 70 (6): 2295–2350.

How America Pays for College 2020. `https://www.salliemae.com/about/leading-research/how-america-pays-for-college/`.

Shimer, Robert, and Lones Smith. 2000. "Assortative matching and search." *Econometrica* 68 (2): 343–369.

Ulbricht, Robert. 2016. "Optimal delegated search with adverse selection and moral hazard." *Theoretical Economics* 11 (1): 253–278.

Williams, Alex. 2020. *Our Ultimate Guide to the Best Data Science Bootcamps,* April. `https://www.coursereport.com/blog/best-data-science-bootcamps-the-complete-guide`.

8 Appendix

8.1 Proof of Lemma 1

Denote V as the value function of the firm. Denote the marginal distribution of μ as F. The dynamic programming problem of the firm is given by:

$$V = -c + \int \max\{E[a|\mu = u], V\} dF(\mu)$$

Note that if there was recall (so that the highest previously viewed μ could be carried as a state variable) the firm would never exercise the option. Because costs are already sunk, if it was previously optimal to search again it will still be optimal to search again the next period if the drawn μ is elss than the last

$$V = -c + \int \max\{\mu, V\} dF(\mu)$$

Re-writing, this yields:

$$0 = -c + \int \max\{\mu - V, 0\} dF(\mu)$$

So the optimal strategy is a reservation rule characterized by μ^*, where $V = \mu^*$. Thus:

$$c = \int \max\{\mu - V, 0\} dF(\mu) \leftrightarrow c = \int_{\mu > \mu^*} \mu - \mu^* dF(\mu)$$

Integration by parts gives:

$$c = -[(1 - F(\mu))(\mu - \mu^*)]_{\mu^*}^{\bar{\mu}} + \int_{\mu^*}^{\bar{\mu}} (1 - F(\mu)) d\mu$$

Since the first term is 0, this simplifies to:

$$c = \int_{\mu^*}^{\bar{\mu}} (1 - F(\mu)) d\mu$$

As an aside, note that we can re-arrange the intermediate equation this way:

$$c = \int_{\mu > \mu^*} \mu dF(\mu) - (1 - F(\mu^*))\mu^* \leftrightarrow \mu^* = \frac{1}{1 - F(\mu^*)} \left(\int_{\mu > \mu^*} \mu dF(\mu) - c \right)$$

which can compactly be re-written as:

$$\mu^* = E[\mu | \mu \geq \mu^*] - \frac{c}{Pr(\mu \geq \mu^*)}$$

8.2 Proof of Lemma 2

The dynamic programming problem of the firm is given by:

$$U = -c + \int \max\{\beta E_a[\mathbb{I}\{a \leq 0\} | (u, s)], U\} dG(\mu, \sigma)$$

$$U = -c + \int \max\{\beta \Phi(-\mu/\sigma), U\} dG(\mu, \sigma)$$

Re-writing, this yields:

$$0 = -c + \int \max\{\beta \Phi(-\mu/\sigma) - U, 0\} dG(\mu, \sigma)$$

Observe that utility only depends on μ/σ, so we can reduce the problem to one-dimensional search. As long as β is negative, utility will be increasing in μ/σ. The firm will always set $\beta \leq 0$ when $E[a|\tilde{\mu} = x]$ is increasing in x (which is what we assumed). Thus we will have a reservation rule strategy in the ratio μ/σ, that is we will select all μ, σ that satisfy:

$$u^* \leq \beta\Phi(-\mu/\sigma) \leftrightarrow \beta\Phi^{-1}(\frac{u^*}{\beta}) \leq \frac{\mu}{\sigma}$$

where u^* can be uniquely obtained in a similar argument to the last proof. Simply define M as the cdf of $u := \beta\Phi(-\mu/\sigma)$. Then:

$$c = \int \max\{u - V, 0\} dM(u) \leftrightarrow c = \int_{u>u^*} u - u^* dM(u)$$

Integration by parts gives:

$$c = -[(1 - M(u))(u - u^*)]_{u^*}^{\bar{u}} + \int_{u^*}^{\bar{u}} (1 - M(u))du$$

Since the first term is 0, this simplifies to:

$$c = \int_{u^*}^{\bar{u}} (1 - M(u))du$$

8.3 Proof of Theorem 1

Lemma 3 ***No-atom optimal search.*** *Let one search over a pool of uniformly distributed x with a payoff $f(x)$, ($f'(x) \geq 0$)), and a cost $c > 0$ per search. Let $\tilde{\mu}^* \in (0, 1)$ be a unique optimal search threshold. Then $\forall \varepsilon > 0 : f(\tilde{\mu}^* - \varepsilon) < f(\tilde{\mu}^* + \varepsilon)$.*

Proof. The intuition of the statement is that one being able to set a threshold on the CDF of the search variable (rather than the variable itself) would never strictly prefer to set it within an atom than anywhere else. The problem described in the lemma can be stated as

$$\max_{x'}\{\mathbb{E}[f(x)|x \geq x'] - \frac{c}{1 - x'}\}$$

The derivative with respect to x' is

$$(\mathbb{E}[f(x)|x \geq x'] - f(x')) * (1 - x') - c = (*)$$

Let us suppose that $\tilde{\mu}^*$ is the unique maximizer and that $\exists\ \varepsilon > 0 :$ s.t. $f(x)$ is flat on $(\tilde{\mu}^* - \varepsilon; \tilde{\mu}^* + \varepsilon)$. Let

$\bar{x} = \tilde{\mu}^* + \varepsilon$. Locally for $x' \in (\tilde{\mu}^* - \varepsilon; \tilde{\mu}^* + \varepsilon)$

$$\mathbb{E}[f(x)|x \geq x'] = \frac{(1-\bar{x}) * \mathbb{E}[f(x)|x \geq \bar{x}] + (\bar{x} - x') * f(\tilde{\mu}^*)}{1 - x'}$$

Than simplifying the derivative of the outcome with respect to x' gives

$$(*) = (1-\bar{x}) * \mathbb{E}[f(x)|x \geq \bar{x}] + (\bar{x} - x') * f(\tilde{\mu}^*) - f(\tilde{\mu}^*) * (1 - x') - c$$

$$= (1-\bar{x}) * (\mathbb{E}[f(x)|x \geq \bar{x}] - f(\tilde{\mu}^*)) - c$$

which apparently does not depend on x' and is constant for $x' \in (\tilde{\mu}^* - \varepsilon; \tilde{\mu}^* + \varepsilon)$. Then $\tilde{\mu}^*$ cannot be a unique maximizer since depending on the sign of the derivative one should either increase the threshold or decrease it or is indifferent in some small neighborhood around $\tilde{\mu}^*$. ■

We apply Theorem 1 to the firm's problem which is given by Equations OBJ, IR, IC and VAL:

$$\max_{\alpha,\beta,\mathcal{D}_R} E[a - \beta\mathbb{I}\{a > 0\}|(\mu,\sigma) \in D_R] - \alpha$$

s.t.

$$\alpha + u^* \geq 0 \tag{IR}$$

$$c = \int_{u \geq u^*} (1 - M(u))du \tag{IC}$$

$$\mathcal{D}_R = \{\mu, \sigma | \mu/\sigma \geq \Phi^{-1}\left(\frac{u^*}{\beta}\right)\} \tag{REGION}$$

First we prove the IR constraint must bind. Suppose it does not. Then the firm could lower α by ϵ and increase maximized profit without violating any other constraints. This contradicts optimality. Thus IR binds at the optimum. From the end of the proof of Lemma 2, we have that:

$$u^* = E[u|u \geq u^*] - \frac{c}{Pr(u \geq u^*)}$$

Plugging this into binding IR and solving for α:

$$\alpha = -E[u|u \geq u^*] + \frac{c}{Pr(u \geq u^*)}$$

Substituting the result into the objective obtains:

$$\max_{\beta,\mathcal{D}_R} E[a|(\mu,\sigma)\in D_R] - \frac{c}{Pr((\mu,\sigma)\in D_R}$$

which is the desired form of the objective. Using Lemma 2, the modified problem becomes:

$$\max_{\beta,u^*} E[a|\mu/\sigma \geq \Phi^{-1}(u^*/\beta)] - \frac{c}{Pr(\mu/\sigma \geq \Phi^{-1}(u^*/\beta))}$$

$$c = \int_{u\geq u^*} (1-M(u))du \tag{IC}$$

This makes apparent that the objective is no longer constrained by the constraints (since we have an extra degree of freedom), and in fact only depends on $x := \Phi^{-1}(u^*/\beta)$. Thus we can maximize the objective without constraints to derive x, then use the definition of x and the IC constraint to derive β, u^*. Finally, α can be retrieved from the binding IR constraint. Thus the problem reduces in the way stated in the proposition.■

8.4 Proof of Proposition 1

Proof. Note that under independence, $\sigma|\mathcal{D}_F$ is the same as the unconditional distribution of σ. Then:

$$\begin{aligned}
Pr(\sigma \leq y|(\mu,\sigma)\in\mathcal{D}_R) &= Pr(\mu \leq y\tilde{\mu}^*|(\mu,\sigma)\in\mathcal{D}_R) * Pr(\sigma \leq y|\mu \leq y\tilde{\mu}^* \text{ \& } (\mu,\sigma)\in\mathcal{D}_R) \\
&+ Pr(\mu > y\tilde{\mu}^*|(\mu,\sigma)\in\mathcal{D}_R) * Pr(\sigma \leq y|\mu > y\tilde{\mu}^* \text{ \& } (\mu,\sigma)\in\mathcal{D}_R) \\
&= Pr(\mu \leq y\tilde{\mu}^*|(\mu,\sigma)\in\mathcal{D}_R) * 1 + (1 - Pr(\mu \leq y\tilde{\mu}^*|(\mu,\sigma)\in\mathcal{D}_R)) * G_\sigma(y) \\
&> G_\sigma(y) = Pr(\sigma \leq y|(\mu,\sigma)\in\mathcal{D}_F)
\end{aligned}$$

Notice the first quantity is the conditional CDF in the recruiter acceptance region. The second to last line shows that the this CDF is essentially a weighted average of 1 and $G_\sigma(y)$ which is always weakly greater than $G_\sigma(y)$. This proves first-order stochastic dominance of σ by $\sigma|\mathcal{D}_F$. ■

8.5 Joint Pareto

From Lemma 1, we know the general form of the acceptance region, what remains is to find μ^*. The equation characterizing μ^* from Lemma 1 can be re-written as:

$$c = \int_{\mu^*}^{\infty} 1 - F_\mu(x)dx = \int_{\mu^*}^{\infty} \left(\frac{x_\mu}{x}\right)^{\theta_\mu} dx$$

Integration and solving for μ^* yields the result. Note the first-best solution does not depend on the distribution of σ.

Suppose two variables, X, Y are distributed independently Pareto. That is, suppose their joint distribution is:

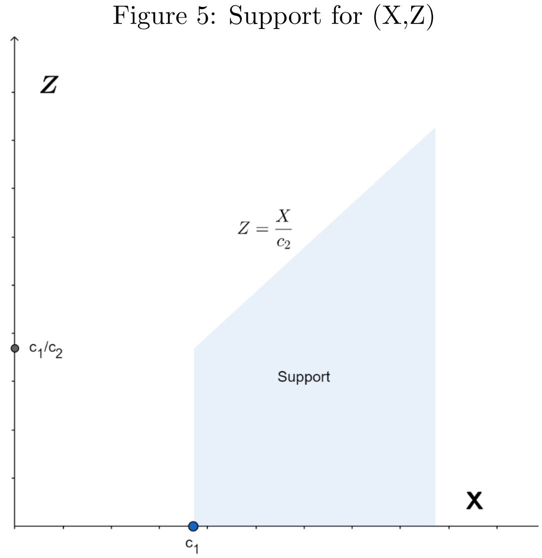
$$f(x,y) = \frac{\alpha_1 c_1^{\alpha_1}}{x^{\alpha_1+1}} \frac{\alpha_2 c_2^{\alpha_2}}{y^{\alpha_2+1}} \mathbb{I}\{x \geq c_1\}\mathbb{I}\{y \geq c_2\}$$

Consider the joint distribution of $X, Z := X/Y$. By the transformation theorem, this is given by:

$$g(x,z) = f(x, x/z) \cdot \frac{x}{z^2}$$

$$g(x,z) = \frac{\alpha_1 \alpha_2 c_1^{\alpha_1} c_2^{\alpha_2}}{x^{\alpha_1+\alpha_2+1}} z^{-1+\alpha_2} \mathbb{I}\{x \geq c_1\}\mathbb{I}\{x/z \geq c_2\}$$

Figure 5: Support for (X,Z)



Now we derive the marginal distribution of Z. Consider first when $z \leq c_1/c_2$. Then the first indicator implies the second is satisfied, and we can get the marginal:

$$g(z) = \int_{c_1}^{\infty} g(x,z)dx = \frac{\alpha_1 \alpha_2}{(\alpha_2 + \alpha_1)} z^{-1+\alpha_2} \left(\frac{c_2}{c_1}\right)^{\alpha_2}$$

In the other case, the second indicator impleis the first, so:

$$g(z) = \int_{c_2 z}^{\infty} g(x,z)dx = \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} z^{-1-\alpha_1} \left(\frac{c_1}{c_2}\right)^{\alpha_1}$$

Now we get the marginal CDF by cases:

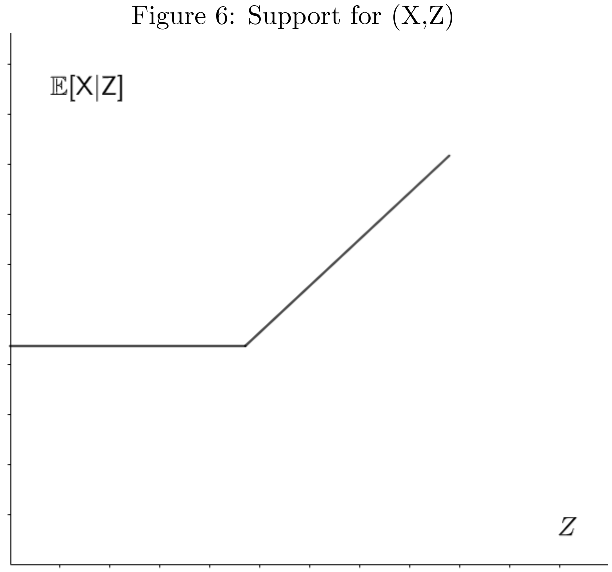
$$G(z) = \begin{cases} \frac{\alpha_1}{\alpha_1+\alpha_2}\left(\frac{c_2}{c_1}\right)^{\alpha_2} z^{\alpha_2} \text{ if } z \leq c_1/c_2 \\ 1 - \frac{\alpha_2}{\alpha_1+\alpha_2} z^{-\alpha_1}\left(\frac{c_1}{c_2}\right)^{\alpha_1} \text{ else} \end{cases}$$

The conditional distribution:

$$g(x|z) = \frac{g(x,z)}{g(z)} = \begin{cases} \frac{c_1^{\alpha_1+\alpha_2}(\alpha_1+\alpha_2)}{x^{\alpha_1+\alpha_2+1}} \mathbb{I}\{x \geq c_1\} \text{ if } z \leq c_1/c_2 \\ \frac{(c_2 z)^{\alpha_1+\alpha_2}(\alpha_1+\alpha_2)}{x^{\alpha_1+\alpha_2+1}} \mathbb{I}\{x \geq c_2 z\} \text{ else} \end{cases}$$

$$E[X|Z=z] = \begin{cases} \frac{(\alpha_1+\alpha_2)}{(\alpha_1+\alpha_2-1)} c_1 \text{ if } z \leq c_1/c_2 \\ \frac{(\alpha_1+\alpha_2)}{(\alpha_1+\alpha_2-1)} c_2 z \text{ else} \end{cases}$$

Figure 6: Support for (X,Z)



For $z > c_1/c_2$

$$\mathbb{E}[X|Z > z] = \frac{\alpha_1+\alpha_2}{\alpha_1+\alpha_2-1} \cdot \frac{\alpha_1}{\alpha_1-1} c_2 z$$

$$\mathbb{E}[X|Z > z] - \mathbb{E}[X|Z = z] = \frac{\alpha_1+\alpha_2}{\alpha_1+\alpha_2-1} \cdot \frac{1}{\alpha_1-1} c_2 z$$

Thus the First Order Condition determining SB search threshold z^* – $([\mu|\tilde{\mu} > z^*] - [\mu|\tilde{\mu} = z^*]) * \Pr(\tilde{\mu} > z^*) = c$ – for independently Pareto distributed μ and σ with parameters (c_1, α_1) and (c_1, α_1), can be rewritten as

$$\frac{\alpha_1+\alpha_2}{\alpha_1+\alpha_2-1}\cdot\frac{1}{\alpha_1-1}c_2z^*\cdot\frac{\alpha_2}{\alpha_1+\alpha_2}{z^*}^{-\alpha_1}\left(\frac{c_1}{c_2}\right)^{\alpha_1}=c$$

or

$$\frac{\alpha_2}{(\alpha_1+\alpha_2-1)(\alpha_1-1)}\cdot\frac{c_1^{\alpha_1}}{c_2^{\alpha_1-1}}\cdot\frac{1}{c}={z^*}^{\alpha_1-1}$$

$$z^*=\left(\frac{\alpha_2}{(\alpha_1+\alpha_2-1)(\alpha_1-1)}\right)^{\dfrac{1}{\alpha_1-1}}\cdot\frac{(c_1^{\alpha_1}/c)^{\dfrac{1}{\alpha_1-1}}}{c_2}$$

Re-arrange:

$$z^*=\left(\frac{c_1^{\alpha_1}\alpha_2}{c(\alpha_1+\alpha_2-1)(\alpha_1-1)}\right)^{\dfrac{1}{\alpha_1-1}}\cdot\frac{1}{c_2}$$

is increasing in c_1,α_2 and decreasing in c_2,c (and probably increasing in α_1 - not clear).

Note that the firm will select $z\geq c_1/c_2$ if and only if:

$$\frac{c_1\alpha_2}{(\alpha_1+\alpha_2-1)(\alpha_1-1)}\geq c$$

That is, as long as costs are not too large.

Note $n+1$: for $z=c_1/c_2$

$$\frac{\alpha_1+\alpha_2}{\alpha_1+\alpha_2-1}\cdot\frac{1}{\alpha_1-1}\cdot\frac{\alpha_2}{\alpha_1+\alpha_2}<\frac{c}{\alpha_1c_1/(\alpha_1-1)}*\frac{\alpha_1}{\alpha_1-1}$$

$$\frac{\alpha_2}{(\alpha_1+\alpha_2-1)\alpha_1}<\frac{c}{\mathbb{E}[\mu]}<1$$

– possible. Thus the other FOC is for $z<c_1/c_2$ is as follows

$$((p*\mathbb{E}[\mu|\tilde{\mu}=c_1/c_2]+(1-p)*\mathbb{E}[\mu|\tilde{\mu}>c_1/c_2])-\mathbb{E}[\mu|\tilde{\mu}=c_1/c_2])*Pr(\tilde{\mu}>z^*)=c$$

Where $p=Pr(\tilde{\mu}<c_1/c_2|\tilde{\mu}>z^*)\;\Rightarrow\;(1-p)*Pr(\tilde{\mu}<c_1/c_2|\tilde{\mu}>z^*)=\dfrac{\alpha_2}{\alpha_1+\alpha_2}$. Thus the FOC is equivalent to

$$(1-p)*Pr(\tilde{\mu}<c_1/c_2|\tilde{\mu}>z^*)*(\mathbb{E}[\mu|\tilde{\mu}>c_1/c_2]-\mathbb{E}[\mu|\tilde{\mu}=c_1/c_2])=\frac{\alpha_2}{\alpha_1+\alpha_2}\cdot\frac{\alpha_1+\alpha_2}{\alpha_1+\alpha_2-1}\cdot\frac{1}{\alpha_1-1}*c_1=c$$

$$\frac{c_1\alpha_2}{(\alpha_1+\alpha_2-1)(\alpha_1-1)}=c$$

8.6 Proof of Equalizing Search Costs

Suppose a principal can either delegate search to an agent who performs search at unit cost c_A or perform search themselves at search cost c_p. We wish to find the ratio c_p/c_A where the principal is indifferent between delegation and direct search. That is, we wish to solve for search costs where profit in the first and second best is equalized. To do this, note that:

$$\frac{\Pi_{SB}(c_A)}{\Pi_{FB}(c_P)}=\frac{\frac{(\alpha_1+\alpha_2)}{(\alpha_1+\alpha_2-1)}\left(\frac{c_1^{\alpha_1}\alpha_2}{c_A(\alpha_1+\alpha_2-1)(\alpha_1-1)}\right)^{\frac{1}{\alpha_1-1}}}{\left(\frac{c_1^{\alpha_1}}{c_P(\alpha_1-1)}\right)^{\frac{1}{\alpha_1-1}}}=\left[\frac{c_P}{c_A}\right]^{\frac{1}{\alpha_1-1}}(\alpha_1+\alpha_2)\left(\frac{\alpha_2}{(\alpha_1+\alpha_2-1)^{\alpha_1}}\right)^{\frac{1}{\alpha_1-1}}$$

We wish to solve for the ratio of costs where this is equal to 1:

$$\left[\frac{c_P}{c_A}\right]^{\frac{1}{\alpha_1-1}}(\alpha_1+\alpha_2)\left(\frac{\alpha_2}{(\alpha_1+\alpha_2-1)^{\alpha_1}}\right)^{\frac{1}{\alpha_1-1}}=1$$

$$(\alpha_1+\alpha_2)^{\alpha_1-1}\frac{\alpha_2}{(\alpha_1+\alpha_2-1)^{\alpha_1}}=\frac{c_A}{c_P}$$

Call this ratio c^*. If $\frac{c_A}{c_P}<c^*$ then the principal will choose to delegate. If not, the principal will conduct search itself. Then we wish to compute comparative statics in c^*. Note that it does not depend on c_1, c_2. Taking logs first we have that:

$$log(c^*)=(\alpha_1-1)log(\alpha_1+\alpha_2)+log(\alpha_2)-\alpha_1 log(\alpha_1+\alpha_2-1)$$

Taking derivatives:

$$\frac{\partial log(c^*)}{\partial\alpha_2}=\frac{\alpha_1-1}{\alpha_1+\alpha_2}+\frac{1}{\alpha_2}-\frac{\alpha_1}{\alpha_1+\alpha_2-1}$$

We want to know the sign of this derivative. All denominators are positive, so we will combine like terms:

$$\frac{\partial log(c^*)}{\partial \alpha_2} = \frac{\alpha_1^2 + \alpha_1\alpha_2 - \alpha_1 + \alpha_1\alpha_2 + \alpha_2^2 - \alpha_2 + \alpha_1^2\alpha_2 + \alpha_1\alpha_2^2 - \alpha_1\alpha_2 - \alpha_1\alpha_2 - \alpha_2^2 + \alpha_2 - \alpha_1^2\alpha_2 - \alpha_1\alpha_2^2}{(\alpha_1 + \alpha_2)\alpha_2(\alpha_1 + \alpha_2 - 1)}$$

which reduces to:

$$\frac{\partial log(c^*)}{\partial \alpha_2} = \frac{\alpha_1^2 - \alpha_1}{(\alpha_1 + \alpha_2)\alpha_2(\alpha_1 + \alpha_2 - 1)}$$

Note that as long as $\alpha_1 > 1$ this is positive. This is also the condition required for the marginal distribution to have a finite moment. Thus c^* is increasing in α_2, meaning that as the distribution of σ shifts to the left (collapsing towards a degenerate distribution) delegated search becomes optimal in a wider range of circumstances.

Now for α_1. Taking derivatives:

$$\frac{\partial log(c^*)}{\partial \alpha_1} = log(\alpha_1 + \alpha_2) - log(\alpha_1 + \alpha_2 - 1) + \frac{\alpha_1 - 1}{\alpha_1 + \alpha_2} - \frac{\alpha_1}{\alpha_1 + \alpha_2 - 1}$$

When we have the most uncertainty, $\alpha_2 = 1$, then:

$$\frac{\partial log(c^*)}{\partial \alpha_1} = log(\alpha_1^{-1} + 1) + \frac{\alpha_1 - 1}{\alpha_1 + 1} - 1$$

This is always negative when $\alpha_1 > 1$, so as the distribution of μ shifts to the left delegated search is optimal in a smaller range of circumstances.

When we have that the distribution of uncertainty is degenerate

8.7 Proof of Relative Agency Loss

using the work from the last section but setting costs equal:

$$RAL = 1 - \frac{\Pi_{SB}}{\Pi_{FB}} = 1 - (\alpha_1 + \alpha_2)\left(\frac{\alpha_2}{(\alpha_1 + \alpha_2 - 1)^{\alpha_1}}\right)^{\dfrac{1}{\alpha_1 - 1}}$$

$$\lim_{\alpha_2 \to \infty} \tilde{\mu}^*(\alpha_2) = \frac{1}{c_2}\mu^* \lim_{\alpha_2 \to \infty} \left(\frac{\alpha_2}{\alpha_1 + \alpha_2 - 1}\right)^{\frac{1}{\alpha_1 - 1}} = \frac{1}{c_2}\mu^* \left(\lim_{\alpha_2 \to \infty} \frac{\alpha_2}{\alpha_1 + \alpha_2 - 1}\right)^{\frac{1}{\alpha_1 - 1}} = \frac{1}{c_2}\mu^*$$

Since:

$$\Pi_{SB} = E[a|\tilde{\mu} = \tilde{\mu}^*] = \frac{\alpha_1 + \alpha_2}{\alpha_1 + \alpha_2 - 1} c_2 \tilde{\mu}^*$$

Then:

$$\lim_{\alpha_2 \to \infty} \Pi_{SB} = \lim_{\alpha_2 \to \infty} \frac{\alpha_1 + \alpha_2}{\alpha_1 + \alpha_2 - 1} \mu^* = \mu^*$$

Thus in the limit RAL goes to 0.

8.8 Search intensity

Proof of Proposition 3.

$$\mu^* := \left(\frac{x_\mu^{\theta_\mu}}{c(\theta_\mu - 1)} \right)^{\frac{1}{\theta_\mu - 1}}$$

$$Pr_1 \equiv Pr(\mu \geq \mu^*) = \frac{x_\mu^{\theta_\mu}}{\mu^{*\theta_\mu}}$$

$$\tilde{\mu}^* = \frac{1}{x_\sigma} \left(\frac{x_\mu^{\theta_\mu} \theta_\sigma}{c(\theta_\mu + \theta_\sigma - 1)(\theta_\mu - 1)} \right)^{\frac{1}{\theta_\mu - 1}}$$

$$Pr_2 = Pr(\tilde{\mu} \geq \tilde{\mu}^*) = \frac{\theta_\sigma}{\theta_\mu + \theta_\sigma} \cdot \tilde{\mu}^{*-\theta_\mu} \left(\frac{x_\mu}{x_\sigma} \right)^{\theta_\mu} = Pr_1 \cdot \frac{\theta_\sigma}{\theta_\mu + \theta_\sigma} \cdot \left(\frac{\theta_\mu + \theta_\sigma - 1}{\theta_\sigma} \right)^{\frac{\theta_\mu}{\theta_\mu - 1}}$$

$$\frac{Pr_2}{Pr_1} = \frac{\theta_\sigma}{\theta_\mu + \theta_\sigma} \cdot \left(\frac{\theta_\mu + \theta_\sigma - 1}{\theta_\sigma} \right)^{\frac{\theta_\mu}{\theta_\mu - 1}}$$

$$\frac{\partial \log(Pr_2/Pr_1)}{\partial \theta_\sigma} = -\frac{\theta_\mu}{(\theta_\mu + \theta_\sigma - 1)(\theta_\mu + \theta_\sigma)\theta_\sigma} < 0$$

Pr_1 – the first best acceptance region probability does not depend on the distribution of σ. The ratio Pr_2/Pr_1 is decreasing in θ_σ. Also

$$\lim_{\theta_\sigma \to \infty} \frac{Pr_2}{Pr_1} = 1$$

Thus we can also conclude that Pr_2 is always larger than Pr_1 (which ends the proof of proposition 3).

8.9 First and Second Best Acceptance Regions Comparison

Proof of Proposition 4. We will use the notations from Figure 2.

$$p = \frac{Pr(A)}{Pr(A) + \Pr(B)}$$

$$q = \frac{Pr(C)}{Pr(C) + \Pr(B)}$$

From proposition 3, we can conclude that $p > q$ since $Pr(\text{FB}) < Pr(\text{SB})$.

$$\mu|\text{SB} \sim (\mu|A)p(\mu|B)$$

$$\mu|\text{FB} \sim (\mu|C)q(\mu|B)$$

where that notations on the RHS are used for mixture distribution. In other words, one could right each of them as a three-component mixture:

$$\mu|\text{SB} \sim (\mu|A)(\text{w/p } q) + (\mu|A)(\text{w/p } p - q) + (\mu|B)(\text{w/p } 1 - p)$$

$$\mu|\text{FB} \sim (\mu|C)(\text{w/p } q) + (\mu|B)(\text{w/p } p - q) + (\mu|B)(\text{w/p } 1 - p)$$

Given the support of $\mu|A,\ \mu|B,\ \mu|C$, it is trivial to conclude that $\mu|B \succ_{\text{FOSD}} \mu|A$ and $\mu|C \succ_{\text{FOSD}} \mu|A$. Thus, each of the components in the first best μ mixture first order stochastically dominates the components in the second best μ mixture. Given that the mixture probabilities are identical, that implies that the whole FB mixture dominates the SB mixture

$$\mu|\text{FB} \succ_{\text{FOSD}} \mu|\text{SB}$$

(this simply follows from the formula of a mixture CDF).