

# Costly Job Search with Inattentive Workers

Siting Estee Lu\*

October 22, 2024

[\[CLICK HERE FOR LATEST VERSION\]](#)

## Abstract

Labour market mismatch can arise from workers having limited attention. This paper proposes a Generalized Partially Directed Search model, extending on existing literature by allowing inattentive workers to have diverse priors and heterogeneous attention costs. I show that mismatch can be inherited from bias in workers' default search strategies, and heterogeneous attention costs could contribute to greater variability in the equilibrium outcomes. I also explore equilibrium multiplicity that was not adequately accounted for in previous studies. Equilibrium where workers adopt different application strategies may generate both higher market efficiency and lower monopsony power than when workers employ the same application strategies. This information-theoretic approach to model job search offers new policy insights on the basis of attention.

**JEL:** D83, D91, J64

**Keywords:** Rational Inattention, Partially Directed Search

---

\*School of Economics, The University of Edinburgh. Contact: s.lu@ed.ac.uk.

I am indebted to my supervisors, Ed Hopkins and Axel Gottfries, for their invaluable suggestions and continuous guidance. I would also like to thank Tatiana Kornienko, Michael Woodford, Rava Azeredo da Silveira and audiences at SGPE 2021, QMUL Economics and Finance Workshop, SABE-IAREP Conference 2023, 33rd Advanced School in Economic Theory, FUR Conference 2024, and Cognitive Foundations of Decision-Making 2024 for their comments and suggestions, as well as Scottish Economic Society for its support. All errors are my own.

# 1 Introduction

Workers' imperfect search behaviours can be driven by incomplete information and cognitive biases, leading to congestion of applications at certain firms. While online job search platforms improved information accessibility, assuming workers to be fully attentive to all information is impractical. Whilst Wu (2024) introduced a search model where information availability depends on how much workers choose to acquire and process at a cost, its prediction of broader search behaviour does not align with empirical finding of inattentive workers exhibiting significant bias and focusing on a narrow set of opportunities (Belot et al. (2019)). Stringent assumptions could have limit the model's explanatory power. Like other decision problems, workers' search behaviour is likely influenced by prior knowledge (Erev et al. (2010), Sáenz-Royo et al. (2022)), assuming that they search randomly by default overlooks the implication of ex-ante information. Workers are also commonly assumed to apply analogously as differentiated strategies of applying to different firms may pose potential coordination difficulties (Wright et al. (2021)), but there are no definite evidence that one equilibrium is distinctly harder to achieve than the other, and diversity of strategies may be suppressed in contrast to real world observations. There lacks a generalized information-theoretic framework that fully captures workers' imperfect cognition on labour market outcomes, which could provide richer predictions and better support for policies to tackle information frictions.

In this paper, I propose a Generalized Partially Directed Search (GPDS) model with two firms and two attention-constrained workers. Wages are not directly observed, but workers can learn about the wages at a precision level of their choice before making their application decisions. The model builds on top of existing partially directed search model, put forth by Wu (2024), but additionally accounts for diverse priors and possibility of equilibrium multiplicity. Considering informative priors effectively incorporates a behavioural component that embraces the possibility of workers holding implicit knowledge about wage distribution. This feeds into the model as exogenous, non-random default choice. It affects attention cost, which is characterized by the Kullback–Leibler divergence between the default and chosen strategy, and would have an impact on the labour market outcome. I also allow the possibility of workers adopting different application strategies in addition to identical ones. This non-behavioural refinement accounts for variability arising from intrinsic feature of the model that has been previously neglected.

By encompassing the above factors, I show that workers' imperfect search and the resulting labour market mismatch may be inherited from one's fundamental bias in default strategies, particularly when it is expensive to acquire and process information. I also make a theoretical contribution by proving the existence of multiple equilibrium. I demonstrate that asymmetric strategies can arise when information is cheap to acquire relative to expected payoffs. Their presence can explain for workers' displayed tendency to choose one firm over the other with much higher probability beyond what wage differences alone would predict.

I also explore the less conventional horizontal-differentiation in attention costs through variability in default strategies or cognitive costs. This offers a novel angle in analyzing imperfect search stemming from worker heterogeneity that differs from the usual skills or productivity disparity (Shi (2002), Wright et al. (2021)). I found workers tend to adopt asymmetric strategies in view of different information costs, providing another explanation for the observed narrow search behaviour. Such heterogeneity may also contribute to greater diversity of equilibrium strategies.

The presence of information friction is non-trivial to the matching problem and can significantly impact individual and social welfare. In presence of multiple equilibria, comparing the equilibrium outcomes, I found social welfare to be higher when workers adopt different application strategies than applying analogously. Since wage offer differs across equilibria, there are also variations in monopsony power, which highlights the trade-off between firms' and workers' welfare. This

finding could be interesting for policymakers as some equilibrium may be deemed more “desirable” than others depending on policy objectives, and serves as basis for further exploration of equilibrium selection measures. Nonetheless, I also extend the model to investigate a constrained social planner problem, where the socially optimal strategy can be used as a benchmark to evaluate the efficiency level of the equilibrium outcomes. I found that equilibrium where workers use the same application strategy is generally inefficient. This supports policies to improve efficiency by targeting information frictions, such as manipulating attention through information presentation. However, I also show that neglecting attention heterogeneity could be detrimental to policy-making as it can result in over- or under- estimation of efficiency prevailing in the market. Therefore, there is a need for accurate assessment of the underlying market state for informed and effective policies.

The rest of the paper is structured as follows: Section 2 introduces the GPDS model. Section 3 explores worker heterogeneity in attention costs. Section 4 investigates the social planner problem. Section 5 discusses the applicability, implications, limitations and extensions of the model. I then conclude in Section 6.

*Related literature.* This paper contributes to literature on frictional search models to explain for the prevalence of labour market mismatch. Miscoordination in the labour market is often ascribed to the presence of skills gap, and workers sorting themselves based on skills, leading to overcrowding at some firms, but there is little evidence to support this. Handel (2003) shows that skills demanded by jobs do not exceed workers’ capabilities. While Cappelli (2015) and Albrecht and Vroman (2002) indicate over-education, where workers have higher skills than job requirements, may be possible in causing unemployment for the less educated workers, such pattern is likely to be transitory with on-the-job search (Dolado et al. (2009)), and a causal relationship between over-education and unemployment is not justifiable (Leuven and Oosterbeek (2011)). Not only there is a lack of apparent skills mismatch, Patterson et al. (2016), Şahin et al. (2014) and Belot et al. (2019) also highlight that workers may not have precise knowledge about which jobs offer favourable remunerations and if their skills are transferable to these roles. Therefore, even if skills are perfectly aligned, mismatch can still happen as workers do not know where to sort themselves. This implies mismatch may be driven by a more fundamental force of information availability, which this paper focuses on.

There are limited search models that study the impact of noisy information on labour market outcome, particularly one that is validated by empirical observations. Menzio (2007) mentioned two well-established theoretical models that sit on extreme ends of information availability – McCall (1970)’s random search model, where workers do not have wage information and random application choices may cause mismatch; and directed search model, recapitulated by Wright et al. (2021), where workers perfectly observe wages and one-to-one matching is possible. Wu (2024) proposed a partially directed search model that falls in the intermediary information range. However, it has imposed strict assumptions on workers, resulting in predictions that partially aligns with real world findings (Belot et al. (2019)). For instance, Wu (2024) found inattentive workers search broadly, applying to every job with positive probability. This suggests if attention costs are reduced, one should apply to firms more selectively. However, Belot et al. (2019)’s survey results reveal that workers naturally apply narrowly, and implementing a search tool with job recommendations that essentially manipulates one’s attention cost, would increase their application scope, except for those starting with broad searches. This indicates Wu (2024)’s model may have missed crucial factors in explaining why workers apply narrowly in practice, and the varied effect from search tool intervention. Therefore, my model offers a comprehensive framework that could bridge the theoretical predictions and empirical results.

This limited attention model also contributes to both search and decision-making literature. It strengthens the predictive power of search models through postulating a novel form of search

costs back-boned by strong cognitive foundations, and it also demonstrates individual-level thought processes can have broader implications on macro issues. This study rides on the vast rational inattention literature on decision-making (e.g. Matějka and McKay (2015), Ravid (2020), Caplin and Dean (2015))), as well as cognitive imprecision (e.g. Woodford (2020), Heng et al. (2020)), where workers could have noisy representation of wages in their internal cognitive system even if wages are perfectly visible. Traditional search costs, such as payments to agencies in Gronau (1971), did not explicitly capture the granularity in information gathering, whereas rational inattention approach, as reviewed by Mackowiak et al. (2020), allows for continuous information precision that directly conceptualizes partial information driven by individual decisions. Conventional search cost also focuses more on acquiring previously inaccessible data that can already be tackled by the proliferation of internet (Denzer et al. (2021), Gürtzgen et al. (2018)). However, better external information availability does not imply improved perception of information environment which individuals base their decisions on. Attention cost is also not equivalent to time cost (Gronau (1971)), as more time invested does not automatically signify learning more precise wages. The above distinctions indicate attention cost to be a new form of search cost that captures the complexities of information acquisition and processing, in which workers are constrained by their own ability and willingness to learn. Furthermore, the inclusion of bias in defaults also draws from neuro and behavioural literature, where prior beliefs about wage distribution can be influenced by past experiences (Erev et al. (2010)) or consolidated knowledge (Sáenz-Royo et al. (2022)). By establishing a link between information-based mismatch and imperfect search behaviour driven by individual decision-making mechanisms, I showcase the viability of tackling macro problem from a micro perspective.

Despite inattention being a pervasive problem with substantial implications on firm-worker matching outcomes, its application in this area remain limited. Cheremukhin et al. (2020) is one of the first to incorporate limited attention in matching, but in the marriage market, where agents pay a cost to locate potential partners. It models observed aspirational dating patterns. The GPDS model is similar in that it also depicts biased search behaviour, but efficiency analysis is centered more on vacancy filling than matching quality as in the marriage market. In the labour market context, studies tend to focus on inattentive firms and assume fully attentive workers. Acharya and Wee (2020) explores firms' selective hiring decisions when they might not be able to perfectly identify suitable candidates, and Habermalz (2014) investigates firms' allocation of attention across current employees who have heterogeneous impact on profits. However, information barrier exists for workers as well. Current literature that studied inattentive workers, such as Briggs et al. (2017), looks at workers' decisions when they hold job offer of unknown type and have to learn about non-wage information, such as overtime pay and work flexibility; and Giezek (2019) explores workers' learning about the probabilities of rejecting a good job and accepting a bad one before making a decision about the offer. Even though they both examine inattention on workers' side, the focus is on their decisions after receiving an offer. The job application stage is equally important and inattentiveness could have a more drastic impact given the substantial larger amount of information to process. Mismatch due to a lack of coordination in application could directly contribute to unemployment and unfilled vacancies, therefore, it is a more pertinent issue that is addressed by this paper.

## 2 The Generalized Partially Directed Search Model (GPDS)

This paper adopts Wu (2024)'s  $2 \times 2$  model with two workers,  $i = 1, 2$ , and two firms,  $j = 1, 2$ . Each firm offers one vacancy.

*Timing and Set-up.* Firms and workers act sequentially. Firms simultaneously set their wages first, followed by workers concurrently select their application strategies.

*Stage 1:* Firms observe realization of productivity,  $\mathbf{z}$ , drawn from an exogenous distribution  $Z$ . Firm  $j$ 's productivity is denoted as  $z_j$ , such that  $\mathbf{z} = (z_1, z_2)$ ,  $\mathbf{z} \in Z$ ,  $Z < \infty$ . In anticipation of workers' application strategies, firms choose the wages to be offered,  $\mathbf{w} = (w_1, w_2)$ ,  $\mathbf{w} \in W$ ,  $0 \leq W \leq Z$ .

*Stage 2:* Workers cannot observe wages or productivities directly. They can pay an attention cost to learn about the wages at a precision level of their choice. This process involves them acquiring a signal,  $\mathbf{s}$ ,  $\mathbf{s} \in \mathbb{R}^J$ , drawn from a distribution conditional on the actual wages. Higher cost is associated with greater precision of the observation. Workers select an application strategy upon observing the signal.

*Stage 3:* Matching and hiring take place. If a match is successfully formed between firm  $j$  and worker  $i$ , then firm  $j$  receives a payoff of  $(z_j - w_j)$  and worker  $i$  receives a payoff of  $w_j$ . If not matched, both parties receive their outside option of 0. If both workers apply to the same firm, one of them will be randomly selected and hired. It is assumed that when wages are the same as the outside option, workers accept the offer.

Both parties understand the game, the information availability and the optimization problem. Non-wage benefits are not considered in this model as they can be subjective, and there are no well-defined space where complete information of which is accessible.

### 2.1 Workers' and Firms' Maximization Problem

The main narrative adopted by this paper follows Wu (2024) closely, where workers' attention costs are perceived as *information foraging and processing costs*. (see Appendix B.1) Workers would enter the job search process holding some prior beliefs about the wage distribution, they then observe two job postings without any wage information. They undergo a two-stage decision process: In the first stage, since wages are not directly ascertainable, they devise an attention strategy, which determines how much time and energy to invest in gathering further information at a cost. This could involve searching on other websites or connecting with current employees of the companies, which helps to narrow down the wage range. In the second stage, workers refine their prior beliefs about the wage distribution with the acquired information and choose an application strategy given their beliefs about the opponent's strategy.

An alternative narrative is to interpret workers' attention costs as *cognitive efforts*. Woodford (2020) indicates that evidence from psychophysics and neurophysiology have shown that in one's cognitive system, at least a small degree of randomness is present in the internal representation of quantities observed. Even if wage information are perfectly visible on the job postings, Heng et al. (2020) highlights that workers could still have noisy representation of these information in the nervous system due to limited neural resources. Therefore, there will always exist some information imprecision. Azeredo da Silveira et al. (2024) also suggests that imprecise information could arise from noisy retrieval of memories. When making application decisions, one can be recalling from their memories about wage information they just saw, and the more cognitive effort one exert, the more precisely one can retrieve these information.

In all cases, firms, knowing workers could behave with some noise, be it from not acquiring precise information, or having imperfect internal representation or noisy recall of wage information, could exploit this constraint by extracting higher wage markdowns.

**Workers' Search Problem with Information Acquisition.** The key component to the GPDS model is workers' attention cost.<sup>1</sup> I characterize attention cost for worker  $i$ :

- (*Initialization*) Worker  $i$  form prior beliefs over the wage distribution, defined by  $G^i(\mathbf{w})$ , where  $G^i(\mathbf{w}) \in \Delta W$ , and  $\Delta W$  is the set of Borel probability distributions on  $W$ .
- (*Decision Problem*) He/she choose an information structure,  $F^i(\mathbf{s}|\mathbf{w})$ , and a search strategy given signals,  $q_j^i(\mathbf{s})$ .
- (*Cost Modelling*) The cost of acquiring and processing wage information is modelled using Shannon's entropy, adopted widely in the rational inattention literature, such as Sims (2003) and Matějka and McKay (2015). It is measured using Kullback-Leibler (KL) divergence between prior beliefs,  $G^i(\mathbf{w})$ , and the posterior distribution,  $F^i(\mathbf{w}|\mathbf{s})$ , which pertains to the updated beliefs about wages given the signals acquired.

$$D(F^i||G^i) = \lambda_i(\mathbb{H}(G) - \mathbb{E}_s\mathbb{H}(F(\mathbf{w}|\mathbf{s}))), \text{ where } D(F^i||G^i) \geq 0 \quad (1)$$

$\lambda_i$  is a scalar factor;  $\mathbb{H}(\cdot)$  represents the entropy, and information content is measured in natural log, following Matějka and McKay (2015). Murphy (2012) indicates that higher entropy implies greater uncertainty. It is maximized when there is a uniform distribution over possible wages, and minimized when a wage is almost certain to be realized.

Wu (2024) has defined a Symmetric Perfect Bayesian Equilibrium (SPBE) with information acquisition, assuming symmetry in workers' strategy. I relax this assumption and provide a generalized fixed-point problem in Definition 2.1.

**Definition 2.1** (Signal-Based Optimization Problem, Wu (2024)). *Perfect Bayesian Equilibrium (PBE) is a tuple  $\{G^e(\mathbf{w})^i, F^e(\mathbf{s}|\mathbf{w})^i, \{q_j^e(\mathbf{s})^i\}_j, G^e(\mathbf{w})^{-i}, F^e(\mathbf{s}|\mathbf{w})^{-i}, \{q_j^e(\mathbf{s})^{-i}\}_j, \{w_j^i(z|\mathbf{z})\}_j\}$ ,*

1. (*Optimal Wage*)  $\{w_j^i(z|\mathbf{z})\}_j$  maximizes firm's profit given workers' equilibrium information structure,  $\{F^e(\mathbf{s}|\mathbf{w})^i, F^e(\mathbf{s}|\mathbf{w})^{-i}\}$ , and strategies,  $\{\{q_j^e(\mathbf{s})^i\}_j, \{q_j^e(\mathbf{s})^{-i}\}_j\}$ , and productivity of all firms,  $\mathbf{z}$ ;
2. (*Optimal Search*)  $\{F^e(\mathbf{s}|\mathbf{w})^i, F^e(\mathbf{s}|\mathbf{w})^{-i}\}$  and  $\{\{q_j^e(\mathbf{s})^i\}_j, \{q_j^e(\mathbf{s})^{-i}\}_j\}$  maximize workers' payoff given equilibrium beliefs about the wage distribution,  $\{G^e(\mathbf{w})^i, G^e(\mathbf{w})^{-i}\}$ .
3. (*Full Consistency*)  $\{G^e(\mathbf{w})^i, G^e(\mathbf{w})^{-i}\}$  need to be consistent with the productivity distribution  $G(\mathbf{z})$  and wages  $\{w_j^i(z|\mathbf{z})\}_j$  on the equilibrium path.

Workers' solution comprises of a pair of strategies  $\{F^e(\mathbf{s}|\mathbf{w})^i, \{q_j^e(\mathbf{s})^i\}_j\}, \{F^e(\mathbf{s}|\mathbf{w})^{-i}, \{q_j^e(\mathbf{s})^{-i}\}_j\}$  that solve the fixed-point problem. The signal-based optimization problem faced by worker  $i$  is

$$\begin{aligned} \max_{\{F^e(\mathbf{s}|\mathbf{w})^i, \{q_j^e(\mathbf{s})^i\}_j\}} & \int_{\mathbf{w}} \int_{\mathbf{s}} [q_1^i(\mathbf{s})(1 - q_1^{-i}(\mathbf{s}) + \frac{q_1^{-i}(\mathbf{s})}{2})w_1 \\ & + q_2^i(\mathbf{s})(1 - q_2^{-i}(\mathbf{s}) + \frac{q_2^{-i}(\mathbf{s})}{2})w_2] F^i(d\mathbf{s}|\mathbf{w}) G^{-i}(d\mathbf{w}) - \lambda_i[\mathbb{H}G^{-i}(\mathbf{w}) - \mathbb{E}_s\mathbb{H}(F_{\mathbf{w}|\mathbf{s}}^i)] \end{aligned} \quad (2)$$

<sup>1</sup>A range of literature on entropy and rational inattention were studied in determination of the relevant notations. (Fosgerau et al. (2020))

$$s.t. \sum_j q_j^i(\mathbf{s}) = 1, \int_{\mathbf{s}} F^i(d\mathbf{s}|\mathbf{w}) = 1$$

*More details in Appendix A.1.*

*Moving to a simplified framework.* Wu (2024) has drawn an equivalence between the signal-based optimization problem and costly directed search with observable wage, making the problem more tractable. In the signal-based problem, precision is explicit in the choice of information structure. In the reformed problem, precision is implicit, workers behave as if they observe the true wage, but incur an information cost that correspond to the amount they would have paid to gain a certain precision of wages. Given workers' choice of information structures and strategies based on signals, the conditional probability of applying to  $j$  is  $q_j^i(\mathbf{w}) = \int_{\mathbf{s}} q_j^i(\mathbf{s}) F^i(\mathbf{s}|\mathbf{w}) d\mathbf{s}$ . It accounts for all possible signals that could arise given true wages and how workers would respond to each signal. It represents the outcome of a two-step process, combining information acquisition and search decision. The reformulation treats wages as if they are fully revealed, and the imprecision of wage information is captured in the “randomness” of  $q_j^i(\mathbf{w})$ . If signal received is precise, then  $q_j^i(\mathbf{w})$  can approach 1 or 0.  $q_j^i(\mathbf{w})$  can also be viewed as the conditional probability of signal directly recommending worker  $i$  to apply to firm  $j$  given wage realization, and thus can be denoted as the recommendation strategy.

Based on Ravid (2020)'s refinement highlighted in Wu (2024), workers' equilibrium strategies are well-defined and optimal for all possible wage profiles, including those off the equilibrium path for some perturbation in firms' wage posting strategies. It is also established that for any  $F^i(\mathbf{s}|\mathbf{w})$  and  $q_j^i(\mathbf{s})$ , there exists a  $q_j^i(\mathbf{w})$  that generates the same expected payoffs and information cost, and vice versa. Therefore, the outcome achieved is equivalent between workers incurring a cost to acquire a signal of certain precision and actually observing the wages after paying an optimal information cost, and subgame perfect equilibrium (SPE) can be used in place of PBE.

Through this transformation, the reformed model is able to circumvent the need to explicitly specify the beliefs and updating process, thus steer clear of potential misalignment problem between the beliefs and the true environment, which is a core challenge under PBE formulation highlighted by Steiner et al. (2017). For example, if workers possess incorrect initial beliefs and have biased information acquisition, they could undergo suboptimal updating of beliefs, leading to off the equilibrium path behaviours and systematic deviation from optimal search. The costly directed search framework, on the other hand, avoids reliance on prior beliefs for equilibrium concepts, instead focuses on alignment of workers' actions with the actual wage distribution in the equilibrium, thereby capturing the optimal information acquired even if beliefs are incorrect, and encapsulating the essence of how information impact decisions without all the complexities.

Correspondingly, the attention cost can be reformulated as a measure of mutual information between  $q_j^i(\mathbf{w})$  and the baseline probability of selecting firm  $j$ ,  $p_j^i \equiv \sum_{\mathbf{w}} q_j^i(\mathbf{w}) G^i(d\mathbf{w})$ . (Matějka and McKay (2015))

$$\hat{c}_i = \lambda_i \int_{\mathbf{w}} \sum_j q_j^i(\mathbf{w}) \log\left(\frac{q_j^i(\mathbf{w})}{p_j^i}\right), \sum_j q_j^i(\mathbf{w}) = 1 \quad (3)$$

This cost function resembles a control problem, as per Mattsson and Weibull (2002) and Steiner et al. (2017), where workers have noisy strategies, and incur a control cost to stabilize the “trembling hand”. The axiomatic derivation of a control cost is equivalent to the relative entropy between the chosen and default strategies, which establishes a link between a logit choice model and entropy from information theory. Information cost in this model can thus be perceived as how much effort workers expend on refining their strategies away from default.

*Components of attention cost.* One of the core objectives of my generalized model is to include more variations in attention costs to provide a richer and more realistic analysis of workers'

behaviour under limited attention. There are two essential components to (3): default strategies and cognitive costs.

In Wu (2024)’s model, firms’ identities are uninformative, therefore, workers are assumed to use a default random search strategy due to insufficient reasoning (Dubs (1942)). However, Banner et al. (2020) highlights that a universal adoption of non-informative priors is not sufficiently justified. Furthermore, assuming random search strategy by default overlooks the possibility of workers entering job search having past experiences of working with similar firms or at identical positions, or having some impressions about wages that correspond to certain industrial norms. In a broader sense, workers could possess prior beliefs about firms’ “goodwill” or reputation, and they would expect higher wages from firms with better “goodwill”. As a result, firms’ identities do reveal some wage information, and workers’ default search behaviour can be less random. This argument is supported by Dietrich and List (2013), who proposed that doxastic reasons, such as prior knowledge or subjective judgements could justify holding onto certain prior beliefs, leading to non-random choice of options. Also mentioned in Faghihi et al. (2015), the dual cognitive process system brought forth by Kahneman (2011) indicates that individuals have a compartment specializes in fast, intuitive thinking, and another for slower, analytical reasoning. Choices made based on prior beliefs could come from the “rapid decision-making system”, influenced by previous experiences (Erev et al. (2010)) and consolidated knowledge (Sáenz-Royo et al. (2022)). Instead of having non-informative priors, workers may have weakly informative priors, their formation can either be rational – basing on past observations, or behavioural – corresponding to subjective judgements, not accounting for this possibility could leave out valuable information about matching outcomes. However, a potential challenge of adopting informative priors is the selection of prior distribution, which might substantially affect the generalizability and validity of the equilibrium outcome. The GPDS model does not face such challenge since it uses default choice probability ( $p_j$ ) as an exogeneously determined parameter in this set-up, any bias in prior beliefs are implicit in this strategy. In addition, in contrast to the difficulty of estimating priors based on available data, default probabilities can be easily approximated. This affords the model better prediction capability given that the underlying market conditions can be more accurately captured.

The cost function (3) also includes a scalar factor,  $\lambda_i \in \Lambda$ . This is an exogenous parameter that depicts the cognitive cost or unit cost of processing information. It can be interpreted as workers’ cognitive ability, and also as an estimate of the degree of complexity or difficulty in accessing and processing wage information. For example, if firms impose pay secrecy clauses in employment contracts to prevent current employees from disclosing wage information, this could make the process of learning more difficult, thereby increasing the unit information cost. Higher cost could also be attributed to complexity in wage structuring. Wages may consist of many components such as basic rate, stock options and performance pegged pay, thus it can be harder to decipher the package value while factoring out the non-monetary benefits. All of these could contribute to higher  $\lambda_i$ .

**Workers’ Costly Directed Search Problem.** Following Matějka and McKay (2015) and Wu (2024), worker  $i$ ’s maximization problem, given worker  $-i$ ’s strategy and wages, is characterized as a costly directed search problem.

$$\begin{aligned} \max_{q_1^i, q_2^i \in [0,1]} & [q_1^i(1 - q_1^{-i} + \frac{q_1^{-i}}{2})w_1 + q_2^i(1 - q_2^{-i} + \frac{q_2^{-i}}{2})w_2 - \lambda_i(q_1^i \log \frac{q_1^i}{p_1^i} + q_2^i \log \frac{q_2^i}{p_2^i})] \\ \text{s.t. } & q_1^i + q_2^i = 1, \quad q_1^{-i} + q_2^{-i} = 1, w_1 \geq 0, w_2 \geq 0 \end{aligned} \quad (4)$$

$(p_1^i, p_2^i)$  and  $(p_1^{-i}, p_2^{-i})$  represent the collections of worker 1 and 2’s default choice probabilities for firm 1 and 2. The chosen application probability of worker 1 is a set of reaction functions,  $(q_1(\mathbf{s}; p_1, \lambda_1), q_2(\mathbf{s}; p_2, \lambda_1))$ , that depend on the default application probabilities, cognitive costs



and the acquired signals. Given the equivalence between signal-based search and costly directed search, the functions are reformulated as  $(q_1(\mathbf{w}; p_1, \lambda_1), q_2(\mathbf{w}; p_2, \lambda_1))$ , and henceforth, simplified to  $(q_1, q_2)$  for worker 1, and  $(q_1^{-i}, q_2^{-i})$  for worker 2. The first two terms of (4) denotes worker  $i$ 's expected payoff. By applying to firm 1 with probability  $q_1^i$ , worker  $i$  obtains the job if worker  $-i$  does not apply it, but if worker  $-i$  does, then one of them will be randomly picked. As a result,  $q_j^i(1 - q_j^{-i} + \frac{q_j^{-i}}{2})$  can be interpreted as worker  $i$ 's job finding rate.

For a given set of wages  $(w_1, w_2)$ , workers' FOCs and optimal strategies are:

$$q_1^i : (1 - q_1^{-i} + \frac{q_1^{-i}}{2})w_1 - (q_1^{-i} + \frac{1 - q_1^{-i}}{2})w_2 = \lambda_i \log(\frac{1 - p_1^i}{p_1^i} \frac{q_1^i}{1 - q_1^i}) \quad (5)$$

$$q_1^{-i} : (1 - q_1^i + \frac{q_1^i}{2})w_1 - (q_1^i + \frac{1 - q_1^i}{2})w_2 = \lambda_{-i} \log(\frac{1 - p_1^{-i}}{p_1^{-i}} \frac{q_1^{-i}}{1 - q_1^{-i}}) \quad (6)$$

$$q_1^i = \frac{1}{1 + \frac{1 - p_1^i}{p_1^i} \exp(-(\frac{(1 - q_1^{-i} + \frac{q_1^{-i}}{2})w_1 - (q_1^{-i} + \frac{1 - q_1^{-i}}{2})w_2}{\lambda_i}))} \quad (7)$$

$$q_1^{-i} = \frac{1}{1 + \frac{1 - p_1^{-i}}{p_1^{-i}} \exp(-(\frac{(1 - q_1^i + \frac{q_1^i}{2})w_1 - (q_1^i + \frac{1 - q_1^i}{2})w_2}{\lambda_{-i}}))} \quad (8)$$

The probabilities of choosing firm 1,  $(q_1^i, q_1^{-i})$ , are defined for  $p_1^i, p_1^{-i} \in (0, 1)$ ,  $\lambda_i, \lambda_{-i} > 0$  and positive expected payoff. Default strategies are assumed to be non-deterministic. Information processing cost may be negligible but cannot be non-existent. This is supported by Woodford (2020) and Heng et al. (2020), where some degree of noise is always present in the internal representation of observations even if wages are perfectly observable. Christie and Schrater (2015) and Jacob et al. (2023) also indicate that any cognitive process and brain activity are metabolically costly, even at rest, there is constant cost to maintain continuous neural activity and brain functionality. Therefore, although cognitive costs can theoretically be 0, it cannot realistically be 0, so  $\lambda_i, \lambda_{-i} > 0$  always hold.

**Firms' Maximization Problem.** The profit maximization problem faced by firm  $j$  is:

$$w_j = \underset{w_j}{\operatorname{argmax}} [1 - (1 - q_j)(1 - q_j^{-i})](z_j - w_j), \quad \text{st. } z_j \geq w_j \geq 0 \quad (9)$$

Firm  $j$  observe a realization of its productivity  $z_j$ , and maximizes its profit by posting a wage,  $w_j$ , given wage posting strategy of the other firm,  $w_{-j}$ . The probability of at least one worker applying to firm  $j$  is  $[1 - (1 - q_j)(1 - q_j^{-i})]$ . Firms can only produce when they are matched with at least one worker. While being matched with both workers would not give them additional benefit of higher productivity, attracting higher application probabilities from the workers would increase the chances of them being able to produce, thereby providing incentives for them compete on wage offers. Each firm is subjected to the feasibility condition of setting its wage in between its productivity,  $z_j$ , and workers' outside option of 0. The equilibrium wages are determined in anticipation of workers' strategies given potential wage realizations. In view of workers constrained by attention, firms may choose to exploit this information asymmetry by extracting higher wage markdowns, which constitutes as a form of monopsony power.

In this model, firms are assumed to be playing pure strategies. I do not consider firms' incentive to signal and discriminate between workers, who may have different attention costs, because there are no additional benefit in hiring the worker with lower attention costs.

## 2.2 Second-stage Game: Competition Between Workers

Since this is a sequential game, I first examine the subgame, where workers make simultaneous moves in selecting which firm to apply to.

### 2.2.1 Complete Information

The subgame can be portrayed as a cooperation game between 2 workers for a given set of wages.

		Worker 2	
		F1	F2
Worker 1	F1	$\frac{w_1}{2}, \frac{w_1}{2}$	$w_1, w_2$
	F2	$w_2, w_1$	$\frac{w_2}{2}, \frac{w_2}{2}$

Suppose workers are fully attentive, then there will be complete information about wages. Three equilibria would be possible: Two pure Nash equilibria (NEs), (F1, F2) and (F2, F1) if  $2w_1 > w_2 > \frac{w_1}{2}$ ; and one mixed NE, (F1, F2;  $\frac{2w_1-w_2}{w_1+w_2}, \frac{2w_2-w_1}{w_1+w_2}$ ) if  $2w_1 \geq w_2 \geq \frac{w_1}{2}$ . The mixed NE would correspond to the conventional directed search equilibrium where workers are assumed to be using the same application strategy. If one of the firms increase wage relative to the other, workers' application probability to that firm increases. The application probability is positively correlated with wages. As for the pure NEs, in expectation of workers applying exclusively to different firms, firms are certain that they will be matched with one worker, thus would have no incentive to increase wages. Instead, they would push wages down to 0 to maximize profit. Workers would not deviate from pure strategies as this would not yield any better payoffs. Therefore, application probability is independent of wages.

On the other hand, when inattention is included, the game turns into an incomplete information game, where wages may not be perfectly visible. This would result in workers' application strategy being perturbed best response to wages offered by the firms and the strategy of the other worker. Unlike pure NEs in complete information, there is less certainty about workers applying strictly to different firms, so firms are not guaranteed to have at least one worker, and workers' application strategies are responsive to wage changes, therefore, wages would remain positive instead of being driven down to 0. Only in the limit of cognitive costs being negligible, then the solutions converge to that of the complete information game. The incorporation of attention costs postulates a possibility for workers to coordinate on fulfilling both vacancies with high probability without being penalized drastically.

### 2.2.2 Relation to Quantal Response Functions

I hope to draw a relation between workers' strategy with inattention, which is denoted as rationally inattentive functions (RIFs) and discussed extensively in this paper, and the quantal response functions (QRFs) that models decision-making with bounded rationality, which was studied comprehensively in other literature such as McKelvey and Palfrey (1995) and Goeree et al. (2016). If I use QRFs to model workers' search behaviours, it means I'm extending the standard NE by allowing for errors in workers' choices instead of them having attention constraint. Their strategies would be captured by smoothed-out response functions as follows:

$$q_{F1}^i = \frac{\exp(c\pi_{F1}(q_{F1}^{-i}))}{\exp(c\pi_{F1}(q_{F1}^{-i})) + \exp(c\pi_{F2}(q_{F1}^{-i}))} = \frac{1}{1 + \exp(-c(\pi_{F1}(q_{F1}^{-i}) - \pi_{F2}(q_{F1}^{-i})))} \quad (10)$$

$$q_{F1}^{-i} = \frac{\exp(c\pi_{F1}(q_{F1}^i))}{\exp(c\pi_{F1}(q_{F1}^i)) + \exp(c\pi_{F2}(q_{F1}^i))} = \frac{1}{1 + \exp(-c(\pi_{F1}(q_{F1}^i) - \pi_{F2}(q_{F1}^i)))} \quad (11)$$

where  $q_{F1}^i$  and  $q_{F1}^{-i}$  are the probabilities of applying to firm 1,  $\pi$  is the payoff from applying to the firms given opponent's strategy, and  $c$  is a precision measure that defines the degree of rationality or randomness in one's choice.  $c = 0$  indicates complete randomness and  $c \rightarrow \infty$  suggests complete rationality. It can also be interpreted as sensitivity to differences in payoffs. For a set of wages,  $(w_1, w_2)$ , the intersections of (10) and (11) determines the quantal response equilibrium (QRE). Figure 1 shows that as precision is lowered, the functions become more smoothed-out, indicating reduction in sensitivity to payoff difference, and the equilibrium solutions can be more diffuse and diverge further away from standard NE under complete information.

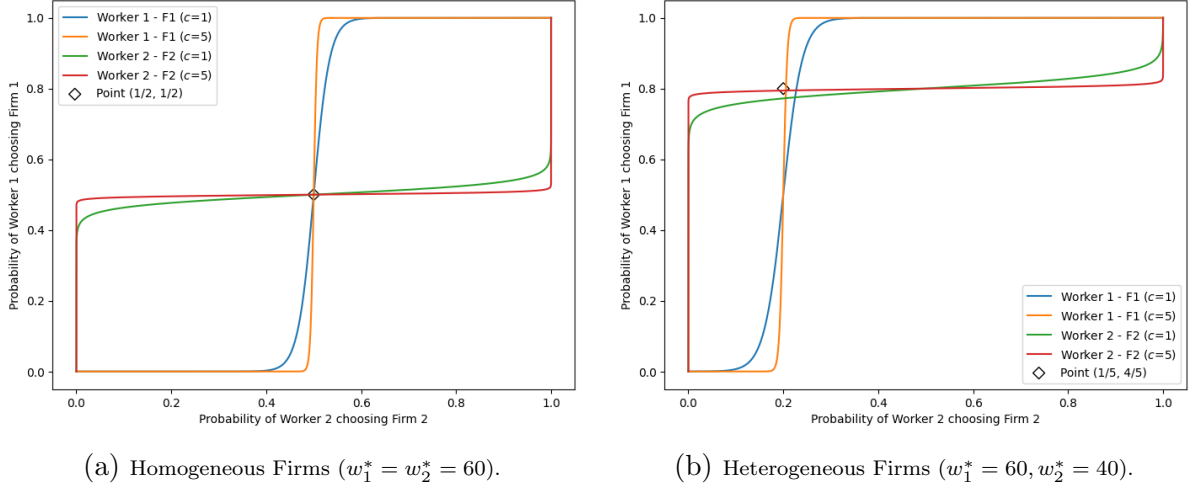


Figure 1: Workers' Quantal Response Functions for varying degree of randomness.

Comparing QRFs (10), (11) and RIFs (7), (8), cognitive cost in RIFs is essentially the inverse of precision in QRFs (i.e.  $c = \frac{1}{\lambda_i}$ ). Higher  $\lambda_i$  indicates higher unit cost of processing information, which implies smaller  $c$ , lower precision in learning about wages or lower sensitivity to payoff difference. In RIFs, workers' dependence on the expected payoff is additionally weighted by the defaults. Higher default probability of selecting firm 1 would lead to lower weight assigned to the expected payoff.<sup>2</sup> While these parallels have not been explicitly established, Goeree et al. (2020) has highlighted the potential link between the concepts.

Even though there are mathematical similarities between the equilibrium found using QRFs and RIFs (denoted as Rationally Inattentive Equilibrium (RIE)), the source of noise modelled and the implications are slightly different. In QRE, workers observe the wages precisely and better respond to beliefs about "randomness" in opponent's strategies and the expected payoffs. In RIE, wages are not directly ascertainable, workers learn about the wages by acquiring a costly signal. They may choose to remain partially informed considering the cost and benefit of doing so, thus their beliefs about the wages and the strategies of their opponent are shaped by the amount information they choose to acquire and process, and they are best responding to opponent's strategies given the noisy environment. The benefit of modelling workers' behaviours using RIFs are two-fold. Firstly, it not only rides on top of QRE's advantage as a statistical model, where one can infer the level of rationality or the degree of errors in individuals' decision-making by fitting the model to empirical data. RIE can also be perceived as an econometric model, where cognitive costs can be approximated with IQ scores, and be used to evaluate the relationship between IQ and search decisions. The second plus point is RIE's policy impact. In the QRE model, the source of randomness in workers' strategies could be difficult to decipher, but the

<sup>2</sup>Figure 1 can similarly be used to illustrate the case of workers with homogeneous attention costs, having random default strategies, and cognitive costs of either 1 or  $\frac{1}{5}$ .

backdrop to the RIE model is one of information-theoretic, the source of noise is attributed to endogenous decision of information acquisition and processing, leaving room for policy tools to improve information provision on the basis of attention allocation.

**Workers Learning from experience.** In repeated interactions and learning from experiences, to model the dynamics towards a RIE, the mechanism goes as follows: Workers could enter the market uninformatively, or having some prior beliefs. They then decide to acquire and process information, considering for the costs involved. Since the source of noise comes from not knowing the wages, which can be interpreted as the underlying state of the game, workers are acquiring information to revise their beliefs about the game states and best responding to their opponents given information they received. By learning the noise in the other worker's strategy, one is essentially capturing the underlying state that governs their opponent's choices, and adapting their own strategies. The long-run equilibrium consists of workers optimally allocating their attention resources and eventually converge to RIE for a given set of wages. While it is possible to provide policy recommendations that influence the learning process and equilibrium outcomes in the RIF framework, the QRE interpretation, which models the randomness or errors in workers' strategies, is harder to generate policy interventions that controls this degree of noise.

### 2.2.3 Deterministic Search Behaviour

**Applying to both firms with certainty.** While some may believe that applications are almost costless, one do not have to make deliberate choice between different firms, and can simply apply to all firms with probability 1, this is a misconception. On the very contrary, both tangible and intangible cost associated with each application decision prevent one from applying to all firms and have to make a selection. These costs include: documentation costs (i.e. preparing CV, customizing cover letters, time spent filling out online applications and following up emails); screening test costs (i.e behavioural and situational judgements tests, psychometric tests, HireVue assessments, Leetcode training); and social capital costs (i.e. job referee or reference letter consists of favor asking, which involve both time and emotional costs, such as favor reciprocity and depletion of goodwill). These costs are assumed to be equally imposed on all job applications and will not affect which firm a worker choose to apply for, but they have an impact on the overall application strategy, which will not be a "hasty click-and-submit" one.

**Finite cost for precise information.** In Mattsson and Weibull (2002), it was assumed that marginal cost to achieve zero mistake is infinite, so one will optimally choose positive mistake probabilities. This is analogous to the model put forth by this paper. As long as information cost is positive, one should optimally decide to tolerate some imprecision and behave probabilistically as a result. However, it may be possible for one to incur finite attention cost to acquire precise information that lead to deterministic behaviour. Given bounded defaults,  $p_j^i \in (0, 1)$ , if there is 0 expected payoff ( $\mathbb{E}$ ) or that  $\lambda_i$  and  $\lambda_{-i}$  are sufficiently high, then there are no incentive to acquire information, equilibrium strategy would be the same as default.

$$q_1^i = p_1^i \text{ if } \begin{cases} \mathbb{E} = (1 - q_1^{-i} + \frac{q_1^{-i}}{2})w_1 - (q_1^{-i} + \frac{1-q_1^{-i}}{2})w_2 = 0 \\ \lim_{\lambda_i \rightarrow \infty} \exp(-\frac{\mathbb{E}}{\lambda_i}) = 1, \mathbb{E} \neq 0 \end{cases}$$

For non-zero expected payoff, workers would have incentive to acquire information at a cost. If the expected payoff is substantially high, then it could be worthwhile to learn the precise information, and the probability of choosing a firm may tend to 1 or 0.

$$q_1^i = 1 \text{ if } \lim_{d\mathbb{E} \rightarrow \infty} \exp(-\frac{d\mathbb{E}}{\lambda_i}) = 0, q_1^i = 0 \text{ if } \lim_{d\mathbb{E} \rightarrow -\infty} \exp(-\frac{d\mathbb{E}}{\lambda_i}) = \infty$$

However, there is lack of incentive for firms to set substantially different wages unless they differ considerably in productivity and the fear of losing a worker is high. Therefore, workers will likely choose imprecise information that leads to probabilistic search behaviour.

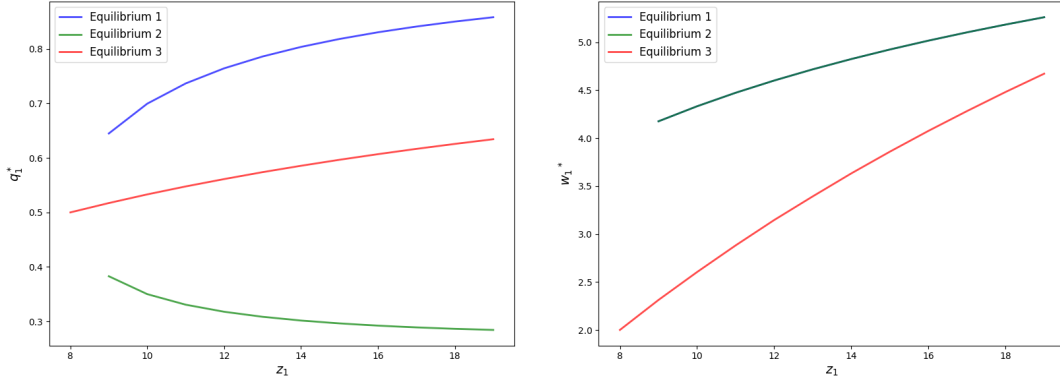
**Deterministic default strategies.** Workers may possess strong initial bias and apply with conviction to one of the firms by default. They can choose to acquire information and deviate away from this deterministic choice if the expected payoff of doing so is higher than the cost. But in the case of high information costs, inattentive workers would not acquire any information, deterministic default would thus translate into deterministic equilibrium behaviour. If both workers apply to the same firm, there will be a severe overcrowding problem. Since it is unlikely for firms to have such a strong goodwill ex-ante, this model assumes bounded defaults (i.e.  $p_j^i \in (0, 1)$ ), which remove this possibility.

## 2.3 Subgame Perfect Equilibrium

I have specified the GPDS framework and explained the second-stage game, where workers take wages as given and select a set of application probabilities after information acquisition. In the full game, the application probabilities are plugged back into firms' maximization problems to determine the equilibrium wages.

**Definition 2.2** (Subgame Perfect Equilibrium (SPE)). *A subgame perfect equilibrium is defined to be a tuple,  $\{q_j^i(w_1, w_2; p_j^i, \lambda_i), q_j^{-i}(w_1, w_2; p_j^{-i}, \lambda_{-i}), w_1^e, w_2^e\}$ , where  $q_j^i$  is a function of default choice probabilities,  $p_j^i$ , cognitive cost,  $\lambda_i$ , and wages,  $w_1, w_2$ .*

1.  $q_j^i(w_1, w_2; p_j^i, \lambda_i)$  maximizes worker  $i$ 's payoff, take as given  $w_1, w_2$  and  $q_j^{-i}(w_1, w_2; p_j^{-i}, \lambda_{-i})$ .
2.  $w_j^e$  maximizes firm  $j$ 's payoff, take as given  $q_j^i(w_1, w_2; p_j^i, \lambda_i), q_j^{-i}(w_1, w_2; p_j^{-i}, \lambda_{-i})$  and  $w_{-j}^e$ .



(a) Worker 1's Equilibrium Probabilities

(b) Firm 1's Equilibrium Wages

Figures show bifurcation graphs with  $q_1^*$  and  $w_1^*$  plot against changes in  $z_1$ .  $z_1 \in [8, 20]$ ,  $z_2 = 8$ ,  $p_1 = p_1^{-i} = \frac{1}{2}$ ,  $\lambda = 1$ .

Figure 2: Effects of Varying Productivity on Equilibrium Probabilities and Wages

Figure 2 demonstrates for a given set of exogenous parameters,  $\lambda, z_1, z_2, p_1, p_1^{-i}$ , worker 1's equilibrium selection of application probability to firm 1 and firm 1's equilibrium wage. In the subsections after, I will explore workers' behaviour, particularly the fact there could be three possible equilibria, followed by firms' behaviour, where equilibrium wages offered by firm 1 seems to be constantly higher for some equilibria than the other.

### 2.3.1 Workers' Equilibrium Search Behaviours

There could exist different types of SPE, which depends on the symmetric or asymmetric nature of workers' application strategies. Figure 3 shows there are two possible types of equilibrium,

one where both workers apply more randomly, but as  $z_1$  grows against  $z_2$ , they apply with higher probability to the more productive firm. This is characterized as the symmetric subgame perfect equilibrium (SSPE), which was addressed in Wu (2024). There is another equilibrium type, where workers apply with higher probability towards different firms. Growing firm heterogeneity is coupled with growing asymmetry in workers' equilibrium strategies, where they apply with higher probability to different firms. This is characterized as the asymmetric subgame perfect equilibrium (ASPE), which was not accounted for. I will first explore SSPE.

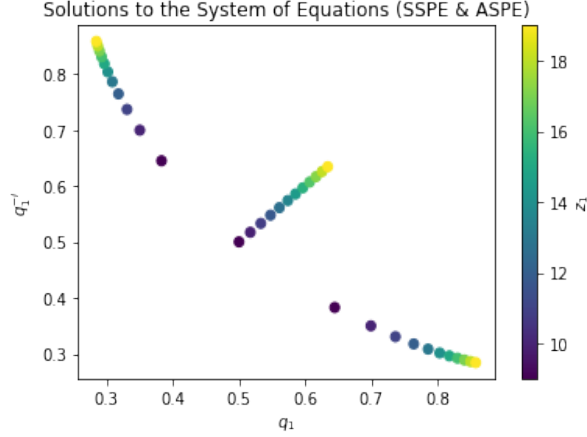


Figure shows both the SSPE and ASPE solutions when  $\lambda = 1$ ,  $z_1 \in [8, 20]$ ,  $z_2 = 8$ ,  $p_1 = p_1^{-i} = \frac{1}{2}$ . The colour spectrum displays a range of  $z_1$  while fixing  $z_2$  to show the impact of heterogeneous productivity on workers' equilibrium behaviour.

Figure 3: Workers' Application Probability

**Definition 2.3** (Symmetric Subgame Perfect Equilibrium (SSPE)). *Symmetric subgame perfect equilibrium is defined to be a case of SPE that  $q_j^i(w_1, w_2; p_j^i, \lambda_i) = q_j^{-i}(w_1, w_2; p_j^{-i})$ . Both workers adopt the same strategy in choosing which firm to apply to in the equilibrium. Firms are assumed in this setting to only use pure strategies in wage posting.*<sup>3</sup>

**Symmetric Subgame Perfect Equilibrium (SSPE).** Retaining the original assumptions from Wu (2024) where workers are homogeneous in attention costs,  $\hat{c}_1 = \hat{c}_2$ , such that  $\lambda_i = \lambda_{-i} = \lambda$ ,  $p_1^i = p_1^{-i} = p$ ; and they adopt symmetric application strategies,  $q_1^i = q_1^{-i} = q$ . I relax the assumption on random search default strategies, such that default strategies can vary but bounded ( $p \in (0, 1)$ ). From workers' search problem (4), worker  $i$ 's FOC:

$$(1 - q + \frac{q}{2})w_1 - (q + \frac{1 - q}{2})w_2 = \lambda \log(\frac{1 - p}{p} \frac{q}{1 - q}) \quad (12)$$

The derivatives with respect to  $w_1$  (13) and  $w_2$  (14) show that higher wage posted by firm 1 increase probability of applying to firm 1, and higher wage posted by firm 2 will decrease probability of applying to firm 1.

$$\frac{dq}{dw_1} = \frac{(2 - q)q(1 - q)}{2\lambda + q(1 - q)(w_1 + w_2)} > 0 \quad (13)$$

$$\frac{dq}{dw_2} = \frac{-(1 + q)q(1 - q)}{2\lambda + q(1 - q)(w_1 + w_2)} < 0 \quad (14)$$

<sup>3</sup>While workers are playing a game similar to coordination game, firms are reacting based on workers' strategies, which in essence is a different type of game. Therefore, any symmetry or asymmetry in firms' wage offers, in terms of whether they are offering the same or different wages, are not denoted in the equilibrium definition that focus on the symmetric or asymmetric nature of workers' strategies.

Given the implicit functions and firms' maximization problem (9),  $w_1$  and  $w_2$  can be expressed in terms of one another.

$$w_1 = \frac{2(1-q)^2 z_1 - 2\lambda - q(1-q)w_2}{q(1-q) + 2(1-q)^2} \quad (15)$$

$$w_2 = \frac{2q^2 z_2 - 2\lambda - q(1-q)w_1}{q(1-q) + 2q^2} \quad (16)$$

*SSPE System of Equations.* The SSPE consists of a system of simultaneous equations, (12), (15), (16), which can be solved for  $(q^*, w_1^*, w_2^*)$ , where  $q^*$  is the equilibrium probability of applying to firm 1,  $w_1^*$  and  $w_2^*$  are the equilibrium wages selected by firms.

$$w_1^* = \max[(1-q^{*2})z_1 - q^{*2}z_2 - \frac{2q^*}{1-q^*}\lambda, 0] \quad (17)$$

$$w_2^* = \max[q^*(2-q^*)z_2 - (1-q^*)^2 z_1 - \frac{2(1-q^*)}{q^*}\lambda, 0] \quad (18)$$

$$q^* = \frac{1}{1 + \frac{1-p}{p} \exp(-\frac{2w_1^* - w_2^* - q^*(w_1^* + w_2^*)}{2\lambda})} \quad (19)$$

The more expensive information is, the less likely workers will acquire and process them, and thus are less sensitive to wage changes. Since workers' strategies are less dependent on wages, firms are more likely to offer low wages. Higher the  $\lambda$ , higher the markdown on wages. Workers are penalized for being inattentive.

**Lemma 2.1** (Uniqueness of Symmetric Subgame Equilibrium). *Given any wage announcement  $(w_1, w_2)$ , there is a unique symmetric subgame equilibrium (SSE) for any  $p$ ,  $p \in (0, 1)$ .*

*Proof.* Shown in Appendix A.2. □

In Wu (2024)'s model, workers are unbiased, and adopt a default random search strategy. As I instigated that workers have potentially biased initial beliefs about wage distribution, they could have non-uniform default choice probabilities across the firms.

**Definition 2.4** (Belief Distortion). *Any deviation of default application strategy away from random is denoted as a result of belief distortion based on informative priors.*

**Definition 2.5** (Correct Prior Beliefs). *Prior beliefs are defined to be correct if workers' default choice probabilities matches the application rate firms should receive given their productivity levels (i.e.  $p = \frac{z_1}{z_1 + z_2}$ ).*

The information cost associated with revision of beliefs can be perceived as one that facilitate more accurate alignment of decision-making. However, workers can receive a signal that is in conflict with their prior beliefs. The cost incurred to deviate away from default strategy could therefore have an additional interpretation of psychological cost from one's belief clashing with new information, coined as "cognitive dissonance". (Festinger (1962)) Workers are not only updating their beliefs with new information, but they are also rectifying their past biases. Changes in default strategies on equilibrium outcome could highlight how innate beliefs may bias search behaviours and affect wage-setting.

**Proposition 1** (Belief Distortion on Equilibrium Application Probability). *For homogeneous workers adopting identical search strategies, biased defaults lead to biased search.*

*Proof.* Based on (17) and (18), the indirect impact of default application probability on equilibrium wages are

$$\frac{dw_1^*}{dp} = -2[q^* z_1 + q^* z_2 + \frac{1}{(1-q^*)^2} \lambda] \frac{dq^*}{dp} \quad (20)$$

$$\frac{dw_2^*}{dp} = 2[(1 - q^*)z_1 + (1 - q^*)z_2 + \frac{1}{q^{*2}}\lambda] \frac{dq^*}{dp} \quad (21)$$

And based on workers' FOC (12),

$$\frac{dq^*}{dp} = \frac{\frac{\lambda}{(1-p)p} + (1 - \frac{q^*}{2})\frac{dw_1^*}{dp} - (\frac{q+1}{2})\frac{dw_2^*}{dp}}{\frac{\lambda}{q^*(1-q^*)} + \frac{1}{2}(w_1^* + w_2^*)} \quad (22)$$

Combine (22),(20) and (21) shows the effect of changing default probabilities on equilibrium probabilities.

$$\frac{dq^*}{dp} = \frac{\frac{1}{p(1-p)}}{\frac{1}{q^*(1-q^*)} + \frac{\frac{1}{2}(w_1^* + w_2^*) + (2q^* - 2q^{*2} + 1)(z_1 + z_2)}{\lambda} + [\frac{2-q^*}{(1-q^*)^2} + \frac{1+q^*}{q^{*2}}]} > 0 \quad (23)$$

where  $\lambda > 0$  and  $p \in (0, 1)$ .

Higher default application probability to firm 1 would increase the equilibrium application probability to firm 1.  $\square$

There is higher chances of overcrowding at a single firm if one of the firms receive higher application probability by default, leading to higher likelihood of mismatch. As for the equilibrium wages, since  $\frac{dq^*}{dp} > 0$ , for (20),  $\frac{dw_1^*}{dp} < 0$ ; for (21),  $\frac{dw_2^*}{dp} > 0$ . Firms would set a lower wage than their rival if there is higher probability of meeting a worker in the equilibrium.

**Proposition 2** (Cognitive Costs on Equilibrium Application Probability). *Increasing cognitive costs lead to stronger reliance on defaults.*

*Proof.* From (19), define expected payoff to be  $\mathbb{E}$ .

$$\mathbb{E} = (1 - \frac{q^*}{2})w_1^* - (\frac{1+q^*}{2})w_2^* = \frac{2w_1^* - w_2^* - q^*(w_1^* + w_2^*)}{2} \quad (24)$$

$$\lim_{\lambda \rightarrow 0^+} q^* = \begin{cases} 1, & \text{if } \mathbb{E} > 0, w_1^* > 2w_2^* \\ 0, & \text{if } \mathbb{E} < 0, 2w_1^* < w_2^* \\ p, & \text{if } \mathbb{E} = 0, p = \frac{2w_1^* - w_2^*}{w_1^* + w_2^*} \end{cases}, \lim_{\lambda \rightarrow \infty} q^* = p \forall p \quad (25)$$

Workers are sensitive to  $\mathbb{E}$  in absence of cognitive costs, both workers will be attract by positive expected payoff. For extremely expensive information, workers are not sensitive to expected payoff at all, equilibrium strategy converges to the default strategy in the limit. Based on (22) and (23), accounting for  $\lambda \rightarrow \infty$ ,  $w_1^* \rightarrow 0$ ,  $w_2^* \rightarrow 0$ ,  $\frac{dw_1^*}{dp} \rightarrow 0$ ,  $\frac{dw_2^*}{dp} \rightarrow 0$ ,

$$\lim_{\lambda \rightarrow 0^+} \frac{dq^*}{dp} = 0, \lim_{\lambda \rightarrow \infty} \frac{dq^*}{dp} = 1 \quad (26)$$

Equilibrium application probabilities are more reliant on defaults as cognitive costs increase.  $\square$

By allowing workers to have varying default search strategies, I demonstrated that workers could carry innate biases forward into the equilibrium strategies when information is costly. I also highlighted, on the firms' side, their wage postings are influenced by workers' default strategies. They could also possess more power in suppressing wages when information costs are high. This provides new insights on labour market mismatch, where potential overcrowding at a single firm could be attributed to biased defaults, on top of information frictions. (Appendix B.2)

**Existence of Multiple Equilibria.** One of the key contributions of this paper is to explore the possibility of multiple equilibria (as shown in Figure 19). Workers are often assumed in



search models to behave symmetrically by choosing the same application strategy, presence of multiplicity has largely been discarded to prevent potential coordination problems. However, there are no definite indication that symmetric equilibrium is easier to reach than asymmetric ones, where workers could choose different application strategies. By exploring the presence of equilibrium multiplicity, it is possible to evaluate social welfare under each scenario, and opens up the possibility of equilibrium selection.

**Definition 2.6** (Asymmetric Subgame Perfect Equilibrium (ASPE)). *Asymmetric subgame perfect equilibrium is defined to be a case of SPE that  $q_j^i(w_1, w_2; p_j^i, \lambda_i) \neq q_j^{-i}(w_1, w_2; p_j^{-i}, \lambda_{-i})$ . Workers use different mixed strategies in applying to firms. Firms are assumed in this setting to only use pure strategies in wage posting.*

**Asymmetric Subgame Perfect Equilibrium (ASPE).** Assuming workers remain homogeneous in attention costs, the ASPE solution, which comprises of a tuple  $(q_1^*, q_1^{-i*}, w_1^*, w_2^*)$ , can be derived from workers' search problem (4) and firms' maximization problem (9). Workers' FOCs are,

$$\text{Worker 1 } (q_1): (1 - q_1^{-i} + \frac{q_1^{-i}}{2})w_1 - (q_1^{-i} + \frac{1 - q_1^{-i}}{2})w_2 = \lambda \log(\frac{1 - p}{p} \frac{q_1}{1 - q_1}) \quad (27)$$

$$\text{Worker 2 } (q_1^{-i}): (1 - q_1 + \frac{q_1}{2})w_1 - (q_1 + \frac{1 - q_1}{2})w_2 = \lambda \log(\frac{1 - p}{p} \frac{q_1^{-i}}{1 - q_1^{-i}}) \quad (28)$$

where  $q_1^i$  is written as  $q_1$  for worker  $i$ .

Solving for firms' maximization problem,

$$\text{Firm 1: } [1 - (1 - q_1)(1 - q_1^{-i})] = (1 - q_1^{-i})(z_1 - w_1) \frac{dq_1}{dw_1} + (1 - q_1)(z_1 - w_1) \frac{dq_1^{-i}}{dw_1} \quad (29)$$

$$\text{Firm 2: } (1 - q_1 q_1^{-i}) = -(z_2 - w_2) [q_1^{-i} \frac{dq_1}{dw_2} + q_1 \frac{dq_1^{-i}}{dw_2}] \quad (30)$$

From (27) and (28), the derivatives with respect to  $w_1$  and  $w_2$  are,

$$\frac{dq_1}{dw_1} = \frac{(w_1 + w_2)(2 - q_1)q_1^{-i}(1 - q_1^{-i})q_1(1 - q_1) - 2\lambda(2 - q_1^{-i})(1 - q_1)q_1}{(w_1 + w_2)^2 q_1^{-i}(1 - q_1^{-i})q_1(1 - q_1) - 4\lambda^2} \quad (31)$$

$$\frac{dq_1^{-i}}{dw_1} = \frac{(2 - q_1^{-i})(1 - q_1)q_1(1 - q_1^{-i})q_1^{-i}(w_1 + w_2) - 2\lambda q_1^{-i}(1 - q_1^{-i})(2 - q_1)}{(w_1 + w_2)^2 q_1^{-i}(1 - q_1^{-i})q_1(1 - q_1) - 4\lambda^2} \quad (32)$$

$$\frac{dq_1}{dw_2} = \frac{2\lambda(1 + q_1^{-i})q_1(1 - q_1) - q_1 q_1^{-i}(1 - q_1)(1 - q_1^{-i})(1 + q_1)(w_1 + w_2)}{q_1 q_1^{-i}(1 - q_1)(1 - q_1^{-i})(w_1 + w_2)^2 - 4\lambda^2} \quad (33)$$

$$\frac{dq_1^{-i}}{dw_2} = \frac{2\lambda(1 + q_1)q_1^{-i}(1 - q_1^{-i}) - q_1 q_1^{-i}(1 - q_1)(1 - q_1^{-i})(1 + q_1^{-i})(w_1 + w_2)}{q_1 q_1^{-i}(1 - q_1)(1 - q_1^{-i})(w_1 + w_2)^2 - 4\lambda^2} \quad (34)$$

Substituting the derivatives into (29) and (30),

$$\frac{1 - (1 - q_1)(1 - q_1^{-i})}{z_1 - w_1} = \frac{(w_1 + w_2)(1 - q_1)q_1(1 - q_1^{-i})q_1^{-i}[(2 - q_1^{-i})(1 - q_1) + (2 - q_1)(1 - q_1^{-i})] - 2\lambda(1 - q_1^{-i})(1 - q_1)[q_1^{-i}(2 - q_1) - q_1(2 - q_1^{-i})]}{(w_1 + w_2)^2 q_1(1 - q_1)q_1^{-i}(1 - q_1^{-i}) - 4\lambda^2} \quad (35)$$

$$\frac{1 - q_1 q_1^{-i}}{z_2 - w_2} = -\left\{ \frac{(w_1 + w_2)(1 - q_1^{-i})q_1^{-i}(1 - q_1)q_1[(2 - q_1)q_1^{-i} + q_1(2 - q_1^{-i})] - 2\lambda q_1 q_1^{-i}[(1 - q_1)(2 - q_1^{-i}) + (1 - q_1^{-i})(2 - q_1)]}{(w_1 + w_2)^2 q_1(1 - q_1)q_1^{-i}(1 - q_1^{-i}) - 4\lambda^2} \right\} \quad (36)$$

ASPE System of Equations. ASPE solution, consisting a set of simultaneous equations, are formulated<sup>4</sup>:

$$q_1^* = \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2-q_1^{-i*})w_1^* - (1+q_1^{-i*})w_2^*}{2\lambda}))} \quad (37)$$

$$q_1^{-i*} = \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2-q_1^*)w_1^* - (1+q_1^*)w_2^*}{2\lambda}))} \quad (38)$$

$$w_1^* = \max\{z_1 - \lambda \frac{(q_1^* + q_1^{-i*} - q_1^* q_1^{-i*})[1 - (\frac{1}{q_1^* - q_1^{-i*}} \log(\frac{q_1^*}{1 - q_1^*} \frac{1 - q_1^{-i*}}{q_1^{-i*}}))]^2 q_1^* q_1^{-i*} (1 - q_1^*)(1 - q_1^{-i*})]}{(1 - q_1^{-i*})(1 - q_1^*)(q_1^* + q_1^{-i*} - q_1^* q_1^{-i*}) - \frac{1}{2(q_1^* - q_1^{-i*})} \log(\frac{q_1^*}{1 - q_1^*} \frac{1 - q_1^{-i*}}{q_1^{-i*}})}}, 0\}$$

$$q_1^* q_1^{-i*} (1 - q_1^*)(1 - q_1^{-i*})(2q_1^* q_1^{-i*} - 3q_1^* - 3q_1^{-i*} + 4) \quad (39)$$

$$w_2^* = \max\{z_2 - \lambda \frac{(1 - q_1^* q_1^{-i*})[1 - q_1^* q_1^{-i*} (1 - q_1^*)(1 - q_1^{-i*})(\frac{1}{q_1^* - q_1^{-i*}} \log(\frac{q_1^*}{1 - q_1^*} \frac{1 - q_1^{-i*}}{q_1^{-i*}}))]^2]}{q_1^* q_1^{-i*} (1 - q_1^* q_1^{-i*}) - q_1^* q_1^{-i*} (1 - q_1^*)(1 - q_1^{-i*})(q_1^* + q_1^{-i*} + 2q_1^* q_1^{-i*})} - \frac{1}{2(q_1^* - q_1^{-i*})} \log(\frac{q_1^*}{1 - q_1^*} \frac{1 - q_1^{-i*}}{q_1^{-i*}})}}, 0\}$$

$$\quad (40)$$

As before, increasing information cost would lower workers' sensitivity to expected payoff. Firms would have monopsony power and extract larger markdowns and put downward pressure on wages. Workers, in response, could lack incentive to acquire any information and adjust away from defaults, thus rendering ASPE non-existent. While small  $\lambda$  unambiguously has a positive impact on wages, it is less intuitive in ASPE as workers are expected to manoeuvre towards pure coordination, and firms should set lower wages in anticipation of high probability in meeting at least one worker. However, when workers apply more to different firms, the second term of (39) and (40) becomes large and pushes the wages down. As wages are lowered, workers could again lack incentive to acquire any information, and ASPE cease to exist.<sup>5</sup>

While SSPE solutions are conditional on the wage announcement, its existence and uniqueness are not dependent on any specific values of wages. However, as described in the analysis about information costs, the necessary condition for the existence of asymmetric strategies in the subgame is dependent on specific values of wage announcements. Before exploring ASPE, I first characterize the subgame equilibrium in the second stage given a set of wage announcements  $(w_1, w_2)$ . If they exist for a set of wages in the subgame, then there will be mapping from the wages to the productivity, and ASPE would exist as a result.

In the subgame, workers could adopt asymmetric strategies,  $(q_1, q_1^{-i})$ , where  $q_1 \neq q_1^{-i}$ . Their application probabilities go in opposite direction – for example, an increase in worker 1's choice probability of firm 1 will induce a decrease in worker 2's choice probability of firm 1. For a given

<sup>4</sup> $q_1$  and  $q_1^{-i}$  refers to  $q_1^*$  and  $q_1^{-i*}$  respectively in the wage determination equations.

<sup>5</sup>While plotting this set of equations graphically for  $\lambda \in (0, \infty)$  also shows a symmetric solution, it does not correspond to the SSPE solution, because  $q_1 = q_1^{-i}$  in the above wage setting equations would lead to undefined results. It is therefore essential to note that the set of equations, (37), (38), (39), (40) only provide defined solutions if ASPE exists.

set of  $(w_1, w_2)$ , based on (37) and (38), denoting  $q_1 = f(q_1^{-i})$  and  $q_1^{-i} = g(q_1)$  respectively.

$$\frac{df(q_1^{-i})}{dq_1^{-i}} = -\frac{1}{(1 + \frac{1-p}{p} \exp(-(\frac{(2-q_1^{-i})w_1 - (1+q_1^{-i})w_2}{2\lambda})))^2} \frac{1-p}{p} \exp(-(\frac{(2-q_1^{-i})w_1 - (1+q_1^{-i})w_2}{2\lambda})) \frac{w_1 + w_2}{2\lambda} < 0 \quad (41)$$

$$\frac{dg(q_1)}{dq_1} = -\frac{1}{(1 + \frac{1-p}{p} \exp(-(\frac{(2-q_1)w_1 - (1+q_1)w_2}{2\lambda})))^2} \frac{1-p}{p} \exp(-(\frac{(2-q_1)w_1 - (1+q_1)w_2}{2\lambda})) \frac{w_1 + w_2}{2\lambda} < 0 \quad (42)$$

where  $w_1, w_2, \lambda > 0$ , and  $p \in (0, 1)$ . Workers display some degree of loyalty or favouritism towards one of the firms that is different from the other. (Appendix B.3)

**Lemma 2.2** (Existence of Asymmetric Subgame Equilibrium). *For a given set of  $(w_1, w_2)$ , the necessary condition for asymmetric subgame equilibria (ASE) to exist is small  $\lambda$  relative to expected payoff, which is reliant on wage magnitude (for  $w_1 = w_2 = w$ ), or wage differences (for  $w_1 \neq w_2$ ) and that  $2w_1 > w_2 > \frac{w_1}{2}$  is fulfilled.*

*Proof.* See Appendix A.3. □

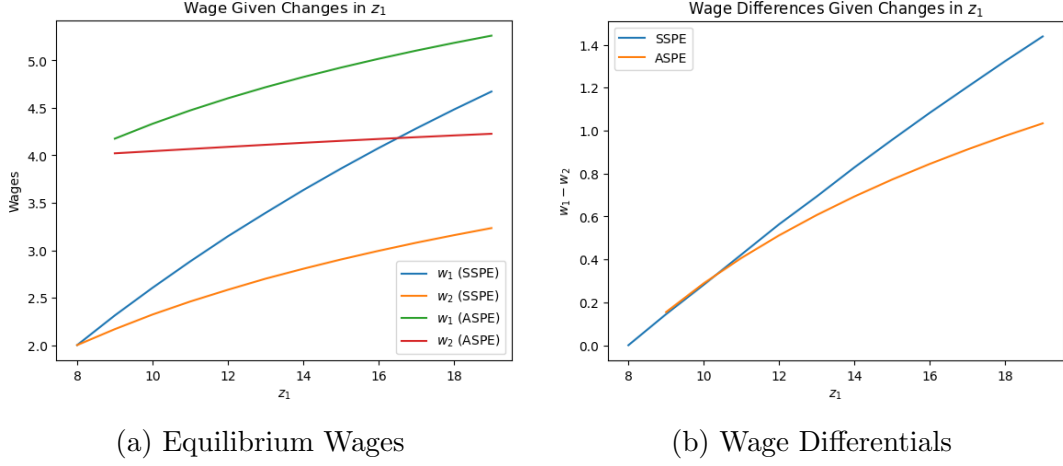
The existence of multiple equilibria shows that despite ex-ante symmetry in default strategy, there is potential for ex-post asymmetry in application strategies, where workers somewhat coordinate on applying to different firms with higher probability. The threshold of  $\lambda$  for existence is an implicit expression of  $w_1, w_2, q^*, p$ . The wage condition in Lemma 2.2 implies  $\lambda$  in turn needs to be small relative to productivities or that productivity difference has to be higher than difference in markdowns in the full game. Broadly speaking, the extent at which the productivity could differ to allow for the existence of ASPE is approximately twice the difference in either direction.

**Strategic Stability.** Suppose workers commit to select a strategy, firms in view of that would decrease their wages to maximize profits. Workers will be making a suboptimal decision, and firms can take advantage by setting lower wages. Such scenario is unstable as workers would be better off acquiring some information and respond to firms' deviation by altering their application strategies. The credible threat of workers acquiring and processing information off-the-equilibrium path supports the stability of the equilibrium on the path.

**Belief Distortion on Equilibrium Asymmetric Application Probability.** In SSPE, workers having a skewed default application strategy towards firm 1 would carry the bias forward, resulting in higher equilibrium application rates towards firm 1. The same applies for ASPE. Even though workers are applying more to different firms, higher default application probability towards firm 1 would also lead to greater "attraction" towards firm 1 in the equilibrium. Therefore, belief distortion can still contribute to biased search. (Appendix A.4.)

### 2.3.2 Firms' Equilibrium Wage Setting

Moving onto the firms' side, Figure 4a demonstrates equilibrium wages in SSPE and ASPE. There are two counter-intuitive points: (1) As shown in Figure 3, growing firm heterogeneity is coupled with growing asymmetry in workers' equilibrium strategies, where workers display more "loyalty" towards different firms. Despite firm 2 being less productive than firm 1 and have increasingly lower ability to match firm 1's wage offer, it can still receive higher application



Figures show both the SSPE and ASPE solutions when  $\lambda = 1$ ,  $z_1$  ranging from 8 to 20,  $z_2 = 8$ ,  $p_1 = p_1^{-i} = \frac{1}{2}$ .

Figure 4: Effects of Varying Productivity Differences on Equilibrium Wages

probability from one of the workers. (2) In ASPE, workers apply with higher probability to different firms, and knowing they will receive high chances of meeting at least one worker, firms should set low wages; whereas in SSPE, workers apply in a more randomized manner, and there could be greater uncertainty of meeting at least one worker, thus higher wages are expected. However, this is not observed in the numerical example when firms have similar productivities.

For the first point, since the relationship between wage differences and productivity differences is linear for SSPE in Figure 4b, the benefit of applying with higher probability to firm 1 consistently outweighs the gain from higher matching probability if one were to direct their search more to firm 2. Whereas in ASPE, changes in wage differentials against productivity disparity is growing at a decreasing rate. The benefit of applying for a job with a slightly lower wage but higher chances of being matched could outweigh the benefit of competing with the other worker at a job with slightly higher wage, as a result, leading to increasing asymmetry in workers' strategies. However, higher firm 1 productivity can still attract both workers to some extent when they are applying asymmetrically, showcased by the asymmetric meeting rate of the firms.

The most intriguing aspect is the magnitude of wages for SSPE and ASPE. Based on the wage determination equations (17) and (39),  $w_1^{\text{SSPE}}$  and  $w_1^{\text{ASPE}}$  are denoted as the equilibrium wage for firm 1 under SSPE and ASPE respectively. For  $w_1^{\text{ASPE}} \geq w_1^{\text{SSPE}}$ :

$$\frac{z_1 + z_2}{\lambda} \geq \frac{(q_1 + q_1^{-i} - q_1 q_1^{-i}) \left[ \left( \frac{1}{q_1 - q_1^{-i}} \log \left( \frac{q_1}{1 - q_1} \frac{1 - q_1^{-i}}{q_1^{-i}} \right) \right)^2 q_1 q_1^{-i} (1 - q_1) (1 - q_1^{-i}) - 1 \right]}{q^2 \left[ \left( \frac{1}{2(q_1 - q_1^{-i})} \log \left( \frac{q_1}{1 - q_1} \frac{1 - q_1^{-i}}{q_1^{-i}} \right) \right) q_1 q_1^{-i} (1 - q_1) (1 - q_1^{-i}) \right.} \\ \left. (2q_1 q_1^{-i} - 3q_1 - 3q_1^{-i} + 4) - (1 - q_1^{-i})(1 - q_1)(q_1 + q_1^{-i} - q_1 q_1^{-i}) \right]} - \frac{2}{q(1 - q)} \quad (43)$$

where  $q$ ,  $q_1$ ,  $q_1^{-i}$  refers to  $q^*$ ,  $q_1^*$  and  $q_1^{-i*}$  respectively.

By Lemma 2.2,  $\lambda$  has to be small relative to expected payoff (and productivity) for ASPE exist, LHS of (43) would most probably be larger than RHS, thus the inequality is likely to hold.

**Proposition 3** (Wages given Asymmetric vs. Symmetric Application Strategies). *Given existence of ASPE, if workers adopt more extreme asymmetric strategies, then  $w_j^{\text{ASPE}} \geq w_j^{\text{SSPE}}$  always hold. If workers adopt less extreme asymmetric strategies, then  $w_j^{\text{ASPE}} \geq w_j^{\text{SSPE}}$  only holds if  $\frac{z_1 + z_2}{\lambda} \geq \frac{1 - 2q^* + 2q^{*2}}{q^{*2}(1 - q^*)^2}$ , where  $q^*$  depends on  $\lambda$  and  $p$ , and indirectly on  $z_1$  and  $z_2$ . This*

condition is more likely to be satisfied with greater symmetry in firms' productivity, high aggregate productivity, low cognitive costs, and uniform default application strategy.

*Proof.* Suppose  $q_1^* \rightarrow 1$  and  $q_1^{-i*} \rightarrow 0$ , let RHS of the inequality (43) to be  $h(\cdot)$ ,

$$\lim_{q_1^* \rightarrow 1, q_1^{-i*} \rightarrow 0} h(\cdot) = \frac{1 - 3q^*}{q^{*2}(1 - q^*)} < 0$$

Then, suppose  $q_1^* \rightarrow 0$  and  $q_1^{-i*} \rightarrow 1$ ,

$$\lim_{q_1^* \rightarrow 0, q_1^{-i*} \rightarrow 1} h(\cdot) = \frac{1 - 3q^*}{q^{*2}(1 - q^*)} < 0$$

For extreme asymmetric cases,  $h(\cdot) < 0$ , and as  $\frac{z_1 + z_2}{\lambda} > 0$ , (43) holds.

Next, suppose  $q_1^* \rightarrow q^{*+}$  and  $q_1^{-i*} \rightarrow q^{*-}$ ,

$$\lim_{q_1^* \rightarrow q^{*+}, q_1^{-i*} \rightarrow q^{*-}} h(\cdot) = \frac{1 - 2q^* + 2q^{*2}}{q^{*2}(1 - q^*)^2} > 0$$

Same applies for the limit where  $q_1^* \rightarrow q^{*-}$  and  $q_1^{-i*} \rightarrow q^{*+}$ .

For intermediary asymmetric case,  $h(\cdot) > 0$ . Upholding of (43) depends on the exact value of  $q^*$ . Graphing  $h(\cdot)$  for  $0 < q^* < 1$  shows a minimum point at  $q^* = \frac{1}{2}$ . As the equilibrium application probability distribution becomes less uniform, the combined productivity relative to cognitive costs,  $\frac{z_1 + z_2}{\lambda}$ , would need to be larger to satisfy the inequality condition. Uniformity of  $q^*$  could be induced by workers having less belief distortion and firms being more homogeneous.

Let  $z_1 = xz_2$ ,  $x \geq 1$ , the threshold of  $x$  would be

$$\bar{x} \geq \frac{\lambda}{z_2} \frac{1 - 2q^* + 2q^{*2}}{q^{*2}(1 - q^*)^2} - 1 \quad (44)$$

In the extreme case of  $q^* \rightarrow 1$  as firm 1 becomes more productive,  $\lim_{q^* \rightarrow 1} \frac{1 - 2q^* + 2q^{*2}}{q^{*2}(1 - q^*)^2} = \infty$ . For  $w_j^{\text{ASPE}} \geq w_j^{\text{SSPE}}$ ,  $\frac{(x+1)z_2}{\lambda}$  needs to be larger than infinity. Disparity in firms' productivity could result in an inverse in relationship,  $w_j^{\text{SSPE}} \geq w_j^{\text{ASPE}}$ .  $\square$

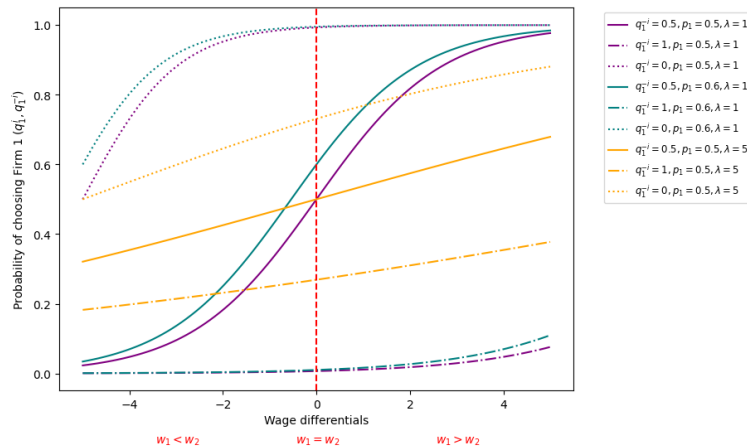


Figure 5: Changes in application probability to firm 1 given changes in wage differentials.

**Labour Supply and Sensitivity to Wage Differentials.** Figure 5 shows changes in workers' probability of choosing firm 1, which also represents how much labour one is willing to supply, given changes in wage differentials (fixing  $w_2$  and increasing  $w_1$ ) when opponent's application rate to firm 1 is set at 0, 0.5 and 1. Focusing on the comparison between (1)  $q_1^{-i} = 0.5$  (worker 2 has randomized search behaviour) and (2)  $q_1^{-i} = 0$  (worker 2 applies strictly to firm 2). In (1), worker 1 starts from the low wage sensitivity region of  $w_1 < w_2$ , and as  $w_1$  increases, high wage sensitivity region is reached, a small change in  $w_1$  has a large impact on  $q_1$ . Firms have strong incentive to outbid each other in small increments. There is intense wage competition among firms, coupled with incentive to keep wages low to maximize profits. In (2), worker 1's wage sensitivity is high when  $w_1 < w_2$ ,  $q_1$  rapidly converge to 1 as  $w_1$  increases, reaching the low wage sensitivity region. Since increasing  $w_1$  further generates no additional benefits, firm 1 will choose not to, leading to less intensive competition at the point where wages equalize. From firm 1's perspective, it faces 100% chance of not getting worker 2 in (2), as compared to 50% chance in (1), therefore, it will set a high initial wage to secure worker 1 in (2). While firm 1 may be incentivized to improve profits by decreasing  $w_1$  relative to  $w_2$  slightly at the point where  $w_1 = w_2$ , there is risk of losing both workers at small cost savings, especially if firm 2 choose to counteract by raising  $w_2$  to capitalize on firm 1's wage reduction. From firm 2's perspective, while it is almost guaranteed to have worker 2, if firm 1 set a substantially high relative wage, then worker 2 would be attracted to firm 1 as well, causing firm 2 to enter the high wage sensitivity region and risk losing both workers, therefore, firm 2 has an incentive to match the wage. As a result, to maintain the lock-in of workers, firms would avoid excessive wage cuts and kept the wages high. Basically, each firm is competing for one worker in the asymmetric case and for both in the symmetric case. The perceived "scarcity" of labour could intensify the competition in a slightly different manner that contributes to more aggressive initial wage setting.

However, one crucial thing about this pattern is that it depends on the ability of firms to compete. As firm 1's productivity increases against firm 2, firm 2 will experience less ability to match up the wage. Although firm 1 can still increase wage to benefit from the high wage sensitivity in (1) even if firm 2 is unable to match up, there is little incentive for firm 1 to do the same in (2) given it will be in the low wage sensitivity region. Therefore, lower wages could be observed in ASPE as compared to SSPE if firms become sufficiently different in productivity.

Figure 5 also demonstrates that higher cognitive costs would lower wage sensitivity in both cases. And an increase in default choice probabilities to firm 1 does not affect much when wages are equalized, but for larger wage differential, such as  $w_1 < w_2$ , wage sensitivity could increase in (1) as workers are more biased and responsive to firm 1's wage; and could lower wage sensitivity in (2), where their labour supply is already close to maximum, so the rate of change is slower.

To put wage sensitivity into perspective.

In SSE, as per (13) and (14),  $\frac{dq}{dw_1} > 0$ ,  $\frac{dq}{dw_2} < 0$ . For higher  $\lambda$ ,  $\frac{dq}{dw_1} \rightarrow 0$ ,  $\frac{dq}{dw_2} \rightarrow 0$ .

In ASE, (31), (32), (33), (34),  $\frac{dq_1}{dw_1} > 0$ ,  $\frac{dq_1^{-i}}{dw_1} > 0$ ,  $\frac{dq_1}{dw_2} < 0$ ,  $\frac{dq_1^{-i}}{dw_2} < 0$  when  $\lambda \rightarrow 0^+$ ; and as  $\lambda \rightarrow \infty$ ,  $\frac{dq_1}{dw_1} \rightarrow 0$ ,  $\frac{dq_1^{-i}}{dw_1} \rightarrow 0$ ,  $\frac{dq_1^*}{dw_2} \rightarrow 0$ ,  $\frac{dq_1^{-i}}{dw_2} \rightarrow 0$ .

Labor supply of a firm increases as its wage increases, and as information becomes more expensive, labour supply naturally become less responsive, ASE may also cease to exist. This is intuitive and may be less interesting, so instead, I compare the two equilibrium types for  $\lambda \rightarrow 0^+$ .

From (31), (32) and (13), and for a given set of  $(w_1, w_2)$ ,

$$\text{ASE: } \frac{dq_1}{dw_1} = \frac{2 - q_1}{w_1 + w_2}, \frac{dq_1^{-i}}{dw_1} = \frac{2 - q_1^{-i}}{w_1 + w_2}; \text{SSE: } \frac{dq}{dw_1} = \frac{2 - q}{w_1 + w_2}$$

Knowing  $q_1$  and  $q_1^{-i}$  go in opposing direction, one of the workers is always more sensitive to wages in ASE. Therefore, the previous argument about competing for the single more sensitive worker applies, which could drive up the wages.

Relating to Section 2.2.1. In complete information, only workers' symmetric strategies would be affected by wage changes and pure strategy would be independent of wages. With inattentive workers, the equilibrium can still comprises of close to pure strategies, but these are influenced by wages. This shows that a small perturbation in the system could lead to more wage competition as firms need to compete on wages regardless of workers' behaving symmetrically or asymmetrically. Particularly, workers would earn a wage of 0 when using pure strategies in complete information environment, but in this case, wages would be positive, which benefits the workers. Furthermore, the difference in equilibrium wages set in ASPE and SSPE implies variation in monopsony power across equilibrium types, which showcase the potential implications from equilibrium selection.

### 2.3.3 Social Welfare Comparison

The next core question would be to evaluate social welfare of different types of equilibrium.

**Social Welfare under SSPE.** I define the welfare function to be  $\mathbb{W}(q^*)$ , which depends on the equilibrium application probability towards firm 1,  $q^*$ .

$$\mathbb{W}(q^*) = (1 - (1 - q^*)^2)z_1 + (1 - q^{*2})z_2 - 2\lambda(q^* \log(\frac{q^*}{p}) + (1 - q^*) \log(\frac{1 - q^*}{1 - p})) \quad (45)$$

$$\mathbb{W}'(q^*) = 2(1 - q^*)z_1 - 2q^*z_2 - 2\lambda \log(\frac{1 - p}{p} \frac{q^*}{1 - q^*}) \quad (46)$$

$$\mathbb{W}''(q^*) = -2(z_1 + z_2) - 2\lambda(\frac{1}{q^*(1 - q^*)}) < 0 \quad (47)$$

$\mathbb{W}'(q^*)$  is concave. Welfare is increasing at a decreasing rate if  $\mathbb{W}'(q^*) > 0$ , and decreasing at increasing rate if  $\mathbb{W}'(q^*) < 0$ , a local maximum is expected if  $\mathbb{W}'(q^*) = 0$ . The impact of changes in equilibrium choice probabilities on welfare would depend on the specific values of the exogenous variables  $z_1$ ,  $z_2$ ,  $\lambda$  and  $p$ .

*Comparative Statistics.* First, in the extreme cases of almost costless information and infinitely expensive information, welfare is maximized when:

$$q^* = \frac{z_1}{z_1 + z_2}, \text{ for } \lambda \rightarrow 0^+; q^* = p = \frac{z_1}{z_1 + z_2}, \text{ for } \lambda \rightarrow \infty$$

For  $\lambda \rightarrow 0^+$ , local maximum welfare can be achieved at  $q^* = \frac{z_1}{z_1 + z_2}$ , which is when the default and equilibrium strategy correctly align with the productivity distribution.<sup>6</sup> As  $\lambda \rightarrow \infty$ ,  $q^*$  converges to  $p$ , if default strategy does not align with the productivity distribution, there will be persistent deviation away from maximum welfare, suggesting higher information cost could be detrimental. Social welfare can be maximized as long as  $q^*$  matches the true underlying productivity distribution. It would not matter what the default strategy is when workers are

---

<sup>6</sup>Even if default strategy does not align ( $p \neq \frac{z_1}{z_1 + z_2}$ ),  $q^*$  is not dependent on  $p$  when information is costless. Since  $\mathbb{E} > 0$ ,  $q^*$  is sensitive to the payoff differences and will adjust accordingly. Welfare will still be maximized at  $q^* = \frac{z_1}{z_1 + z_2}$ .

almost fully attentive, but when workers are completely inattentive, default strategies have to be correct for this to happen. As for intermediary  $\lambda$ , the critical point is found by weighing between the social benefit and cost. From (46),  $\mathbb{W}'(q^*) > 0$  if  $q^* < p$ ,  $p \leq \frac{z_1}{z_1+z_2}$ , and if  $q^* > p$ ,  $q^* < \frac{z_1}{z_1+z_2}$ . Both are cases where application to firm 1 is lower than needed, workers should be more targeted in order to maximize welfare.

The impact of an increase default application rate to firm 1,  $p$ , on  $\mathbb{W}(q^*)$  is

$$\frac{d\mathbb{W}(q^*)}{dp} = \frac{dq^*}{dp} [2(1-q^*)z_1 - 2q^*z_2 - 2\lambda \log(\frac{q^*}{p}) - 2\lambda \log(\frac{1-q^*}{1-p})] - 2\lambda(\frac{1-q^*}{1-p} - \frac{q^*}{p}) \quad (48)$$

Workers are not reliant on defaults at all when information is costless and depend completely on defaults if information is infinitely expensive. For the intermediary  $\lambda \in (0, \infty)$ , from (23),  $\frac{dq^*}{dp} > 0$ , as  $\lambda$  increases,  $\frac{d\mathbb{W}(q^*)}{dp}$  is likely going to be more positive for  $q^* > p$ , and more negative for  $q^* < p$ . Biased defaults adversely affect welfare when in the equilibrium, workers apply in a less targeted manner after information acquisition, suggesting there is additional welfare loss when workers have to “correct” their initial bias.

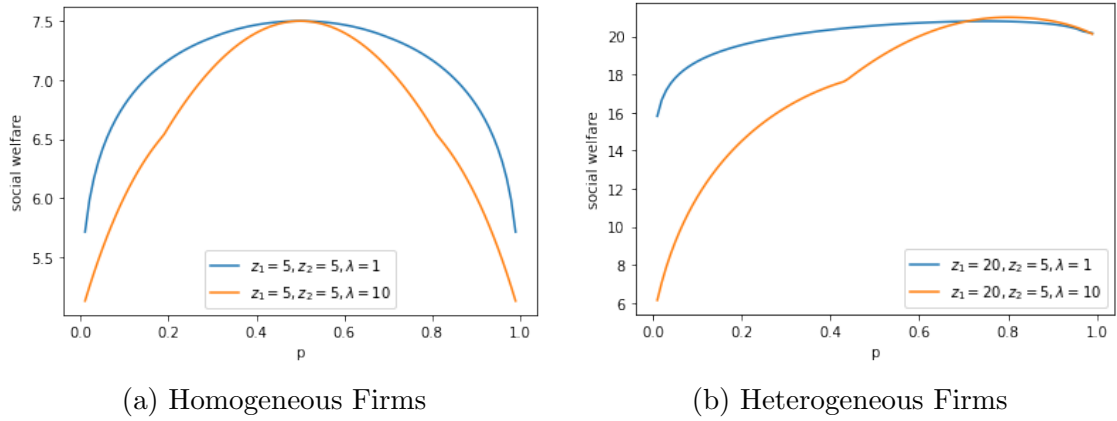


Figure 6: Changing Default Application Probability to Firm 1 on Social Welfare

Figure (6) shows that social welfare is generally higher when cognitive costs are low. Belief distortion could lead to a greater dip in social welfare when the default application rate does not match the productivity distribution. For heterogeneous firms, Figure 6b demonstrates that the peak of social welfare when workers have high cognitive costs could be slightly higher than when workers have low cognitive costs. This could be because workers with high cognitive costs are expected not to obtain any information and stick to their default strategy, which correctly matches the productivity distribution. On the other hand, when cognitive costs are low, workers would attempt to acquire some information, which affects the wage distribution, and in turn influence the alignment between equilibrium application probability and the productivity distribution.

**Social Welfare under ASPE.** To evaluate social welfare under ASPE, where  $q_1^* \neq q_1^{-i*}$ :

$$\begin{aligned} \mathbb{W}(q_1^*, q_1^{-i*}) = & (1 - (1 - q_1^*)(1 - q_1^{-i*}))z_1 + (1 - q_1^*q_1^{-i*})z_2 \\ & - \lambda[q_1^* \log(\frac{q_1^*}{p}) + (1 - q_1^*) \log(\frac{1 - q_1^*}{1 - p}) + q_1^{-i*} \log(\frac{q_1^{-i*}}{p}) + (1 - q_1^{-i*}) \log(\frac{1 - q_1^{-i*}}{1 - p})] \end{aligned} \quad (49)$$

**Proposition 4** (Social Welfare Comparison ASPE vs. SSPE). *Given necessary conditions for existence of multiple equilibria are satisfied, social welfare is higher under ASPE than SSPE.*



*Proof.* Shown in Appendix A.5.

In order for  $\mathbb{W}(q_1^*, q_1^{-i*}) > \mathbb{W}(q^*)$ , the relative social benefit of meeting a worker needs to outweigh the relative attention costs between the two types of equilibrium strategies.

$$\begin{aligned} & ((1 - q^*)^2 - (1 - q_1^*)(1 - q_1^{-i*}))z_1 + (q^{*2} - q_1^*q_1^{-i*})z_2 \\ & > \lambda \left[ \sum_{j=1}^2 q_j^* \log \frac{q_j^*}{p_j} + \sum_{j=1}^2 q_j^{-i*} \log \frac{q_j^{-i*}}{p_j} - 2q^* \log \left( \frac{q^*}{p} \right) - 2(1 - q^*) \log \left( \frac{1 - q^*}{1 - p} \right) \right] \end{aligned} \quad (50)$$

In the proof, I first demonstrate that expected productivity/ex-post welfare is higher under ASPE than SSPE. I then study the relative attention costs between the two equilibrium strategies, which indicates that more extreme asymmetric strategies tend to constitute higher attention cost, particularly when workers are biased. However, since such strategies emerge only when  $\lambda$  is relatively low (Lemma 2.2), by comparing the  $\lambda$  requirement for (50) to hold and the  $\lambda$  condition, formulated based on (37) and (38), assuming ASPE exist, I verify that existence condition is stricter than (50). As a result, social welfare is higher under ASPE than SSPE.  $\square$

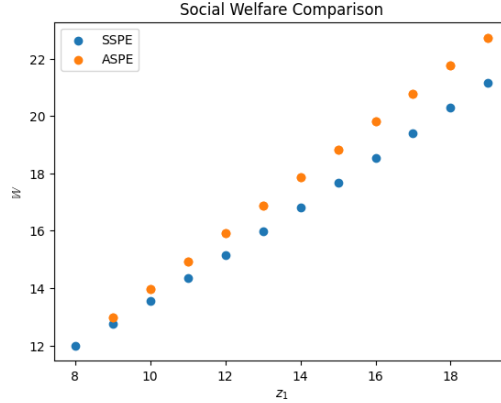


Figure shows welfare comparison for SSPE and ASPE when  $\lambda = 1$ ,  $z_1$  ranging from 8 to 20,  $z_2 = 8$ ,  $p_1 = p_1^{-i} = \frac{1}{2}$ .

Figure 7: Effects of Varying Productivity Differences on Social Welfare

Figure 7 shows some numerical examples where social welfare under ASPE is higher than SSPE. Intuitively, there is a higher chance of one-to-one matching when workers apply more to different firms, thus leading to more efficient social outcome, where both vacancies are more likely to be filled, than when workers apply symmetrically and risk crowding at one firm. This applies even as the degree of heterogeneity in productivity increases. Coupled with Figure 4a from before, which shows equilibrium wages to be higher in ASPE than SSPE, as well as the other way around when firm productivity difference widens. There is a range of low productivity differences, where ASPE can be both socially efficient and welfare-maximizing for workers.

**Concluding Remarks.** In this section, I lay down the fundamental building blocks for the GPDS model. As my focus is on workers' search behaviours, I devoted a part to the explore the second-stage game. In the subgame perfect equilibrium analysis, I show that workers' biased default could contribute to biased search in the equilibrium, implying that labour market mismatch could be attributed workers' innate bias. I also highlighted the possibility of multiple equilibria, where workers could adopt the same or different application strategies, contributing to the prevalence of both randomized and narrow search behaviour in the equilibrium. Proposition 3 and 4 are strong results. While it remains true that firms have power in extracting high markdowns when workers' attention costs are high, Proposition 3 showcases that workers may be

better off using asymmetric application strategies in the equilibrium than behaving symmetrically, especially when information costs are low and firms have similar productivity levels. It also highlights that unlike search under complete information, where pure strategies would lead to an equilibrium wage of 0, by adding an attention cost as a perturbation of the system, asymmetric mixed strategies that are close to the pure ones could in fact lead to a better outcome for workers, where they earn positive wage. On top of this, Proposition 4 indicates that ASPE is not only better off for workers, it can also generate higher social welfare than SSPE. The variation in efficiency for the equilibria has important policy implications. It advocates the potential of improving social welfare through equilibrium selection in presence of multiple equilibria. Though the mechanism is not explored in the current paper, this model establishes the potential to dive into it.

### 3 Variations in Default Strategies and Cognitive Costs

In this section, the GPDS model accommodates ex-ante worker heterogeneity in attention costs, and explore how this would affect equilibrium behaviour and contribute to possible mismatch. Worker disparity are often explored along the lines of skills, differences in attention or information costs have yet to be analyzed. Even though it is possible that high-skilled workers incur lower information processing cost than low-skilled workers when studying the same environment, and they may also possess better social connections and more insider information about the wage distribution, I do not include skills heterogeneity in this model, because such distinction might pose complications about how firms treat workers who applied to them equally. They could discriminate between workers, choosing high-skilled workers due to their superior capabilities instead of their access to better information. This would defeat the purpose of analyzing labour outcome on the basis of attention constraint, which could provide an alternative policy tool. Therefore, I focus on investigating horizontal differentiation in attention among workers without going into vertical differentiation in skills.

#### 3.1 Worker Heterogeneity in Default Strategies

Workers may differ in prior beliefs over the wage distribution as a result of insider information or firms having asymmetric reputation influence over the workers, thus resulting in workers adopting different default strategies, and this in turn affects the amount of information workers choose to acquire. Therefore, I hope to understand how heterogeneous defaults could contribute to workers' equilibrium behaviour in comparison to when they have the same default strategies, as well as its impact on social welfare.

**Model Assumption.** I relax the assumption of homogeneous default choice probability, and instead assume workers choose different firms with higher probability by default due to reasons, such as past experiences, prior information from social network or differences in beliefs about goodwill of the firms. Henceforth, I impose  $p_1 > \frac{1}{2}$  and  $p_1^{-i} < \frac{1}{2}$ , such that by default, worker 1 choose firm 1 with higher probability and worker 2 choose firm 2 with higher probability.

**Definition 3.1** (Strength of Belief Distortion). *Strength of belief distortion is defined to be the degree of deviation away from random search default strategy.*

*E.g.  $p_1 > \frac{1}{2}$  and  $p_1^{-i} < \frac{1}{2}$ : if  $p_1 \in (1 - \epsilon, 1)$  and  $p_1^{-i} \in (0, \epsilon)$ , where  $\epsilon$  is a small, positive value, belief distortion is described to strong for both workers; and  $p_1 \in (\frac{1}{2}, \frac{1}{2} + \epsilon)$  and  $p_1^{-i} \in (\frac{1}{2} - \epsilon, \frac{1}{2})$  implies weak belief distortion.*

*Relative belief distortion is defined to be a comparison of the asymmetry in default strategies between the workers:*

$$\text{Cases} = \begin{cases} p_1 = 1 - p_1^{-i}, \text{No relative distortion} \\ p_1 > 1 - p_1^{-i}, \text{Stronger worker 1 distortion} \\ p_1 < 1 - p_1^{-i}, \text{Stronger worker 2 distortion} \end{cases} \quad \text{for } p_1 \in (0.5, 1), p_1^{-i} \in (0, 0.5) \quad (51)$$

**SSPE.** Since workers have heterogeneous defaults, for them to behave symmetrically in the equilibrium, they would need to incur different attention costs for any positive  $\lambda$ . Firms do not discriminate based on workers' information types, the less attentive worker would incur higher cost but does not have an edge over the other worker in getting hired. Therefore, he/she would have incentive to lower information cost and adopt a different application strategy. In the special case of firms having the same productivity and set the same equilibrium wages. For  $\lambda > 0$ , workers with heterogeneous defaults but no relative distortion could incur the same information cost and choose a random search strategy in the equilibrium in view of the wages. But since wages do not differ between firms, the expected payoff would be 0, and workers incur

positive information costs, they would fare better by not acquiring any information, and use their default strategies instead. In contrast to the outcome with homogeneous workers, there are no SSPE as long as there is an information cost.

**ASPE.** Workers' FOCs are modified to include different default strategies, while firms' FOC, (29) and (30), remain unchanged since they are not directly dependent on  $p_1$  and  $p_1^{-i}$ . The ASPE solution,  $(q_1^*, q_1^{-i*}, w_1^*, w_2^*)$ , can be solved with the following system of equations.<sup>7</sup>

$$q_1^* = \frac{1}{1 + \frac{1-p_1}{p_1} \exp(-(\frac{(2-q_1^{-i*})w_1^* - (1+q_1^{-i*})w_2^*}{2\lambda}))} \quad (52)$$

$$q_1^{-i*} = \frac{1}{1 + \frac{1-p_1^{-i}}{p_1^{-i}} \exp(-(\frac{(2-q_1^*)w_1^* - (1+q_1^*)w_2^*}{2\lambda}))} \quad (53)$$

$$w_1^* = \max\{z_1 - \frac{\lambda(q_1^* + q_1^{-i*} - q_1^*q_1^{-i*})[1 - q_1^*q_1^{-i*}(1 - q_1^*)(1 - q_1^{-i*}) (\frac{1}{q_1^* - q_1^{-i*}} \log(\frac{1-p_1}{p_1} \frac{p_1^{-i}}{1-p_1^{-i}} \frac{q_1^*}{1-q_1^*} \frac{1-q_1^{-i*}}{q_1^{-i*}}))^2]}{(1 - q_1^{-i*})(1 - q_1^*)(q_1^* + q_1^{-i*} - q_1^*q_1^{-i*}) - q_1^*q_1^{-i*}(1 - q_1^*)(1 - q_1^{-i*})}, 0\} \\ (2q_1^*q_1^{-i*} - 3q_1^* - 3q_1^{-i*} + 4) \frac{1}{2(q_1^* - q_1^{-i*})} \log(\frac{1-p_1}{p_1} \frac{p_1^{-i}}{1-p_1^{-i}} \frac{q_1^*}{1-q_1^*} \frac{1-q_1^{-i*}}{q_1^{-i*}}) \quad (54)$$

$$w_2^* = \max\{z_2 - \frac{\lambda(1 - q_1^*q_1^{-i*})[1 - q_1^*q_1^{-i*}(1 - q_1^*)(1 - q_1^{-i*}) (\frac{1}{q_1^* - q_1^{-i*}} \log(\frac{1-p_1}{p_1} \frac{p_1^{-i}}{1-p_1^{-i}} \frac{q_1^*}{1-q_1^*} \frac{1-q_1^{-i*}}{q_1^{-i*}}))^2]}{q_1^*q_1^{-i*}(1 - q_1^*q_1^{-i*}) - q_1^*q_1^{-i*}(1 - q_1^*)(1 - q_1^{-i*})(q_1^* + q_1^{-i*} + 2q_1^*q_1^{-i*})}, 0\} \\ \frac{1}{2(q_1^* - q_1^{-i*})} \log(\frac{1-p_1}{p_1} \frac{p_1^{-i}}{1-p_1^{-i}} \frac{q_1^*}{1-q_1^*} \frac{1-q_1^{-i*}}{q_1^{-i*}}) \quad (55)$$

Both workers' and firms' behaviours are dependent on the default choice probabilities. Interestingly, the relative difference between workers' default probabilities, rather than the actual default probabilities, have direct impact on wages.

**Lemma 3.1** (Existence of Asymmetric Subgame Equilibrium with Heterogeneous Default Strategies).

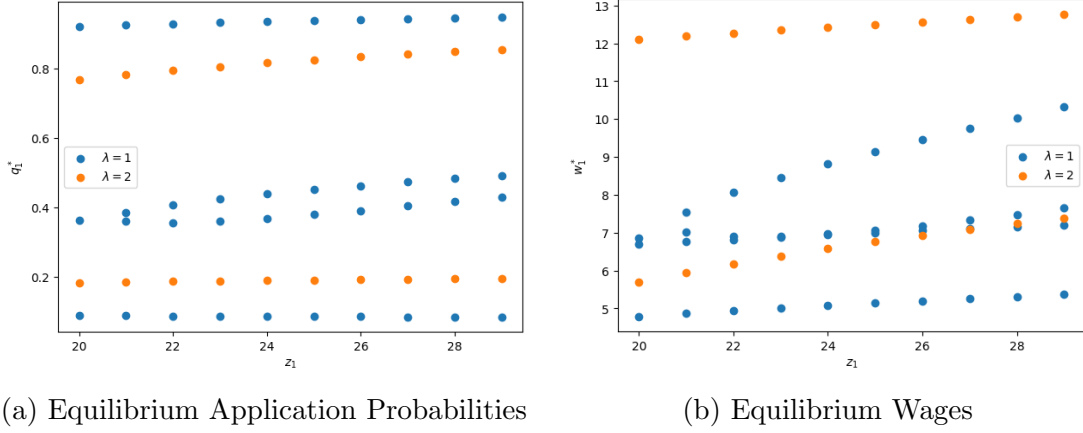
*Given any wage announcement  $(w_1, w_2)$ , there always exist at least one asymmetric subgame equilibrium (ASE) for  $\lambda > 0$ . On top of wage conditions from Lemma 2.2, more extreme asymmetry in strategies are likely to exist when workers are strongly biased to different firms; and greater relative belief distortion may introduce additional dynamical nuances and result in an expanded set of equilibria.*

*Proof.* Shown in Appendix A.6.

I prove there exists at least one equilibrium solution, and at extreme values of  $\lambda$ , workers will only apply symmetrically when information is costless, but as long as  $\lambda > 0$ , they apply asymmetrically, and as information becomes extremely expensive, workers stick to playing defaults, which is inherently asymmetric. Apart from  $\lambda$ , high degree of worker heterogeneity, reflected via extremely biased defaults could lead to workers applying more to different firms in the equilibrium. I also show using implicit function theorem that heterogeneous defaults could affect equilibrium multiplicity, particularly when there is relative belief distortion. There can

<sup>7</sup> $q_1$  and  $q_1^{-i}$  refers to  $q_1^*$  and  $q_1^{-i*}$  for wage determination.

be increased analytical complexities, and the current set of solutions may not fully capture the entire range of equilibria that are possible to exist within the system.  $\square$



Figures show for  $\lambda = 1$ ,  $\lambda = 2$ , and  $z_1$  ranging from 20 to 30,  $z_2 = 20$ ,  $p_1 = 0.61$ ,  $p_1^{-i} = 0.4$ .

Figure 8: Effects of Varying Productivity Disparity on Equilibrium Solutions

The absence of symmetric equilibrium suggests that workers do not employ the same application strategy when they are heterogeneously inattentive. It is therefore necessary to consider asymmetric equilibria in order to accommodate variability in workers' information availability. Figure 8 illustrates the possibility of several sets of  $(q_1^*, q_1^{-i*}, w_1^*, w_2^*)$  for each  $(z_1, z_2, \lambda, p_1, p_1^{-i})$ . The number of equilibria could be higher when information costs are low. While the number of equilibria are not of peculiar interest, it highlights that heterogeneous defaults could be a potential reason behind the greater variability in application strategies, contributing to unpredictability of the matching outcome.

**Proposition 5** (Heterogeneous Defaults on Equilibrium Strategies). *At low cognitive costs, workers' equilibrium strategies are influenced more by their opponent's default strategy than their own, and the reverse happens at high cognitive costs.*

*Proof.* Shown in Appendix A.7.  $\square$

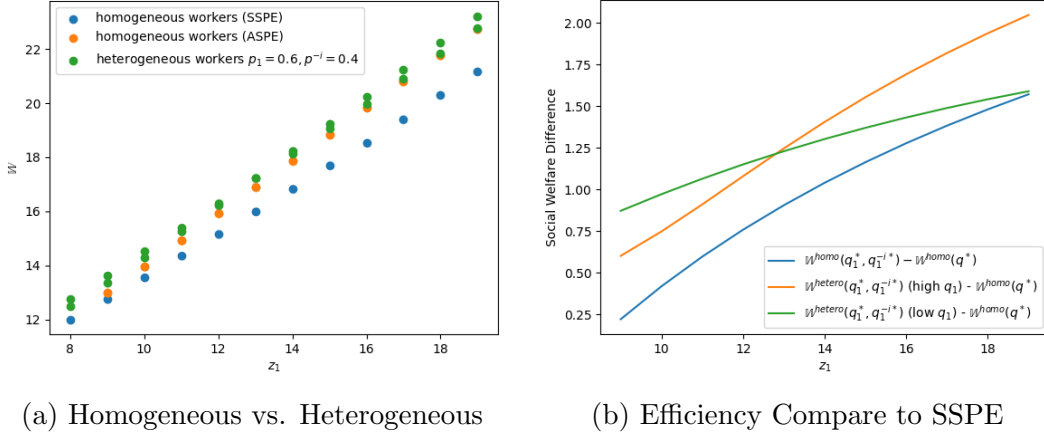
Workers can be more strategic and factoring in their opponent's biases when it is inexpensive to acquire information, and they would be more inclined to follow their own innate biases when information costs increase. This relates strongly to reality, where workers hardly change the type of firms they apply to, and their decisions could be as a result of deep-seated beliefs about the wages as highlighted in Archer (2016), and Briel et al. (2022) suggests such beliefs may even be gender-specific. However, having strong opposite beliefs could pay off as it may serve as a natural mechanism in facilitating better coordination and mitigate overcrowding issues.

This additional behavioural complication in workers' search behaviour could inevitably affect firms' wage setting. While the general pattern regarding wage sensitivity remain the same as the derivatives with respect to  $w_1, w_2$  (i.e. (31), (32), (33), (34)) are independent of  $p_1, p_1^{-i}$ , firms will be dealing with workers who have different responsiveness to expected payoff, characterized by the term  $\frac{1-p_j^i}{p_j^i}$  in (52) and (53). Figure 8b shows that some equilibria, particularly those with equilibrium strategies that differ significantly from default, could lead to higher equilibrium wages. Intuitively, to induce higher application rate from a worker that does not have initial bias towards it, firms need to set higher wage.

### Social Welfare with Heterogeneous Defaults.

$$\mathbb{W}^{hetero}(q_1^*, q_1^{-i*}) = (1 - (1 - q_1^*)(1 - q_1^{-i*}))z_1 + (1 - q_1^* q_1^{-i*})z_2 - \lambda \left[ \sum_{j=1}^2 q_j^* \log \frac{q_j^*}{p_j} + \sum_{j=1}^2 q_j^{-i*} \log \frac{q_j^{-i*}}{p_j^{-i}} \right] \quad (56)$$

Since there are no SSPE for  $\lambda > 0$ , instead of comparing social welfare between different equilibrium types, I compare between equilibrium outcomes for heterogeneous workers (ASPE) and homogeneous workers (ASPE & SSPE) to illustrate the impact of heterogeneous defaults.



Figures show for  $\lambda = 1$ ,  $z_1 \in (8, 20)$ ,  $z_2 = 8$ ,  $p = 0.6$  for homogeneous workers and  $p_1 = 0.6$ ,  $p_1^{-i} = 0.4$  for heterogeneous workers.

Figure 9: Social Welfare with Defaults Variation

Figure 9a shows an example where social welfare under SSPE is persistently lower than ASPE with homogeneous workers (as per Proposition 4). The inclusion of heterogeneous defaults leads to higher social welfare than both equilibrium types with homogeneous workers, which could be due to higher possibility of one-to-one matching arising from greater asymmetry in equilibrium strategies. However, such result may depend on equilibrium selection. As shown in Figure 9b, though  $\mathbb{W}^{hetero}(q_1^*, q_1^{-i*}) > \mathbb{W}^{homo}(q^*)$  for the ASPE with low  $q_1^*$ , the difference grow smaller. Given the trajectory of change as firm heterogeneity increases,  $\mathbb{W}^{homo}(q_1^*, q_1^{-i*})$  could surpass  $\mathbb{W}^{hetero}(q_1^*, q_1^{-i*})$ , indicating that heterogeneous defaults could also contribute to efficiency loss.

*Extreme Cases.* As  $\lambda \rightarrow 0^+$ , the game converge to one with complete information, and workers adopting different strategies would generate higher social welfare than them using the same strategies due to higher chances of one-to-one matching. And as  $\lambda$  increases, having some degree of heterogeneity in defaults might even be beneficial as the amount of information required to direct workers to different firms can be much smaller than when workers are homogeneous and biased in the same direction. For  $\lambda \rightarrow \infty$ , workers do not acquire any information, social welfare would depend on the primitive default value. Figure 10 shows that when facing firms of the same productivity, if workers are homogeneous and apply symmetrically in the equilibrium, then welfare is maximized when  $q^* = \frac{1}{2}$ , and  $q^* = \frac{z_1}{z_1 + z_2}$  if productivities differ. This is where equilibrium application probability matches the productivity distribution. However, if workers hold different defaults, assuming degree of belief distortion is the same, then when facing homogeneous firms, any pair of default strategies (i.e.  $p_1 \neq p_1^{-i}$ ), will lead to higher social welfare than homogeneous defaults. When productivities differ, welfare is higher if  $p_1^*$  and  $p_1^{-i*}$  are sufficiently disperse, such that the lower application rate received by the more productive firm is compensated by higher probability of one-to-one matching. Although this case is less interesting because workers' equilibrium application strategies are completely subjected to

pre-defined default parameter, but it highlights the possibility that efficiency can be completely exogenous in high cognitive costs environment.

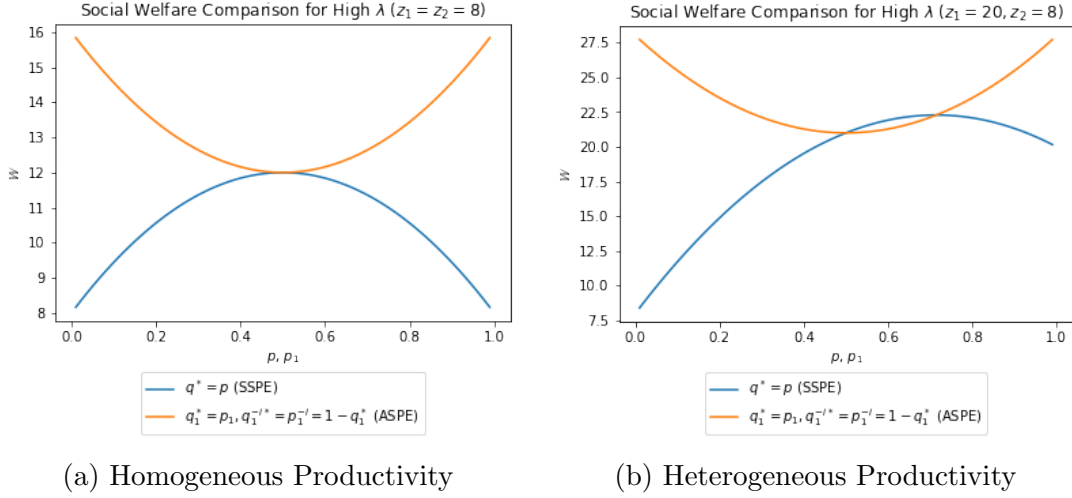


Figure 10: Social Welfare at High Cognitive Costs

### 3.2 Worker Heterogeneity in Cognitive Costs

Workers could also differ in information processing capability, which translate into heterogeneity in attention costs. In this subsection, assuming default strategies to be the same across workers, I explore how variations in cognitive costs could affect equilibrium outcomes.

**Model Assumption.** I relax the assumption of homogeneous cognitive costs, such that  $\lambda_i = [\lambda_1, \lambda_2]$ ,  $\lambda \in \Lambda$ , and  $\lambda_i$  can take different values. Henceforth, I suppose  $\lambda_1 > \lambda_2$ , worker 1 has higher cognitive cost than worker 2. The default strategy is random search as in Wu (2024).

**SSPE.** From workers' FOCs, (5) and (6), if workers are using the same application strategy,  $q_1 = q_1^{-i} = q$ , then  $\lambda_1 \log(\frac{q}{1-q}) = \lambda_2 \log(\frac{q}{1-q})$ . Given that  $\lambda$ s differ, the only possible equilibrium strategies would be random search, such that  $q^* = \frac{1}{2}$ ,  $w_1 = w_2 = w^*$ . Based on workers' and firms' FOCs,

$$w_1 = z_1 - \frac{(w_1 + w_2)^2 q^2 (1 - q) - 4\lambda_1 \lambda_2 \frac{1}{1-q}}{2(1 - q)^2 q (w_1 + w_2) - 2(1 - q)(\lambda_1 + \lambda_2)} \quad (57)$$

$$w_2 = z_2 - \frac{(w_1 + w_2)^2 (1 - q)^2 q - 4\lambda_1 \lambda_2 \frac{1}{q}}{2q^2 (1 - q)(w_1 + w_2) - 2q(\lambda_1 + \lambda_2)} \quad (58)$$

When  $q^* = \frac{1}{2}$ ,

$$w_1^* = z_1 - \frac{(w_1^* + w_2^*)^2 - 64\lambda_1 \lambda_2}{2(w_1^* + w_2^*) - 8(\lambda_1 + \lambda_2)}, w_2^* = z_2 - \frac{(w_1^* + w_2^*)^2 - 64\lambda_1 \lambda_2}{2(w_1^* + w_2^*) - 8(\lambda_1 + \lambda_2)}$$

Since  $w_1^* = w_2^* = w^*$ , for positive equilibrium wages,  $z_1 = z_2$  needs to hold, firms must be homogeneous. If firms are heterogeneous, the more productive firm would want to post higher wage relative to the other firm to attract higher application rate. Correspondingly, the more attentive worker would incur a lower cost to learn about the expected payoff as compared to the other worker, thus workers would not behave similarly. Therefore, the only possible SSPE solution  $(q^*, w^*)$  is

$$q^* = \frac{1}{2}, w^* = \max\left[\frac{(\lambda_1 + \lambda_2 + 0.5z) \pm \sqrt{(\lambda_1 + \lambda_2 + 0.5z)^2 + 32\lambda_1 \lambda_2 - 4\lambda_1 z - 4\lambda_2 z}}{2}, 0\right] \quad (59)$$

where  $z \geq w^* \geq 0$ .

Even though in SSPE,  $q^*$  is independent of  $\lambda_1$  and  $\lambda_2$ , they have an impact on the equilibrium wages. Higher the cognitive costs, lower the equilibrium wage. Firms can extract higher markdowns if the combined cognitive costs are high, the less cognitively constraint worker will be penalized even if he/she is substantially more attentive than the other.

**Proposition 6** (Heterogeneous Cognitive Costs on Equilibrium Wages). *In SSPE, cognitive cost of the less attentive worker has a larger impact on the equilibrium wages than the other worker.*

*Proof.* Shown in Appendix A.8. □

**ASPE.** The ASPE solution  $(q_1^*, q_1^{-i*}, w_1^*, w_2^*)$  can be solved with the following system of equations.

$$q_1^* = \frac{1}{1 + \exp(-(\frac{(2-q_1^{-i*})w_1^* - (1+q_1^{-i*})w_2^*}{2\lambda_1}))} \quad (60)$$

$$q_1^{-i*} = \frac{1}{1 + \exp(-(\frac{(2-q_1^*)w_1^* - (1+q_1^*)w_2^*}{2\lambda_2}))} \quad (61)$$

$$w_1^* = \max\{z_1 - \frac{(q_1^* + q_1^{-i*} - q_1^* q_1^{-i*})[\lambda_1 \lambda_2 - (\frac{1}{q_1^* - q_1^{-i*}}(\lambda_1 \log(\frac{q_1^*}{1 - q_1^*}) - \lambda_2 \log(\frac{q_1^{-i*}}{1 - q_1^{-i*}})))^2 q_1^* q_1^{-i*} (1 - q_1^*)(1 - q_1^{-i*})]}{(1 - q_1^{-i*})(1 - q_1^*)(q_1^* + q_1^{-i*} - q_1^* q_1^{-i*}) - \frac{1}{2(q_1^* - q_1^{-i*})}(\lambda_1 \log(\frac{q_1^*}{1 - q_1^*}) - \lambda_2 \log(\frac{q_1^{-i*}}{1 - q_1^{-i*}})) q_1^* q_1^{-i*} (1 - q_1^*)(1 - q_1^{-i*})(2q_1^* q_1^{-i*} - 3q_1^* - 3q_1^{-i*} + 4)}}, 0\}$$

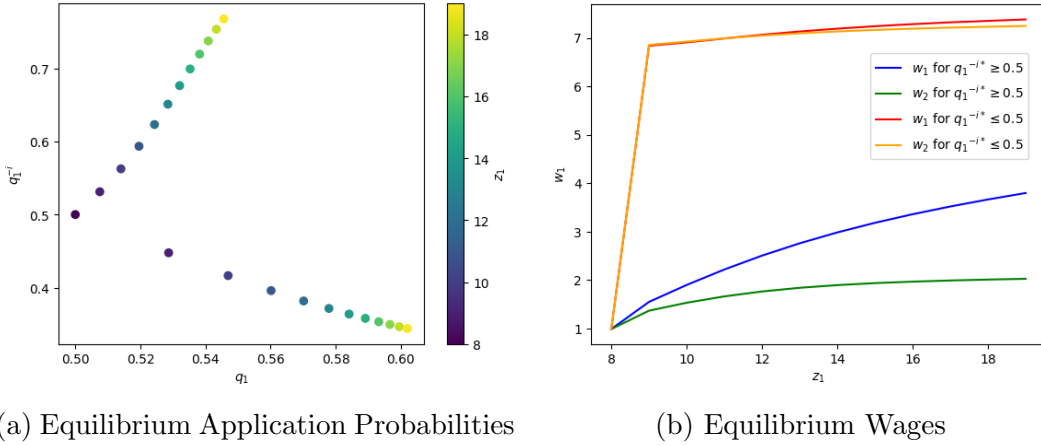
$$w_2^* = \max\{z_2 - \frac{(1 - q_1^* q_1^{-i*})[\lambda_1 \lambda_2 - q_1^* q_1^{-i*} (1 - q_1^*)(1 - q_1^{-i*})(\frac{1}{q_1^* - q_1^{-i*}}(\lambda_1 \log(\frac{q_1^*}{1 - q_1^*}) - \lambda_2 \log(\frac{q_1^{-i*}}{1 - q_1^{-i*}})))^2]}{q_1^* q_1^{-i*} (1 - q_1^* q_1^{-i*}) - q_1^* q_1^{-i*} (1 - q_1^*)(1 - q_1^{-i*})(q_1^* + q_1^{-i*} + 2q_1^* q_1^{-i*}) - \frac{1}{2(q_1^* - q_1^{-i*})}(\lambda_1 \log(\frac{q_1^*}{1 - q_1^*}) - \lambda_2 \log(\frac{q_1^{-i*}}{1 - q_1^{-i*}})) q_1^* q_1^{-i*} (1 - q_1^*)(1 - q_1^{-i*})}}, 0\} \quad (62)$$

**Lemma 3.2** (Existence of Asymmetric Subgame Equilibrium with Heterogeneous Cognitive Costs). *For a given set of  $(w_1, w_2)$ , there always exist at least one subgame equilibrium. The necessary condition for asymmetric subgame equilibrium (ASE) to exist is small  $\lambda_1$  and  $\lambda_2$  relative to expected payoff, and that wages need to fulfill  $2w_2 > w_1 > \frac{w_2}{2}$ .*

*Proof.* Shown in Appendix A.9. □

The condition on  $\lambda$  could also influence the number of equilibria. Decreasing the aggregate  $\lambda$ s or lowering one of the  $\lambda$ s could lead to multiple equilibrium solutions, and the more extreme asymmetric strategies, where workers apply more to different firms, are also more likely to exist. In the example shown in Figure 11, when firms have the same productivities, workers respond by applying symmetrically. As firm 1's productivity raises relative to firm 2's, there could be 2 ASPEs: One where both workers apply with high probability to firm 1, but the more attentive worker, worker 2, applies at a higher rate; the other consists of worker 1 applying with higher





(a) Equilibrium Application Probabilities

(b) Equilibrium Wages

Figures show for  $\lambda_1 = 3$ ,  $\lambda_2 = 1$ ,  $z_1$  ranging from 8 to 20,  $z_2 = 8$ ,  $p_1 = p_1^{-i} = 0.5$ .

Figure 11: Effects of Varying Productivity on Equilibrium Probabilities and Wages

probability to firm 1 and worker 2 to firm 2. The first is fairly intuitive. Worker 2 can take advantage of her lower cost in acquiring and processing information as compared to worker 1, and apply with higher probability to the more productive firm. The second maybe slightly counter-intuitive. In Figure 11b, the equilibrium wages correspond to this type of equilibrium has smaller wage differentials, the benefit of “avoiding” the other worker is thus larger than trying to compete for the higher wage. Therefore, worker 2 can be perceived to be more strategic in choosing to apply with higher probability to the lower wage firm.

**Proposition 7** (Heterogeneous Cognitive Cost on Equilibrium Strategies). *For small wage differential, the more attentive worker is likely to avoid overcrowding by applying more to the less productive firm; For large wage differential, the more attentive worker is likely to take advantage of the lower information cost and apply more to the more productive firm.*

*Proof.* Shown in Appendix A.10. □

The changes in workers’ search behaviour could similarly affect firms’ wage setting. Unlike in Section 3.1, the derivatives (31), (32), (33), (34), and thus wage sensitivity is affected by  $\lambda$ s.

$$\frac{dq_1}{dw_1} = \frac{(w_1 + w_2)(2 - q_1)q_1^{-i}(1 - q_1^{-i})q_1(1 - q_1) - 2\lambda_2(2 - q_1^{-i})(1 - q_1)q_1}{(w_1 + w_2)^2q_1^{-i}(1 - q_1^{-i})q_1(1 - q_1) - 4\lambda_1\lambda_2} \quad (64)$$

$$\frac{dq_1^{-i}}{dw_1} = \frac{(2 - q_1^{-i})(1 - q_1)q_1(1 - q_1^{-i})q_1^{-i}(w_1 + w_2) - 2\lambda_1q_1^{-i}(1 - q_1^{-i})(2 - q_1)}{(w_1 + w_2)^2q_1^{-i}(1 - q_1^{-i})q_1(1 - q_1) - 4\lambda_1\lambda_2} \quad (65)$$

$$\frac{dq_1}{dw_2} = \frac{2\lambda_2(1 + q_1^{-i})q_1(1 - q_1) - q_1q_1^{-i}(1 - q_1)(1 - q_1^{-i})(1 + q_1)(w_1 + w_2)}{q_1q_1^{-i}(1 - q_1)(1 - q_1^{-i})(w_1 + w_2)^2 - 4\lambda_1\lambda_2} \quad (66)$$

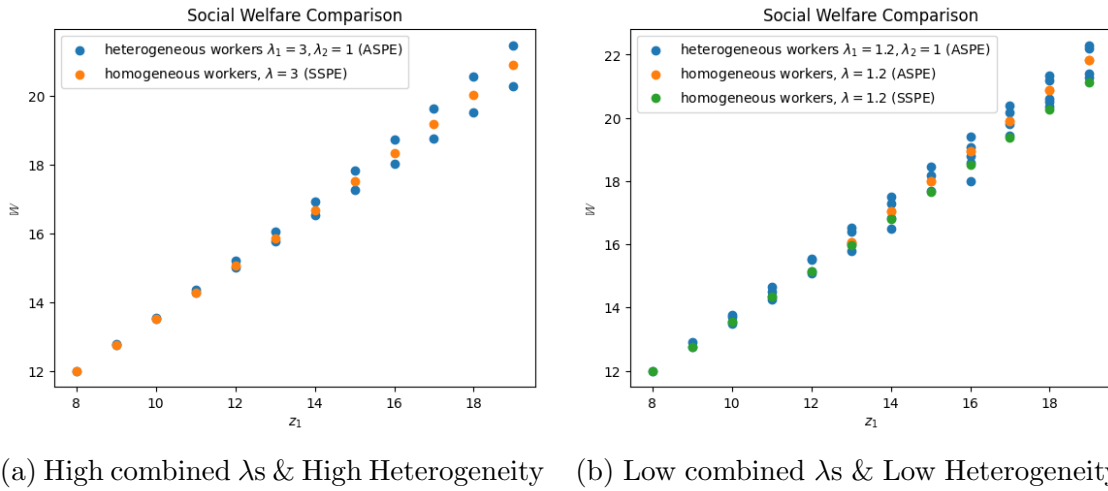
$$\frac{dq_1^{-i}}{dw_2} = \frac{2\lambda_1(1 + q_1)q_1^{-i}(1 - q_1^{-i}) - q_1q_1^{-i}(1 - q_1)(1 - q_1^{-i})(1 + q_1^{-i})(w_1 + w_2)}{q_1q_1^{-i}(1 - q_1)(1 - q_1^{-i})(w_1 + w_2)^2 - 4\lambda_1\lambda_2} \quad (67)$$

If  $\lambda$ s are substantially varied, and  $\lambda_2 \rightarrow 0^+$ , then  $\frac{dq_1}{dw_1} > \frac{dq_1^{-i}}{dw_1} > 0$ ,  $\frac{dq_1}{dw_2} < \frac{dq_1^{-i}}{dw_2} < 0$ . Worker 2, who incur lower information cost, may be less reactive to wage changes than worker 1. It may even be possible for  $\frac{dq_1}{dw_1} < 0$  and  $\frac{dq_1^{-i}}{dw_2} > 0$ , worker 2 apply less to firm 1 when firm 1’s wage increase and more when firm 2’s wage increases. This portrays the possibility of worker 2 having more strategic concerns. Knowing increase in  $w_1$  will induce the more cognitively constrained worker to apply to firm 1 with higher probability, worker 2 would instead increase application

probability to firm 2 to mitigate overcrowding. Figure 11b shows that equilibrium wages differ across ASPEs, and the equilibrium with more asymmetric search strategies could induce higher wages. The bulk of this is still attributed to the perceived scarcity of labour, which leads to aggressive initial wage setting behaviour, but the presence of heterogeneity in cognitive costs could exaggerate this effect. For example, if firm 2 believes worker 1 is almost certain to apply to firm 1, it may set a high initial wage to secure worker 2, but this could also entice worker 1, which might cause worker 2 to switch firm in response, and to prevent this, firm 2 would need to set an even higher wage to retain worker 2's attraction. Firm 1 would also set a high wage to avoid losing both workers and stay competitive. This is in contrast to workers applying more randomly, the initial wage do not need to be set as high as they believe the probability of meeting a worker is much greater than zero.

**Social Welfare with Heterogeneous Cognitive Costs.** When workers are heterogeneous, the only possible SSPE is when workers apply randomly, and welfare in such case is simply  $\frac{3}{2}z$ , which is the same as when workers have identical cognitive costs. For ASPE,

$$\mathbb{W}^{hetero}(q_1^*, q_1^{-i*}) = (1 - (1 - q_1^*)(1 - q_1^{-i*}))z_1 + (1 - q_1^* q_1^{-i*})z_2 - \lambda_1 \sum_{j=1}^2 q_j^* \log \frac{q_j^*}{p_j} - \lambda_2 \sum_{j=1}^2 q_j^{-i*} \log \frac{q_j^{-i*}}{p_j} \quad (68)$$



(a) High combined  $\lambda$ s & High Heterogeneity (b) Low combined  $\lambda$ s & Low Heterogeneity

Figures show homogeneous workers with  $\lambda = 3$ , and heterogeneous workers with  $\lambda_1 = 3, \lambda_2 = 1$  (LHS); Homogeneous workers with  $\lambda = 1.2$ , and heterogeneous workers with  $\lambda_1 = 1.2, \lambda_2 = 1$  (RHS).  $z_1 \in [8, 20]$ ,  $z_2 = 8$ ,  $p_1 = p_1^{-i} = 0.5$ .

Figure 12: Welfare Comparison between Homogeneous and Heterogeneous Workers

Figure 12 compares social welfare between homogeneous and heterogeneous workers. When  $\lambda$ s are high and diverse, social welfare can be either higher or lower than SSPE of homogeneous workers. In Figure 12a, the equilibrium where both workers apply with higher probability to the more productive firm 1, but at different rates, could in fact be slightly more efficient than when workers attempting to coordinate. This could be due to high combined information costs incurred to direct workers to different firms, which might outweigh the benefit of some degree of coordination. Figure 12b shows the comparison when  $\lambda$ s are low and close together. While SSPE is always less efficient than ASPE when workers are homogeneous, incorporation of heterogeneity could complicate the issue, social welfare could be higher or lower depending on equilibrium selection. In this case, the equilibrium with the higher social welfare comprises of workers applying more to different firms. This could be attributed to the low combined information costs, which makes the higher possibility of one-to-one matching more attractive. Though worker heterogeneity makes efficiency analysis more nuanced, it is noticeable in both

Figures that reducing cognitive cost for one of the worker could lead to equilibrium outcome(s) that generates higher social welfare. This suggests that policymakers does not necessarily need to lower information processing costs for the entire workforce, but equilibrium selection is a more important concern, highlighting the need to explore mechanisms to nudge individuals towards more efficient equilibrium. (Appendix B.4.)

## 4 Social Planner Problem

In this section, I would like to introduce a social planner to determine the social optimal application strategy. The difference between this and the equilibrium strategies can enlighten us about the efficiency level prevailing in the market with respect to the societal expectations. On top of Wu (2024)'s social planner problem, I relax the assumption on default choice probabilities, and account for the possibility of worker heterogeneity. There can be two types of social planner – a constrained and an unconstrained one.

**Definition 4.1** (Constrained social planner problem). *The constrained social planner problem involves maximizing social welfare by solving for one optimal application strategy for both workers, where  $\hat{q}$  is the socially optimal choice probability of firm 1.*

$$\max_{q \in [0,1]} (1 - (1-q)^2)z_1 + (1-q^2)z_2 - \lambda_1(q \log \frac{q}{p_1} + (1-q) \log \frac{1-q}{1-p_1}) - \lambda_2(q \log \frac{q}{p_1^{-i}} + (1-q) \log \frac{1-q}{1-p_1^{-i}}) \quad (69)$$

**Definition 4.2** (Unconstrained social welfare problem). *The unconstrained social planner problem involves solving for a set of socially optimal application strategies for each worker, and the solution consists of a tuple  $(\hat{q}_1, \hat{q}_1^{-i})$ , where  $\hat{q}_1$  is the optimal choice probability of firm 1 for worker 1 and  $\hat{q}_1^{-i}$  is that for worker 2, and there are no restriction on workers to use the same or different strategies.*

$$\begin{aligned} \max_{q_1, q_1^{-i} \in [0,1]} & (1 - (1-q_1)(1-q_1^{-i}))z_1 + (1 - q_1 q_1^{-i})z_2 \\ & - \lambda_1[q_1 \log \frac{q_1}{p_1} + (1-q_1) \log \frac{1-q_1}{1-p_1}] - \lambda_2[q_1^{-i} \log \frac{q_1^{-i}}{p_1^{-i}} + (1-q_1^{-i}) \log \frac{1-q_1^{-i}}{1-p_1^{-i}}] \end{aligned} \quad (70)$$

Even though analyzing the unconstrained social planner problem could provide some information for policymakers on the benefit of directing workers asymmetrically, the practicality and scalability of providing individual-specific advice could be costly, thus social welfare that can be achieved might need to be discounted. As a result, unconstrained social planner problem will not be discussed in this paper, but it is important to acknowledge this possibility. (Appendix C.1)

### 4.1 Constrained Social Welfare Problem

**Homogeneous Workers.** Assume workers are homogeneous (same  $\lambda$  and  $p$ ), FOC based on (69),

$$(1-q)z_1 - qz_2 - \lambda \log\left(\frac{1-p}{p} \frac{q}{1-q}\right) = 0 \quad (71)$$

Given exogenously determined value for  $z_1$ ,  $z_2$  and  $\lambda$ ,  $\hat{q}$  that maximizes the social welfare is

$$\hat{q} = \frac{1}{1 + \frac{1-p}{p} \exp\left(-\left(\frac{(1-\hat{q})z_1 - \hat{q}z_2}{\lambda}\right)\right)} \quad (72)$$

$$\frac{d\hat{q}}{dp} = \frac{\frac{\lambda}{(1-p)p}}{z_1 + z_2 + \frac{\lambda}{\hat{q}(1-\hat{q})}} > 0 \text{ for } \lambda > 0, p \in (0, 1) \quad (73)$$

Novel to Wu (2024)'s social planner problem, workers' default may deviate away from random, therefore, socially optimal strategy can be affected accordingly. Since  $\frac{d\hat{q}}{dp} > 0$ , larger belief distortion towards firm 1, higher socially optimal choice probability of firm 1.

The socially optimal welfare given  $\hat{q}$  is

$$\mathbb{W}(\hat{q}) = (1 - (1-\hat{q})^2)z_1 + (1 - \hat{q}^2)z_2 - 2\lambda(\hat{q} \log(\frac{\hat{q}}{p}) + (1-\hat{q}) \log(\frac{1-\hat{q}}{1-p})) \quad (74)$$

**Proposition 8** (Inefficient SSPE). *SSPE is defined as inefficient if  $q^* \neq \hat{q}$ . It is generally inefficient when (1) prior beliefs are incorrect (Definition 2.5); and (2) firms are heterogeneous in productivity.*

*Proof.* Shown in Appendix A.11.

In the proof, I show that when  $\lambda \rightarrow 0^+$ ,

$$\lim_{\lambda \rightarrow 0^+} \hat{q} = \frac{z_1}{z_1 + z_2} \quad (75)$$

$$\lim_{\lambda \rightarrow 0^+} q^* = \frac{2w_1^* - w_2^*}{w_1^* + w_2^*}, \text{ where } q^* \in (0, 1) \quad (76)$$

where  $w_1^*$  and  $w_2^*$  are determined by (17) and (18). As  $\lambda \rightarrow \infty$ ,

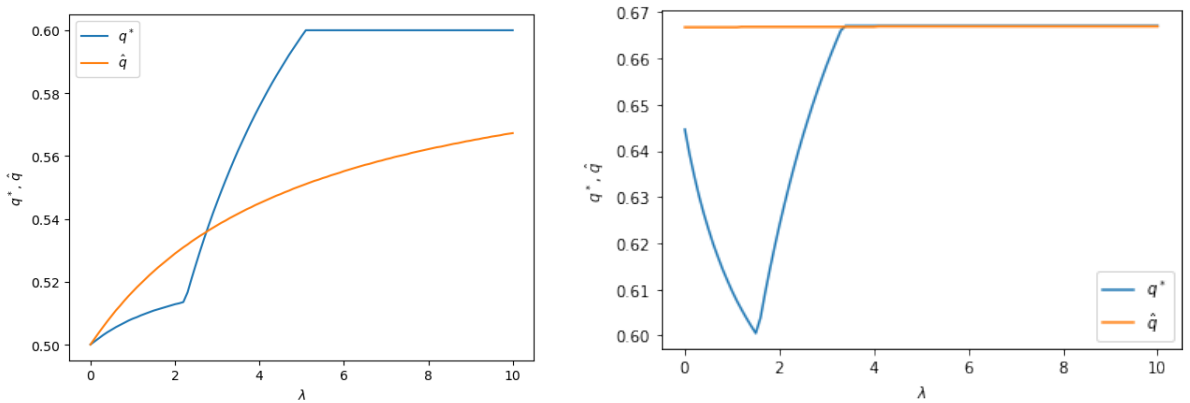
$$\lim_{\lambda \rightarrow \infty} \hat{q} = p, \lim_{\lambda \rightarrow \infty} q^* = p \quad (77)$$

If prior beliefs are accurate and firms are homogeneous, then  $q^* = \hat{q} = \frac{1}{2}$ , SSPE is efficient. However, any other variations could lead to inefficiency. If prior beliefs are incorrect, then as  $\lambda$  increases,  $q^*$  and  $\hat{q}$  converge to  $p$  at different rate, SSPE is inefficient as long as information is affordable. When facing heterogeneous firms, even if prior beliefs are correct, SSPE can still be inefficient. Even for  $\lambda \rightarrow 0^+$ ,  $q^* \neq \hat{q}$ . Also, as  $\hat{q}$  and  $q^*$  converge at different rate to  $p$  for intermediary  $\lambda$ s, there may be a threshold crossing point,  $\bar{\lambda}$ , where  $q^* = \hat{q}$ , depending on the direction of convergence.

$$\bar{\lambda} = \frac{z_1 - \frac{z_1 + (z_1 - w_1^*) - (w_1^* - w_2^*)}{z_1 + z_2 + (z_1 - w_1^*) + (z_2 - w_2^*)}(z_1 + z_2)}{\log\left(\frac{1-p}{p} \frac{z_2 + (z_2 - w_2^*) + (w_1^* - w_2^*)}{z_1 + z_2 + (z_1 - w_1^*) + (z_2 - w_2^*)}\right)} \quad (78)$$

It can be affected by default choice probability and firm heterogeneity.  $\square$

By social planner standard, workers are generally not learning the optimal wage precision, leading to inefficiency. As supplement to Proposition 8, I show 2 examples, where inefficiency could arise from incorrect prior beliefs and heterogeneous firms.



(a) Homogeneous Firms, Incorrect Defaults

(b) Heterogeneous Firms, Correct Defaults

Figures show for  $\lambda \in [0, 10]$ , (a)  $z_1 = 10, z_2 = 10, p > \frac{z_1}{z_1 + z_2} = 0.5$ , (b)  $z_1 = 10, z_2 = 5, p = \frac{z_1}{z_1 + z_2} = 0.667$ .

Figure 13: Comparison between Socially Optimal Solution  $\hat{q}$  and SSPE solution  $q^*$

Figure 13a shows while workers could start off searching efficiently when information cost is negligible, both  $q^*$  and  $\hat{q}$  are affected by incorrect defaults as  $\lambda$  increases, but at different

rates, and equilibrium strategy constantly deviates away from optimal, except at the crossing point. While inaccurate prior beliefs provide a compelling case that workers are affected by the additional cost to correct their initial bias, a natural question is why workers having correct prior beliefs could also arrive at inefficient outcomes when facing heterogeneous firms. Figure 13b demonstrates workers with correct defaults but face heterogeneous firms. Even at low information costs, the equilibrium outcome is inefficient. This maybe counter-intuitive as workers are expected to get more precise information when cost is low and thus generate more efficient outcome than when  $\lambda$  is high. The reason is because the model resembles a Stackelberg leadership model where firms set wages before workers make their moves, therefore, even if workers do not incur any attention costs, firms is able to extract a markdown as the first movers. While homogeneous firms offer the same wages in equilibrium, thus wage-productivity ratio remain the same (i.e.  $\frac{w_1^*}{w_1^*+w_2^*} = \frac{z_1}{z_1+z_2}$ ), when firms are heterogeneous, the more productive firm could extract a slightly higher markdown than the other, resulting in workers, basing their search on wages instead of productivities, to deviate away from the optimal strategy. Such form of inefficiency is inherit in the nature of market even if workers have correct defaults and are completely attentive. With the presence of attention costs, firms could penalize workers by setting lower wages, and workers are directed based on the observed wages. Only when information becomes sufficiently costly that workers no longer learn, then wages would have no impact, and workers would stick to their correct defaults and the outcome is efficient. This mechanism highlights the additional source of inefficiency caused by inattention. (Appendix B.5.)

The next step is to evaluate workers' asymmetric strategies in the equilibrium, against the socially optimal solution. Based on Proposition 4, ASPE is more efficient than SSPE, and for it to be more efficient than the constrained socially optimal outcome,  $\mathbb{W}(q_1^*, q_1^{-i*}) > \mathbb{W}(\hat{q})$ ,

$$\begin{aligned} & ((1 - \hat{q})^2 - (1 - q_1^*)(1 - q_1^{-i*}))z_1 + (\hat{q}^2 - q_1^*q_1^{-i*})z_2 \\ & > \lambda \left[ \sum_{j=1}^2 q_j^* \log \frac{q_j^*}{p_j} + \sum_{j=1}^2 q_j^{-i*} \log \frac{q_j^{-i*}}{p_j} - 2\hat{q} \log \left( \frac{\hat{q}}{p} \right) - 2(1 - \hat{q}) \log \left( \frac{1 - \hat{q}}{1 - p} \right) \right] \quad (79) \end{aligned}$$

Even though the conditions for  $\lambda$  are generally more binding for ASPE to exist, and it is likely for (79) to be fulfilled as long as ASPE exist, the analysis is analytically complex and may be subjected to specific values of the exogenous parameters. Intuitively, there could be more benefit from one-to-one matching than higher likelihood of more productive firm meeting a worker as having an analogous application strategy could contribute significantly to overcrowding.

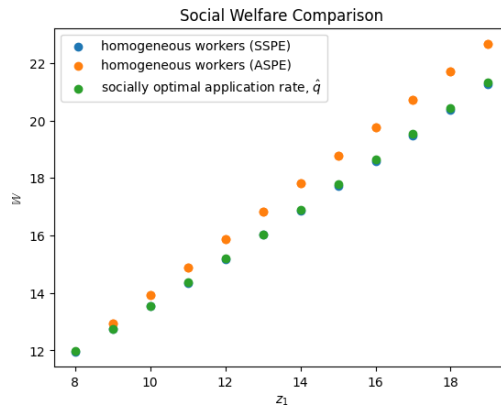


Figure show for homogeneous workers having  $\lambda = 1$ , figure shows  $z_1$  ranging from 8 to 20,  $z_2 = 8$ ,  $p_1 = p_1^{-i} = 0.5$ .

Figure 14: Comparison between Socially Optimal Solution, ASPE Solution and SSPE Solution

Figure 14 illustrates an example where constrained social planner is not as efficient as compared to ASPE, and SSPE is the least efficient. This highlights that focusing solely on comparison between SSPE and socially optimal strategy could be detrimental to social welfare analysis, as the market can be perceived to be more inefficient than it actually is.

**Heterogeneous Workers.** Based on the constrained social welfare problem (69), the FOC to determine socially optimal allocation  $\hat{q}$  considering for *heterogeneous defaults* is

$$2(1 - \hat{q})z_1 - 2\hat{q}z_2 - \lambda[\log(\frac{1 - p_1}{p_1} \frac{\hat{q}}{1 - \hat{q}}) + \log(\frac{1 - p_1^{-i}}{p_1^{-i}} \frac{\hat{q}}{1 - \hat{q}})] = 0 \quad (80)$$

where  $p_1 \neq p_1^{-i}$ , and as before, assume  $p_1 > \frac{1}{2}$  and  $p_1^{-i} < \frac{1}{2}$ . Expressing (80) in terms of  $\hat{q}$  on RHS,  $p_1$  and  $p_1^{-i}$  on LHS, LHS depends on the relative belief distortion (Definition 3.1), changes in which would affect  $\hat{q}$ .

$$2z_1 - \lambda \log(\frac{1 - p_1}{p_1} \frac{1 - p_1^{-i}}{p_1^{-i}}) = 2\lambda \log(\frac{\hat{q}}{1 - \hat{q}}) + 2\hat{q}(z_1 + z_2) \quad (81)$$

**Proposition 9** (Heterogeneous Defaults on Socially Optimal Strategy). *Socially optimal solution,  $\hat{q}$ , is positively affected by  $p_1$  and  $p_1^{-i}$ . Worker with stronger belief distortion can influence the socially optimal strategy in his/her favoured direction.*

*Proof.* Shown in Appendix A.12. □

Another aspect of analysis lays in *heterogeneous cognitive costs*, the FOC of the constrained social welfare problem that incorporates  $\lambda_1 \neq \lambda_2$ , and as before, assuming  $\lambda_1 > \lambda_2$ , is

$$2(1 - \hat{q})z_1 - 2\hat{q}z_2 - (\lambda_1 + \lambda_2) \log(\frac{\hat{q}}{1 - \hat{q}}) = 0 \quad (82)$$

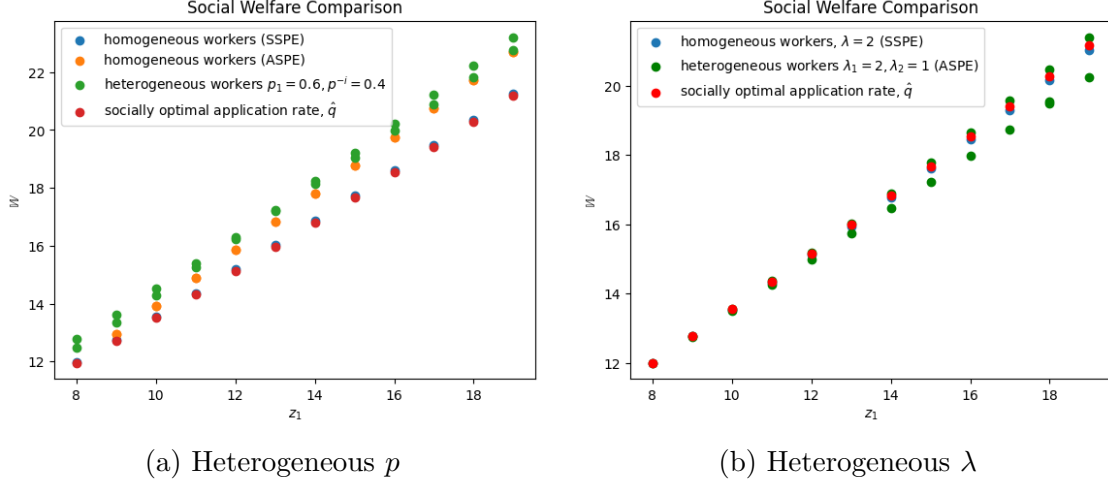
$$\hat{q} = \frac{\exp \frac{2(1 - \hat{q})z_1 - 2\hat{q}z_2}{\lambda_1 + \lambda_2}}{1 + \exp \frac{2(1 - \hat{q})z_1 - 2\hat{q}z_2}{\lambda_1 + \lambda_2}} \quad (83)$$

**Proposition 10** (Heterogeneous Cognitive Costs on Socially Optimal Strategy). *Socially optimal solution,  $\hat{q}$ , is more negatively affected by the worker with higher cognitive cost.*

*Proof.* Shown in Appendix A.13. □

Proposition 9 and 10 demonstrate that the less attentive worker, who either possess a stronger belief distortion or higher cognitive costs that makes information acquisition and processing more expensive, is able to influence the socially optimal strategy more. Furthermore, they also show that without accounting for worker heterogeneity, evaluating equilibrium efficiency against the socially optimal benchmark that assumes worker homogeneity could lead to over or under-estimation of inefficiency, contributing to misguided policies.

Figure 15 illustrates some examples of welfare comparison. With the incorporation of worker heterogeneity, through diverse priors or reducing cognitive costs of one of the workers, could lead to higher upper-bound of social welfare being achieved as compared to equilibrium with homogeneous workers and adjusted- $\hat{q}$  benchmarks (that included worker heterogeneity), but this comparison is intricate due to the possibility of multiple equilibria and could depend on equilibrium selection. In Figure 15a, equilibrium choices with heterogeneous workers could generate better outcome than homogeneous workers and the socially optimal strategy. This can be due to higher possibility of one-to-one matching when workers apply asymmetrically



(a) Figure show for homogeneous workers,  $\lambda = 1$ ,  $p_1 = p_1^{-i} = 0.6$ ; and heterogeneous workers,  $p_1 = 0.6, p_1^{-i} = 0.4$ . (b) Figure show for homogeneous workers,  $\lambda = 2$ ,  $p_1 = p_1^{-i} = 0.5$ , herein, ASPE do not exist; and heterogeneous workers,  $\lambda_1 = 2, \lambda_2 = 1$ . For all,  $z_1 \in [8, 20]$ ,  $z_2 = 8$ .

Figure 15: Social Welfare Comparison between Equilibrium and Socially Optimal Outcomes

than adopting analogous application strategies. But interestingly, harmonizing workers' default strategies might generate better outcome than the adjusted- $\hat{q}$  benchmarks, implying there may be efficiency gain from such policy intervention. When comparing equilibrium choices made by workers with different cognitive costs against adjusted- $\hat{q}$  benchmarks in Figure 15b, the results is indeterminate. However, if both workers have identical and high cognitive costs, social welfare will be lower than adjusted- $\hat{q}$  benchmarks, implying potential benefit of reducing cognitive costs for a subset of worker population. (Appendix A.14)

The above analysis highlights that if the efficiency benchmark has correctly accounted for worker heterogeneity, not only can it accurately assess the efficiency level of the prevailing equilibrium, it also provide basis for more policy discussions. There could be benefits from accommodating or introducing some degree of heterogeneity into the worker population, which makes it possible for more asymmetric strategies to arise that can potentially be more efficient than when workers are homogeneous and the adjusted- $\hat{q}$  benchmark. But since this can depend on equilibrium selection, it also stipulate future discussion on mechanisms to nudge workers to coordinate on the "correct" equilibrium. Furthermore, the adjusted- $\hat{q}$  also highlights a pitfall of constrained social planner, where picking a single optimal level,  $\hat{q}$ , to assess market efficiency can be insufficient. Heterogeneous workers could be applying asymmetrically and generate higher social welfare than the socially optimal solution, but these strategies is considered "inefficient", if defined by its deviation away from the socially optimal level. Measures that caters to this target may in fact lead to lower social welfare than leaving workers' application decisions to the market.



## 5 Discussions

**Bridging Theoretical and Empirical Studies.** Earlier in the paper, I highlighted the limitation of existing theoretical model to describe narrow search behaviour, followed by broadening or sometimes, narrower search behaviour after job centre recommendation, that was put forth by Belot et al. (2019). This divergence can be explained by the GPDS model. For instance, in the simplest case where workers are assumed to be homogeneous, by lifting the restriction on workers adopting identical application strategies in the equilibrium and consider for the possibility of multiple equilibria, GPDS would predict workers to be able to assign different choice probability to the firms, and thus by definition, workers apply narrowly. Since workers in the survey are not pre-specified to only follow the same application strategy, this theoretical model provides a more realistic portrayal of the possible equilibrium outcomes. Adding a social planner, Wu (2024) indicates workers would use a more directed search strategy, whereas Belot et al. (2019) suggests there would be a broadening of search, or sometimes, narrower if the original strategy is relatively broad. This can be perceived as workers being nudged to switch away from asymmetric or symmetric application strategies to a socially optimal strategy, which can either be less or more targeted. By accounting for the possibility of equilibrium multiplicity, the GPDS model would be able to accommodate both results. While it can be argued that the job-seekers that use the job centre services may have high attention costs, and homogeneous workers only behave asymmetrically at fairly low cognitive costs relative to expected payoffs, then simply considering for multiple equilibrium may not be enough as only SSPE emerge under such circumstances. The GPDS model has the additional benefit of considering for workers heterogeneity in attention costs, in which case, ASPE could exist even at high cognitive costs. Any suggestions by job centres to switch away from such equilibrium strategies to a socially optimal one may imply a move from more targeted search to a broadening of scope. The GPDS model is shown to be more generalizable, and might be able to provide slightly more accurate predictions of workers' behaviour and the labour market outcomes.

**Beauty of Small Set-up.** Although GPDS model is a reduced-form model, focused on a simplified  $2 \times 2$  setting, it is effective in painting a picture that mismatch could prevail even with 2 workers and 2 firms, but the silver lining is that some equilibrium may be more efficient than others even if information frictions cannot be completely eliminated. While increasing the number of workers and firms could enlighten us on the labour market as a whole, it is expected that the conclusions drawn from this work remains, but extrapolating this to a multi-agent setting expands the dimensions of mismatch problem, particularly if the number of workers does not match up with the number of vacancies. Furthermore, the small set-up also fits particularly well with the idea of consideration sets put forth by Hauser and Wernerfelt (1990) and Caplin et al. (2019), where workers only consider a subset of the firms when making their application decisions. Drawing parallel to this model, workers could place 0 probability on certain options by default, and reduce their choice sets to a few firms. Since workers are applying with positive probability to all the options in their consideration sets, if the consideration sets overlap, then workers are essentially considering the same set of firms, akin to the set-up portrayed in this paper. Future discussions could be oriented on the overlapping of consideration sets. It can be hypothesized that the less the consideration sets overlap, the more likely workers will be matched up with a firm. There can be mechanism to partition job search interface by showing workers different subsets of job postings to reduce mismatch problem. However, if workers are given many options, they can only accept one job, other firms would obtain 0 payoffs, which is neither beneficial for the firms nor for the society; and if workers are given one option each, though this improves matching, there is a lack of choices, and place a huge responsibility on the social planner to identify the "perfect match" for each worker. As long as the consideration set includes more than one option, the findings from this paper remain applicable.

**Parallels to Other Empirical Studies.** Apart from the literature that motivates the research question, the GPDS model can also be used to describe other empirical results. For instance, Altmann et al. (2018) has conducted a field experiment to explore the impact of targeted information dissemination by handing out brochures that include information about the general labour market conditions. It was found that such measure could increase employment. This can be represented in the GPDS model as lowering the unit information processing cost (or cognitive cost,  $\lambda$ ) for a group of job-seekers through compressing essential information in one place and making it easier to acquire and process. The model predicts possibility of multiple equilibria, including one(s) where workers apply more to different firms. There could be higher chances of matching and social welfare can be improved. Nonetheless, in Abebe et al. (2020), the impact of randomized job-fair interventions on labour market outcome was analyzed. It was found that such events lowers workers' expectations and correct their beliefs about the labour market, as a result, creating more matches at the end, especially for the less educated. This result can also be derived from the GPDS model. When cognitive costs are high, workers are fully reliant on their default strategies. By matching the default strategies with the actual underlying productivity distribution, more efficient outcome can be achieved.

**Policy Implications.** This framework offers opportunities for policy interventions that targets information presentation, rather than just reducing time or monetary search costs. A possible measure to mitigate information friction would be imposing stricter pay transparency laws. For instance, Europa (2023) states that EU has passed the Pay Transparency Directive in 2023, which require companies to inform job seekers about the starting salary or pay range of the advertised positions, either in the vacancy notice or ahead of interview. While the main aim of such policy is to detect and reduce gender pay gap, this model postulates that disclosing wage range, or better, the exact wages, during the advertising stage has the additional benefit of lowering information friction, thereby leading to more efficient outcomes. However, since attention costs in job search can also be interpreted as workers observing the exact wages but failed to process the information or have perfect recall of the precise value when making decisions, pay transparency laws would not be helpful in reducing information friction in this aspect, and completely elimination of attention costs through such policy would be implausible. Furthermore, the GPDS approach models the mental cost of overriding default behaviors, capturing behavioral biases as well as cognitive limitations. Since workers might not enter labour market as a blank slate, and mismatch can be deconstructed to a consequence of innate bias, policies to harmonize prior beliefs to match with market conditions could bring about more efficient outcome. Nonetheless, the possibility of some equilibria generating higher wages and more efficient outcome advocates a need for exploring equilibrium selection measures, which can be a policy area to further explore.

**Limitations.** The GPDS model is not without its limitations. First of all, the theory is silent on what exact information workers learn and how they learn, while this allows for more flexible interpretations, policy formulations based on misinformed supposition might not lead to desired effect. Furthermore, there are no distinctions between cost of searching, processing, retaining and retrieval of information, these are all bundled together. While it is possible to interpret attention costs in many different ways, which makes the model more generalizable to explore workers' utility maximization behaviour subjecting to information constraints. It might be worthwhile to disentangle the different types of costs and explore which of them contributes the most to workers' application behaviour. The level of rationality demanded by this model is also exceptional. Even though it is possible to empirical data may exhibit patterns that resemble or trend towards the predicted equilibrium outcomes, achieving such outcomes in practice could be challenging. Drawing relation to the quantal response equilibrium, for equilibrium to hold in this model, workers need to correctly anticipate how others choose their precision levels, and how such choices translate into randomness in strategies, and they need to best respond to this anticipated randomness, Crawford et al. (2013) notes that such calibration puts stringent

requirements on workers' rationality. Furthermore, workers could have wrong perception of their attention costs, or that they pay more attention to certain firms than others, creating another layer of imprecision that may be better modelled with Gabaix (2019)'s behavioural inattention. However, despite such limitations, this model can serve as a valuable baseline to understand workers' equilibrium behaviour under inattention, and upon which behavioural elements can be further included.

**Extensions.** One of the main contributions of this model is regarding the existence of multiple equilibria, which highlights the equilibrium selection problem. While workers learning from experience as an equilibrium selection mechanism is briefly mentioned in Section 2.2.2, there is great potential to explore more on the dynamics behind which equilibrium are likely to be selected by workers and the measures social planner can adopt to nudge workers towards a socially preferred equilibrium outcome. Another question that was rarely considered is the impact of competition on attention. The GPDS model provided a foundation for the emergence of a competitive landscape, where many workers are simultaneously competing for vacancies. Higher competition among workers could lead to a decrease in wages because firms have greater power in extracting higher markdowns when the probability of them meeting a worker increase, so workers might want to execute less effort and incur less attention costs. Such hypothesis is open for investigation in the future. Last but not least, although current empirical studies effectively study the impact of attention through manipulating information presentation, they do not explicitly analyze the impact of attention costs. Experiments could be carried out as a next step to validate the findings and design mechanisms to induce better matching outcome.

## 6 Conclusion

This paper proposed a Generalized Partially Directed Search model to investigate the role of inattention on job search. It loosens some restrictions on the existing theoretical model, allowing workers to have diverse priors and heterogeneous attention costs. It accounts for the possibility of multiple equilibria that opens up discussion about efficiency of different equilibrium types.

I show that workers' equilibrium application strategies are swayed by their default search strategies, especially when information is expensive. Mismatch may be deconstructed to a consequence of innate bias, and harmonizing prior beliefs to match underlying market conditions may be important consideration for policy-makers. In presence of multiple equilibria, workers using different application strategies generates higher social welfare than them adopting the same ones. Besides, equilibrium wages may also be higher when workers employ asymmetric strategies than applying analogously. This finding highlights that monopsony power could differ across equilibrium, and there is possibility of some equilibrium generating more efficient outcome as well as lower monopsony power, thereby advocating a need to explore equilibrium selection measures. The inclusion of worker heterogeneity in the model complicates the equilibrium analysis. I demonstrate that a variety of equilibrium application strategies may arise, and most of them would be asymmetric as a result of difference in workers' attention costs. Lastly, by devising a socially optimal benchmark, I evaluate the prevailing market efficiency. Symmetric application strategies were found to contribute to inefficient outcome in general. In the case where workers are heterogeneous, there can be over- or under-estimation of efficiency in the market if heterogeneity is not adequately accounted for, and this can lead to misinformed policies. Evaluating the equilibrium outcomes with heterogeneous workers against the adjusted socially optimal benchmark shows potential for higher social welfare when workers have different attention costs. While this highlights potential benefit from accommodating or introducing worker heterogeneity, this analysis is intricate and can be highly dependent on equilibrium selection.

Overall, this paper seeks to enhance the modelling of job search when workers can be inattentive. The information-theoretic approach offers some new policy suggestions that tackle information friction in the labour market.

## References

- Abebe, G., Caria, S., Fafchamps, M., Falco, P., Franklin, S., Quinn, S., and Shilpi, F. (2020). Matching frictions and distorted beliefs: Evidence from a job fair experiment. *Department of Economics, Oxford University (mimeo)*.
- Acharya, S. and Wee, S. L. (2020). Rational inattention in hiring decisions. *American Economic Journal: Macroeconomics*, 12(1):1–40.
- Albrecht, J. and Vroman, S. (2002). A matching model with endogenous skill requirements. *International Economic Review*, 43(1):283–305.
- Altmann, S., Falk, A., Jäger, S., and Zimmermann, F. (2018). Learning about job search: A field experiment with job seekers in germany. *Journal of Public Economics*, 164:33–49.
- Archer, S. (31 May 2016). Why we don’t change jobs enough – and why we should. <https://www.theguardian.com/careers/2016/may/31/change-jobs-risks-dead-end-take-career-happiness>.
- Azeredo da Silveira, R., Sung, Y., and Woodford, M. (2024). Optimally imprecise memory and biased forecasts. *American Economic Review*, 114(10):3075–3118.
- Banner, K. M., Irvine, K. M., and Rodhouse, T. J. (2020). The use of bayesian priors in ecology: The good, the bad and the not great. *Methods in Ecology and Evolution*, 11(8):882–889.
- Belot, M., Kircher, P., and Muller, P. (2019). Providing advice to jobseekers at low cost: An experimental study on online advice. *The review of economic studies*, 86(4):1411–1447.
- Briel, S., Osikominu, A., Pfeifer, G., Reutter, M., and Satlukal, S. (2022). Gender differences in wage expectations: the role of biased beliefs. *Empirical Economics*, pages 1–26.
- Briggs, J., Caplin, A., Martin, D., Tonetti, C., and Stanford, G. (2017). Due diligence: Job search with rationally inattentive workers. Technical report, NYU working paper.
- Caplin, A. and Dean, M. (2015). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203.
- Caplin, A., Dean, M., and Leahy, J. (2019). Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies*, 86(3):1061–1094.
- Cappelli, P. H. (2015). Skill gaps, skill shortages, and skill mismatches: Evidence and arguments for the united states. *ILR review*, 68(2):251–290.
- Cheremukhin, A., Restrepo-Echavarria, P., and Tutino, A. (2020). Targeted search in matching markets. *Journal of Economic Theory*, 185:104956.
- Christie, S. T. and Schrater, P. (2015). Cognitive cost as dynamic allocation of energetic resources. *Frontiers in neuroscience*, 9:289.
- Crawford, V. P., Costa-Gomes, M. A., and Iriberri, N. (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature*, 51(1):5–62.
- Denzer, M., Schank, T., and Upward, R. (2021). Does the internet increase the job finding rate? evidence from a period of expansion in internet use. *Information Economics and Policy*, 55:100900.

- Dietrich, F. and List, C. (2013). Reasons for (prior) belief in bayesian epistemology. *Synthese*, 190(5):787–808.
- Dolado, J. J., Jansen, M., and Jimeno, J. F. (2009). On-the-job search in a matching model with heterogeneous jobs and workers. *The Economic Journal*, 119(534):200–228.
- Dubs, H. H. (1942). The principle of insufficient reason. *Philosophy of Science*, 9(2):123–131.
- Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., Hertwig, R., Stewart, T., West, R., and Lebiere, C. (2010). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making*, 23(1):15–47.
- Europa (2023). Pay transparency in the eu. <https://www.consilium.europa.eu/en/policies/pay-transparency/#:~:text=The%20EU%20wants%20to%20strengthen,EU%20rules%20on%20pay%20transparency.&text=The%20Council%20has%20adopted%20new,transparency%20on%2024%20April%202023>. Accessed: 2024-06-04.
- Faghihi, U., Estey, C., McCall, R., and Franklin, S. (2015). A cognitive model fleshes out kahneman’s fast and slow systems. *Biologically Inspired Cognitive Architectures*, 11:38–52.
- Festinger, L. (1962). *A theory of cognitive dissonance*, volume 2. Stanford university press.
- Fosgerau, M., Melo, E., De Palma, A., and Shum, M. (2020). Discrete choice and rational inattention: A general equivalence result. *International economic review*, 61(4):1569–1589.
- Gabaix, X. (2019). Behavioral inattention. In *Handbook of Behavioral Economics: Applications and Foundations 1*, volume 2, pages 261–343. Elsevier.
- Giezek, M. (2019). *Rational Inattention in Search*. PhD thesis, New York University.
- Goeree, J. K., Holt, C. A., and Palfrey, T. R. (2016). Quantal response equilibrium. In *Quantal Response Equilibrium*. Princeton University Press.
- Goeree, J. K., Holt, C. A., and Palfrey, T. R. (2020). Stochastic game theory for social science: A primer on quantal response equilibrium. In *Handbook of Experimental Game Theory*, pages 8–47. Edward Elgar Publishing.
- Gronau, R. (1971). Information and frictional unemployment. *The American Economic Review*, 61(3):290–301.
- Gürtzgen, N., Nolte, A., Pohlen, L., and van den Berg, G. J. (2018). Does the internet help unemployed job seekers find a job? evidence from the broadband internet expansion in germany.
- Habermalz, S. (2014). Rational inattention and employer learning. *Oxford Economic Papers*, 66(2):605–626.
- Handel, M. J. (2003). Skills mismatch in the labor market. *Annual Review of Sociology*, 29(1):135–165.
- Hauser, J. R. and Wernerfelt, B. (1990). An evaluation cost model of consideration sets. *Journal of consumer research*, 16(4):393–408.
- Heng, J. A., Woodford, M., and Polania, R. (2020). Efficient sampling and noisy decisions. *Elife*, 9:e54962.

- Jacob, M., Ford, J., and Deacon, T. (2023). Cognition is entangled with metabolism: relevance for resting-state eeg-fmri. *Frontiers in Human Neuroscience*, 17:976036.
- Kahneman, D. (2011). *Thinking, fast and slow*. macmillan.
- Leuven, E. and Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. *Handbook of the Economics of Education*, 4:283–326.
- Mackowiak, B., Matejka, F., Wiederholt, M., et al. (2020). Rational inattention: A review. *CEPR Discussion Papers*, (15408).
- Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98.
- Mattsson, L.-G. and Weibull, J. W. (2002). Probabilistic choice and procedurally bounded rationality. *Games and Economic Behavior*, 41(1):61–78.
- McCall, J. J. (1970). Economics of information and job search. *The Quarterly Journal of Economics*, 84(1):113–126.
- McKelvey, R. D. and Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38.
- Menzio, G. (2007). A theory of partially directed search. *Journal of political Economy*, 115(5):748–769.
- Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.
- Patterson, C., Şahin, A., Topa, G., and Violante, G. L. (2016). Working hard in the wrong place: A mismatch-based explanation to the uk productivity puzzle. *European Economic Review*, 84:42–56.
- Ravid, D. (2020). Ultimatum bargaining with rational inattention. *American Economic Review*, 110(9):2948–2963.
- Royster, D. C. (2023). Implicit function theorem.
- Sáenz-Royo, C., Chiclana, F., and Herrera-Viedma, E. (2022). Functional representation of the intentional bounded rationality of decision-makers: A laboratory to study the decisions a priori. *Mathematics*, 10(5):739.
- Şahin, A., Song, J., Topa, G., and Violante, G. L. (2014). Mismatch unemployment. *American Economic Review*, 104(11):3529–3564.
- Shi, S. (2002). A directed search model of inequality with heterogeneous skills and skill-biased technology. *The Review of Economic Studies*, 69(2):467–491.
- Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690.
- Steiner, J., Stewart, C., and Matějka, F. (2017). Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica*, 85(2):521–553.
- Wolfram (2023). Implicit function theorem.
- Woodford, M. (2020). Modeling imprecision in perception, valuation, and choice. *Annual Review of Economics*, 12(1):579–601.

Wright, R., Kircher, P., Julien, B., and Guerrieri, V. (2021). Directed search and competitive search equilibrium: A guided tour. *Journal of Economic Literature*, 59(1):90–148.

Wu, L. (2024). Partially directed search in the labor market. *Journal of Economic Theory*, page 105858.

## A Appendix: Analytical

### A.1 Signal-Based Optimization Problem

Differentiate wrt  $F^i(\mathbf{s}|\mathbf{w})$ ,  $F^{-i}(\mathbf{s}|\mathbf{w})$ ,  $q_1^i(\mathbf{s})$ ,  $q_1^{-i}(\mathbf{s})$  for (2), the FOCs are:

$$\frac{d\mathcal{L}}{dq_1^i(\mathbf{s})} : \int_{\mathbf{w}} [(1 - \frac{q_1^{-i}(\mathbf{s})}{2})w_1 - (q_1^{-i}(\mathbf{s}) + \frac{1 - q_1^{-i}(\mathbf{s})}{2})w_2] F^i(d\mathbf{s}|\mathbf{w}) G^{-i}(d\mathbf{w}) = 0 \quad (84)$$

$$\frac{d\mathcal{L}}{dq_1^{-i}(\mathbf{s})} : \int_{\mathbf{w}} [(1 - \frac{q_1^i(\mathbf{s})}{2})w_1 - (q_1^i(\mathbf{s}) + \frac{1 - q_1^i(\mathbf{s})}{2})w_2] F^{-i}(d\mathbf{s}|\mathbf{w}) G^i(d\mathbf{w}) = 0 \quad (85)$$

$$\frac{d\mathcal{L}}{F^i(\mathbf{s}|\mathbf{w})} : F^i(\mathbf{s}|\mathbf{w}) = \exp[-(1 + \frac{\int_{\mathbf{w}} [q_1^i(\mathbf{s})(1 - \frac{q_1^{-i}(\mathbf{s})}{2})w_1 + (1 - q_1^i(\mathbf{s}))(q_1^{-i}(\mathbf{s}) + \frac{1 - q_1^{-i}(\mathbf{s})}{2})w_2] G^{-i}(d\mathbf{w})}{\lambda_i})] \quad (86)$$

$$\frac{d\mathcal{L}}{F^{-i}(\mathbf{s}|\mathbf{w})} : F^{-i}(\mathbf{s}|\mathbf{w}) = \exp[-(1 + \frac{\int_{\mathbf{w}} [q_1^{-i}(\mathbf{s})(1 - \frac{q_1^i(\mathbf{s})}{2})w_1 + (1 - q_1^{-i}(\mathbf{s}))(q_1^i(\mathbf{s}) + \frac{1 - q_1^i(\mathbf{s})}{2})w_2] G^i(d\mathbf{w})}{\lambda_{-i}})] \quad (87)$$

Assume workers are identical, such that  $\lambda_i = \lambda_{-i} = \lambda$ , and in the equilibrium,  $G^i(d\mathbf{w}) = G^{-i}(d\mathbf{w}) = G^e(d\mathbf{w})$ . A symmetric application strategy could arise where  $F^i(\mathbf{s}|\mathbf{w}) = F^{-i}(\mathbf{s}|\mathbf{w}) = F^e(\mathbf{s}|\mathbf{w})$ ,  $q_1^i(\mathbf{s}) = q_1^{-i}(\mathbf{s}) = q_1^e(\mathbf{s})$ ,

$$F^e(\mathbf{s}|\mathbf{w}) = \exp[-(1 + \frac{\int_{\mathbf{w}} [q_1^e(\mathbf{s})(1 - \frac{q_1^e(\mathbf{s})}{2})w_1 + (1 - q_1^e(\mathbf{s}))(q_1^e(\mathbf{s}) + \frac{1 - q_1^e(\mathbf{s})}{2})w_2] G^e(d\mathbf{w})}{\lambda})] \quad (88)$$

$$q_1^e(\mathbf{s}) = \frac{2\mathbb{E}[w_1] - \mathbb{E}[w_2]}{\mathbb{E}[w_1] + \mathbb{E}[w_2]}, \text{ where } \mathbb{E}[w_1] = \int_{\mathbf{w}} w_1 F(d\mathbf{s}|\mathbf{w}) G(d\mathbf{w}), \mathbb{E}[w_2] = \int_{\mathbf{w}} w_2 F(d\mathbf{s}|\mathbf{w}) G(d\mathbf{w}) \quad (89)$$

Workers choose the same information structure and the same application strategy.

However, it is also possible for different equilibrium application strategies to arise. Even if they have the same consistent beliefs about the wage distribution in the equilibrium, workers can choose different information structures, such that  $F^i(\mathbf{s}|\mathbf{w}) \neq F^{-i}(\mathbf{s}|\mathbf{w})$ ,  $q_1^i(\mathbf{s}) \neq q_1^{-i}(\mathbf{s})$ . This choice of differing information structure could be due to strategic concerns and coordination attempts. It is not only about learning the true wage distribution, it is also about learning how to best respond to the other worker's strategies. Workers may be focusing on different parts of the wage information, thus choosing different information structures that remain optimal given their beliefs and costs. As highlighted in the main text, this result shares similar equilibrium properties as costly directed search with observable wages.

Return to the Section 2.1.

### A.2 Uniqueness of Symmetric Subgame Equilibrium

To show an unique solution to  $q$  exists given any wage announcement  $(w_1, w_2)$  for any  $p$ ,  $p \in (0, 1)$ . From worker  $i$ 's FOC (12),

$$\frac{p}{1-p} \exp(\frac{2w_1 - w_2 - q(w_1 + w_2)}{2\lambda}) = \frac{q}{1-q} \quad (90)$$



Case 1: For  $w_1 = 0$ ,  $w_2 = 0$ ,  $q = p$ . Unique solution exist for  $q$ .

Case 2: For  $w_1 > 0$ ,  $w_2 = 0$ ,

$$\frac{p}{1-p} \exp\left(\frac{(2-q)w_1}{2\lambda}\right) = \frac{q}{1-q} \quad (91)$$

Express  $p$  in terms of  $q$ ,

$$p = \frac{qA^{q-2}}{1-q+qA^{q-2}}, \text{ where } \exp\left(\frac{w_1}{2\lambda}\right) = A, A > 0 \quad (92)$$

Let  $f(q) = \frac{qA^{q-2}}{1-q+qA^{q-2}} - p$ , knowing  $q \in (0, 1)$ ,

$$\lim_{q \rightarrow 0} f(q) = -p < 0, \lim_{q \rightarrow 1} f(q) = 1 - p > 0$$

$f'(q) > 0$  and  $f(q)$  is a strictly increasing function in the interval, therefore by intermediate value theorem, for any  $p \in (0, 1)$ , a unique solution  $q$  exists.

Case 3: For  $w_1 = 0$ ,  $w_2 > 0$ ,

$$\frac{p}{1-p} \exp\left(\frac{-(1+q)w_2}{2\lambda}\right) = \frac{q}{1-q} \quad (93)$$

Express  $p$  in terms of  $q$ ,

$$p = \frac{qB^{1+q}}{1-q+qB^{1+q}}, \text{ where } \exp\left(\frac{w_2}{2\lambda}\right) = B, B > 0 \quad (94)$$

Let  $g(q) = \frac{qB^{1+q}}{1-q+qB^{1+q}} - p$ , knowing  $q \in (0, 1)$ ,

$$\lim_{q \rightarrow 0} g(q) = -p < 0, \lim_{q \rightarrow 1} g(q) = 1 - p > 0$$

$g'(q) > 0$  and  $g(q)$  is a strictly increasing function in the interval, therefore by intermediate value theorem, for any  $p \in (0, 1)$ , a unique solution  $q$  exists.

Case 4: For  $w_1 > 0$ ,  $w_2 > 0$ , use (90) and express  $p$  in terms of  $q$ ,

$$p = \frac{qC^{-1}D^{-q}}{1-q+qC^{-1}D^{-q}}, \text{ where } \exp\left(\frac{2w_1-w_2}{2\lambda}\right) = C, \exp\left(\frac{-(w_1+w_2)}{2\lambda}\right) = D, C, D > 0 \quad (95)$$

Let  $h(q) = \frac{qC^{-1}D^{-q}}{1-q+qC^{-1}D^{-q}} - p$ ,

$$\lim_{q \rightarrow 0} h(q) = -p < 0 \quad (96)$$

$$\lim_{q \rightarrow 1} h(q) = 1 - p > 0 \quad (97)$$

$h'(q) > 0$  and  $h(q)$  is a strictly increasing function in the interval, therefore by intermediate value theorem, for any  $p \in (0, 1)$ , a unique solution  $q$  exist. Therefore, it has been shown that unique solution to  $q$  exist for all cases of wage announcement for any  $p$ ,  $p \in (0, 1)$ .

Return to the Section 2.1.

### A.3 Existence of Asymmetric Subgame Equilibrium

To show that Asymmetric Subgame Equilibrium (ASE) exists under certain conditions.

*Step 1:* Based on (37) and (38), given  $q_1^*$  and  $q_1^{-i*}$  are continuous functions, define  $h(q_1^*, q_1^{-i*}, w_1, w_2)$  as

$$h(\cdot) = \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2-q_1^{-i*})w_1 - (1+q_1^{-i*})w_2}{2\lambda}))} - q_1^* - \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2-q_1^*)w_1 - (1+q_1^*)w_2}{2\lambda}))} + q_1^{-i*} \quad (98)$$

where  $h(\cdot)$  is a continuous function, dependent on  $q_1^*, q_1^{-i*}, w_1, w_2$ . My aim is to show that  $h(\cdot)$  may take positive and negative values at different points, and that it crosses 0 multiple times. By intermediate value theorem, this would prove for multiple combinations of  $(q_1^*, q_1^{-i*})$  that solves for  $h(\cdot) = 0$ .

Evaluate  $h(\cdot)$  at the extreme values of  $q_1^*$  and  $q_1^{-i*}$  and for the intermediary cases:

when  $q_1^* \rightarrow 0$  and  $q_1^{-i*} \rightarrow 1$ ,

$$h(0, 1, w_1, w_2) = \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{w_1 - 2w_2}{2\lambda}))} - \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{2w_1 - w_2}{2\lambda}))} + 1 > 0 \quad (99)$$

and  $q_1^* \rightarrow 1$  and  $q_1^{-i*} \rightarrow 0$ ,

$$h(1, 0, w_1, w_2) = \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{2w_1 - w_2}{2\lambda}))} - \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{w_1 - 2w_2}{2\lambda}))} - 1 < 0 \quad (100)$$

The above holds true for all wage announcements. However, in order to determine the wage condition necessary for fulfilling the equilibrium conditions, the original equations for  $(q_1^*, q_1^{-i*})$  have to be satisfied as well, which indicates that  $\lambda$  has to be small relative to the expected payoff, and  $2w_1 > w_2 > \frac{w_2}{2}$ . To further illustrate this condition, I use workers' FOCs (27) and (28).

For  $q_1^{-i*} \rightarrow 0$  and  $q_1^* \rightarrow 1$ , (27) shows

$$\frac{1}{\lambda} [(1 - q_1^{-i*} + \frac{q_1^{-i*}}{2})w_1 - (q_1^{-i*} + \frac{1 - q_1^{-i*}}{2})w_2] = \log(\frac{1-p}{p} \frac{q_1^*}{1 - q_1^*}) \quad (101)$$

$$\text{LHS: } \frac{1}{\lambda} \lim_{q_1^{-i*} \rightarrow 0, q_1^* \rightarrow 1} [(1 - q_1^{-i*} + \frac{q_1^{-i*}}{2})w_1 - (q_1^{-i*} + \frac{1 - q_1^{-i*}}{2})w_2] = \frac{w_1 - \frac{1}{2}w_2}{\lambda}$$

$$\text{RHS: } \lim_{q_1^{-i*} \rightarrow 0, q_1^* \rightarrow 1} \log(\frac{1-p}{p} \frac{q_1^*}{1 - q_1^*}) = \infty$$

The equality holds iff  $\lambda \rightarrow 0^+$ , and/or  $w_1 - \frac{1}{2}w_2$  is sufficiently large,

$$\lim_{\lambda \rightarrow 0^+} \frac{w_1 - \frac{1}{2}w_2}{\lambda} = \infty, \text{ s.t. } w_1 > \frac{1}{2}w_2$$

And for (28),

$$\frac{1}{\lambda} [(1 - q_1^* + \frac{q_1^*}{2})w_1 - (q_1^* + \frac{1 - q_1^*}{2})w_2] = \log(\frac{1-p}{p} \frac{q_1^{-i*}}{1 - q_1^{-i*}}) \quad (102)$$

$$\text{LHS: } \frac{1}{\lambda} \lim_{q_1^{-i^*} \rightarrow 0, q_1^* \rightarrow 1} [(1 - q_1^* + \frac{q_1^*}{2})w_1 - (q_1^* + \frac{1 - q_1^*}{2})w_2] = \frac{\frac{1}{2}w_1 - w_2}{\lambda}$$

$$\text{RHS: } \lim_{q_1^{-i^*} \rightarrow 0, q_1^* \rightarrow 1} \log(\frac{1-p}{p} \frac{q_1^{-i^*}}{1 - q_1^{-i^*}}) = -\infty$$

The equality holds iff  $\lambda \rightarrow 0^-$ , and/or  $\frac{1}{2}w_1 - w_2$  is sufficiently large, but since  $\lambda$  cannot be negative in this model,  $\frac{1}{2}w_1 < w_2$  has to be fulfilled.

$$\lim_{\lambda \rightarrow 0^+} \frac{\frac{1}{2}w_1 - w_2}{\lambda} = -\infty, \text{ s.t. } \frac{1}{2}w_1 < w_2$$

In addition to the above, if firms are homogeneous and set the same wage  $w_1 = w_2 = w$ , then the limit equations for (27) and (28) would hold iff  $\lambda \rightarrow 0^+$  with no other wage conditions,

$$\lim_{\lambda \rightarrow 0^+} \frac{w}{2\lambda} = \infty, \lim_{\lambda \rightarrow 0^+} -\frac{w}{2\lambda} = -\infty$$

With that, I specified the wage condition, which is small cognitive costs ( $\lambda$ ) relative to expected payoff. The expected payoff depends on wage ( $w$ ) when workers are facing homogeneous firms that offer the same wage; or wage differentials ( $|w_1 - w_2|$ ) when firms are heterogeneous in productivity and could set different wages, and in this case,  $2w_1 > w_2 > \frac{w_1}{2}$  needs to hold. Wages can approximately go as high as two times apart between the firms.

Suppose  $w_1$  is set way higher than  $w_2$ , such that (101) could be met in the limit but not (102), then equilibrium conditions do not hold, there will not be solution for  $q_1^{-i^*} \rightarrow 0$  and  $q_1^* \rightarrow 1$ . Intuitively, this implies when facing high  $w_1$  relative to  $w_2$ , worker 1 would apply with higher probability to firm 1 than to firm 2, and since the wage differential is sufficiently large, it could be more profitable for worker 2 to apply with higher probability to firm 1 as well, therefore, the case where worker 2 use a different strategy from worker 1 and stays “loyal” to firm 2 cease to exist. The same analysis applies to  $q_1^{-i^*} \rightarrow 1$  and  $q_1^* \rightarrow 0$ .

Since (99) and (100) always hold, for multiple solutions to exist, I need to evaluate  $h(\cdot)$  at intermediary values of  $q_1^*$  and  $q_1^{-i^*}$ . Given there always exist a unique symmetric solution, such that  $h(\cdot) = 0$  for some  $q_1^* = q_1^{-i^*} = q^*$ , I can evaluate the variables at a value close to  $q^*$ . Define  $q^{*-} = q^* - \epsilon$  and  $q^{*+} = q^* + \epsilon$ , for  $q_1^* \rightarrow q^{*-}$  and  $q_1^{-i^*} \rightarrow q^{*+}$ ,

$$h(q^{*-}, q^{*+}, w_1, w_2) = \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2w_1 - w_2 - q^{*+})(w_1 + w_2)}{2\lambda}))} - \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2w_1 - w_2 - q^{*-})(w_1 + w_2)}{2\lambda}))} - (q^{*-} - q^{*+}) < 0 \quad (103)$$

where  $(q^{*-} - q^{*+}) = -2\epsilon$ . And for  $q_1^* \rightarrow q^{*+}$  and  $q_1^{-i^*} \rightarrow q^{*-}$ ,

$$h(q^{*+}, q^{*-}, w_1, w_2) = \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2w_1 - w_2 - q^{*-})(w_1 + w_2)}{2\lambda}))} - \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2w_1 - w_2 - q^{*+})(w_1 + w_2)}{2\lambda}))} - (q^{*+} - q^{*-}) > 0 \quad (104)$$

where  $(q^{*+} - q^{*-}) = 2\epsilon$ .

The sign of (103) and (104) would depend on the values of  $w_1$ ,  $w_2$  and  $\lambda$ . If  $\lambda$  is large,

$h(q^{*-}, q^{*+}, w_1, w_2) > 0$ ,  $h(q^{*+}, q^{*-}, w_1, w_2) < 0$ , implying that  $h(\cdot)$  only cuts 0 once, and resulting in only one set of solution for  $(q_1^*, q_1^{-i*})$ . For small  $\lambda$  relative to the expected payoff, then  $h(q^{*-}, q^{*+}, w_1, w_2) < 0$ ,  $h(q^{*+}, q^{*-}, w_1, w_2) > 0$  hold, and  $h(\cdot)$  cuts 0 multiple times. The same analysis applies when firms are homogeneous, except I would be evaluating  $\lambda$  relative to  $w$ .  $h(\cdot)$  could potentially cut 0 three times, resulting in three possible equilibria in the subgame (2 ASE and 1 SSE) for a given set of wages.

*Step 2:* The existence of ASE relies heavily on the upholding of (103) and (104). If the inequalities break down, they combine to form the following inequality, which would result in one unique set of solution for  $(q_1^*, q_1^{-i*})$ ,

$$\frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{2w_1 - w_2 - q^{*-}(w_1 + w_2)}{2\lambda}))} - \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{2w_1 - w_2 - q^{*+}(w_1 + w_2)}{2\lambda}))} < 2\epsilon \quad (105)$$

$$\frac{1}{1 + \frac{1-p}{p} ab} - \frac{1}{1 + \frac{1-p}{p} \frac{a}{b}} < 2\epsilon, a = \exp(-(\frac{2w_1 - w_2 - q^*(w_1 + w_2)}{2\lambda})), b = \exp(-\frac{\epsilon(w_1 + w_2)}{2\lambda}) \quad (106)$$

For any variation in  $\lambda$ ,  $w_1$  and  $w_2$ ,  $a$  moves the first and second term of (106) in the same direction, making the extent of change in  $b$  essential. For large  $\lambda$  and small  $w_1$  and  $w_2$ , it is more likely to fulfill (106), and therefore, a unique solution will be anticipated instead of multiple. The main purpose of (106) is for us to evaluate the potential threshold of  $\lambda$ , such that if  $\lambda$  exceeds certain value,  $\bar{\lambda}$ , then ASE would not exist.

This threshold is an implicit expression of  $w_1$ ,  $w_2$ ,  $q^*$ ,  $p$ .

*Comparative statistics* of the changes in the parameters on  $\bar{\lambda}$ , where higher  $\bar{\lambda}$  indicates workers could be applying asymmetrically for a wider range of information costs:

- $w_1$  and  $w_2$ : While increasing wage announcements always increase  $\epsilon(w_1 + w_2)$ , resulting in higher  $\bar{\lambda}$  in general, the overall impact would depend on  $a$ . If both firms set higher wages, but wage differential is also high, then the value for workers to learn about the precise wage distribution would be higher, it becomes justifiable for workers who previously possess higher information costs to begin acquiring and processing information in view of the additional benefit.
- $q^*$ : The value of  $q^*$  is determined in the symmetric equilibrium. Suppose  $q^*$  increases and worker 1 applies with higher probability to firm 1,  $\bar{\lambda}$  will be lowered. An increase in  $q^*$  could in turn be attributed to an increase in the default probability of selecting firm 1,  $p$ , or increase in  $w_1^*$  relative to  $w_2^*$  in the SSPE – the first suggest an increase in workers' initial bias towards firm 1, thus the threshold for cognitive costs might have to be lower for workers to be directed to different firms in the equilibrium; the second denotes an increase in the equilibrium wage differential, this intuitively makes coordinating on applying to different firms difficult as the potential benefit of applying symmetrically to the higher wage firm could be larger than applying to different firms, therefore the threshold of  $\lambda$  required for asymmetric strategies to emerge would be lowered.

The above shows that  $\bar{\lambda}$  could be higher while ensuring asymmetric strategies exist in the equilibrium if expected payoff remain high and the existence condition for ASE is still satisfied; or when  $q^*$  determined in SSPE is more uniform, which also implies that workers are less biased initially.

*Step 3:* In the full game, wages are endogenously determined by firms. Therefore, given workers'

behaviours in the subgame, there should be a corresponding mapping from wage announcements to the equilibrium wages,  $(w_1^*, w_2^*)$ , where  $w_1^* \in [0, z_1]$ ,  $w_2^* \in [0, z_2]$ , and this pair have to satisfy the wage conditions for ASE previously analyzed.

Drawing relation between wages and productivities to shine a light on condition for ASPE in general. If firms are homogeneous and set the same wages, then as specified in *Step 2*,  $\lambda$  would have to be sufficiently small relative to expected payoff, which can directly mean that  $\lambda$  has to be sufficiently small relative to productivities or productivity difference. For firms offering different wages, then equilibrium wages need to satisfy  $2w_1^* > w_2^* > \frac{w_1^*}{2}$  on top of small  $\lambda$  relative to expected payoffs. Based on the wage determination equations (39) and (40), the wage condition can be reformulated as

$$2 \max[z_1 - m_1(q_1^*, q_1^{-i*}, \lambda), 0] > \max[z_2 - m_2(q_1^*, q_1^{-i*}, \lambda), 0] > \frac{\max[z_1 - m_1(q_1^*, q_1^{-i*}, \lambda), 0]}{2} \quad (107)$$

where the markdowns are denoted as  $m_1$  and  $m_2$  for firm 1 and 2 respectively, and depends workers' cognitive costs and strategies. This reduces to,

$$2z_1 - z_2 > 2m_1 - m_2, 2z_2 - z_1 > 2m_2 - m_1$$

The condition depends on productivity difference between the firms and the markdown difference, which are in turn subjected to  $q_1^*$ ,  $q_1^{-i*}$ ,  $\lambda$ . Generally speaking, the extent at which the productivity could differ is approximately twice the difference in either direction.

Return to the Section 2.3.1.

## A.4 Variations in Defaults Given Existence of Multiple Equilibria

Demonstrating the changes in application probability to firm 1 given a change in defaults,

$$\begin{aligned} \frac{dq_1^*}{dp} = & \frac{q_1^*(1 - q_1^*)[(2 - q_1^*)q_1^{-i*}(1 - q_1^{-i*})(w_1 + w_2) - 2\lambda(2 - q_1^{-i*})]}{(w_1 + w_2)^2 q_1^{-i*}(1 - q_1^{-i*})q_1^*(1 - q_1^*) - 4\lambda^2} \frac{dw_1^*}{dp} \\ & - \frac{q_1^*(1 - q_1^*)[(1 + q_1^*)q_1^{-i*}(1 - q_1^{-i*})(w_1^* + w_2^*) - 2\lambda(1 + q_1^{-i*})]}{(w_1^* + w_2^*)^2 q_1^{-i*}(1 - q_1^{-i*})q_1^*(1 - q_1^*) - 4\lambda^2} \frac{dw_2^*}{dp} \\ & + \frac{[2\lambda(w_1^* + w_2^*)q_1^{-i*}(1 - q_1^{-i*}) - 4\lambda^2]q_1^*(1 - q_1^*)}{p(1 - p)[(w_1^* + w_2^*)^2 q_1^{-i*}(1 - q_1^{-i*})q_1^*(1 - q_1^*) - 4\lambda^2]} \end{aligned} \quad (108)$$

$$\begin{aligned} \frac{dq_1^{-i*}}{dp} = & \frac{q_1^{-i*}(1 - q_1^{-i*})[(2 - q_1^{-i*})q_1^*(1 - q_1^*)(w_1^* + w_2^*) - 2\lambda(2 - q_1^*)]}{(w_1^* + w_2^*)^2 q_1^*(1 - q_1^*)q_1^{-i*}(1 - q_1^{-i*}) - 4\lambda^2} \frac{dw_1^*}{dp} \\ & - \frac{q_1^{-i*}(1 - q_1^{-i*})[(1 + q_1^{-i*})q_1^*(1 - q_1^*)(w_1^* + w_2^*) - 2\lambda(1 + q_1^*)]}{(w_1^* + w_2^*)^2 q_1^*(1 - q_1^*)q_1^{-i*}(1 - q_1^{-i*}) - 4\lambda^2} \frac{dw_2^*}{dp} \\ & + \frac{q_1^{-i*}(1 - q_1^{-i*})[2\lambda(w_1^* + w_2^*)q_1^*(1 - q_1^*) - 4\lambda^2]}{(1 - p)p[(w_1^* + w_2^*)^2 q_1^*(1 - q_1^*)q_1^{-i*}(1 - q_1^{-i*}) - 4\lambda^2]} \end{aligned} \quad (109)$$

$p$  only has indirect impact on  $w_1^*$  and  $w_2^*$  through  $q_1^*$  and  $q_1^{-i*}$ . The signs of  $\frac{dq_1^*}{dp}$  and  $\frac{dq_1^{-i*}}{dp}$  are likely to be positive.<sup>8</sup> Any deviation away from  $p = \frac{1}{2}$  has a direct positive effect on  $\frac{dq_1^*}{dp}$  and  $\frac{dq_1^{-i*}}{dp}$ , suggesting that workers having an initial bias towards firm 1 would apply with higher probability to firm 1 in the equilibrium.

Return to the Section 2.3.1.

---

<sup>8</sup>The first two terms of each equation have a smaller impact than the third term on the overall sign given their denominators are substantially larger, and that the third term is positive.

## A.5 Higher Social Welfare in ASPE than SSPE

To show that  $\mathbb{W}(q_1^*, q_1^{-i*}) > \mathbb{W}(q^*)$ :

*Step 1:* From (50), I first demonstrate that the expected productivity/ex-post welfare is higher under ASPE than SSPE:

$$((1 - q^*)^2 - (1 - q_1^*)(1 - q_1^{-i*}))z_1 + (q^{*2} - q_1^*q_1^{-i*})z_2 > 0 \quad (110)$$

Suppose workers adopt more extreme asymmetric strategies,  $q_1^* \rightarrow 1$ ,  $q_1^{-i*} \rightarrow 0$  or  $q_1^* \rightarrow 0$ ,  $q_1^{-i*} \rightarrow 1$ .

$$\lim_{q_1^* \rightarrow 1, q_1^{-i*} \rightarrow 0} ((1 - q^*)^2 - (1 - q_1^*)(1 - q_1^{-i*}))z_1 + (q^{*2} - q_1^*q_1^{-i*})z_2 = (1 - q^*)^2z_1 + q^{*2}z_2 > 0$$

$$\lim_{q_1^* \rightarrow 0, q_1^{-i*} \rightarrow 1} ((1 - q^*)^2 - (1 - q_1^*)(1 - q_1^{-i*}))z_1 + (q^{*2} - q_1^*q_1^{-i*})z_2 = (1 - q^*)^2z_1 + q^{*2}z_2 > 0$$

(110) is satisfied.

Then, suppose workers adopt less extreme asymmetric strategies,  $q_1^* \rightarrow q^{*-}$  and  $q_1^{-i*} \rightarrow q^{*+}$ , or  $q_1^* \rightarrow q^{*+}$  and  $q_1^{-i*} \rightarrow q^{*-}$ , where  $q^{*-} = q^* - \epsilon$ ,  $q^{*+} = q^* + \epsilon$ ,  $\epsilon$  is negligible.

$$((1 - q^*)^2 - (1 - q_1^*)(1 - q_1^{-i*}))z_1 + (q^{*2} - q_1^*q_1^{-i*})z_2 = \epsilon^2z_1 + \epsilon^2z_2 > 0$$

In the limit, the expected productivity difference between ASPE and SSPE tends to 0, but for any positive  $\epsilon$ , there is higher expected productivity/ex-post welfare in ASPE than SSPE. The difference can be further exaggerated if firms' productivity are high.

*Step 2:* I study the relative attention costs between the two equilibrium strategies. If  $\lambda$  is negligible, then (50) always hold. For  $\lambda > 0$ , the magnitude of the RHS matters.

*Relative cost of different equilibrium strategies.* Obstructing away  $\lambda$ , define RHS of (50) to be a function  $\mathbb{S}(q_j^*, q_j^{-i*}, q^*)$ ,

$$\mathbb{S}(\cdot) = \sum_{j=1}^2 q_j^* \log \frac{q_j^*}{p_j} + \sum_{j=1}^2 q_j^{-i*} \log \frac{q_j^{-i*}}{p_j} - 2q^* \log \left( \frac{q^*}{p} \right) - 2(1 - q^*) \log \left( \frac{1 - q^*}{1 - p} \right) \quad (111)$$

Define  $q_1^* \rightarrow 1$ ,  $q_1^{-i*} \rightarrow 0$  or  $q_1^* \rightarrow 0$ ,  $q_1^{-i*} \rightarrow 1$  as  $\lim_{pure}$ . From (111),

$$\lim_{pure} \mathbb{S}(\cdot) = (1 - 2q^*) \log \left( \frac{1 - p}{p} \right) - 2q^* \log \left( \frac{q^*}{1 - q^*} \right) - 2 \log(1 - q^*) \quad (112)$$

If workers have uniform defaults (i.e.  $p = \frac{1}{2}$ ),

$$\lim_{pure} \mathbb{S}(\cdot) = -2q^* \log \left( \frac{q^*}{1 - q^*} \right) - 2 \log(1 - q^*) > 0, \lim_{pure} \mathbb{S}(\cdot) \in [0, \log(4)]$$

If workers are biased to firm 1 ( $p > \frac{1}{2}$ ) or firm 2 ( $p < \frac{1}{2}$ ),

$\lim_{pure} \mathbb{S}(\cdot) \in [-C, (1 - 2q^*) \log \left( \frac{1 - p}{p} \right) - 2q^* \log \left( \frac{q^*}{1 - q^*} \right) - 2 \log(1 - q^*)]$ , where  $C$  is a positive constant.

$\lim_{pure} \mathbb{S}(\cdot)$  is negative for small values of  $q^*$  when  $p > \frac{1}{2}$ , and  $\lim_{pure} \mathbb{S}(\cdot)$  is negative for large values of  $q^*$  for  $p < \frac{1}{2}$ . By (23), higher  $p$  will not induce a low  $q^*$ , thus if workers are biased,  $\lim_{pure} \mathbb{S}(\cdot)$  is expected to be positive. Further, the maximum point is always higher than  $\log(4)$

for  $p \in (0, 1)$ ,  $p \neq \frac{1}{2}$ , thus if workers are biased,  $\mathbb{S}(\cdot)$  is expected to be larger, and thus higher likelihood of inequality (50) to break down.

Next, define  $q_1^* \rightarrow q^{*-}$  and  $q_1^{-i*} \rightarrow q^{*+}$ , or  $q_1^* \rightarrow q^{*+}$  and  $q_1^{-i*} \rightarrow q^{*-}$  as  $\lim_{mixed}$ . From (111),

$$\lim_{mixed} q^* \log\left(\frac{(q^* - \epsilon)(q^* + \epsilon)}{(1 - (q^* - \epsilon))(1 - (q^* + \epsilon))}\right) - \epsilon \log\left(\frac{(q^* - \epsilon)(1 - (q^* + \epsilon))}{(1 - (q^* - \epsilon))(q^* + \epsilon)}\right) \\ - 2q^* \log\left(\frac{q^*}{1 - q^*}\right) + \log\left(\frac{(1 - q^*)^2 - \epsilon^2}{(1 - q^*)^2}\right)$$

For small  $\epsilon$ , the expression is positive but remains low, with a minimum at  $q^* \rightarrow \frac{1}{2}$ . Workers using asymmetric strategies that are very close to symmetric ones likely to incur small cost relative to the ex-post welfare, therefore the inequality (50) is likely to hold.

*Cognitive costs.* To investigate the magnitude of  $\lambda$  for the inequality (50) to remain true. Reformulating (50):

$$\lambda < \frac{((1 - q^*)^2 - (1 - q_1^*)(1 - q_1^{-i*}))z_1 + (q^{*2} - q_1^* q_1^{-i*})z_2}{\sum_{j=1}^2 q_j^* \log \frac{q_j^*}{p_j} + \sum_{j=1}^2 q_j^{-i*} \log \frac{q_j^{-i*}}{p_j} - 2q^* \log\left(\frac{q^*}{p}\right) - 2(1 - q^*) \log\left(\frac{1 - q^*}{1 - p}\right)} \quad (113)$$

Assuming ASPE exist, based on workers' FOCs (37) and (38), for  $q_1^* \rightarrow 0$ ,  $q_1^{-i*} \rightarrow 1$  and  $q_1^* \rightarrow 1$ ,  $q_1^{-i*} \rightarrow 0$ ,  $\lambda \rightarrow 0$ ; for  $q_1^* \rightarrow q^{*-}$ ,  $q_1^{-i*} \rightarrow q^{*+}$  and  $q_1^* \rightarrow q^{*+}$ ,  $q_1^{-i*} \rightarrow q^{*-}$ ,  $\lambda \rightarrow \frac{(1 - \frac{q^*}{2})w_1^* - (\frac{q^*+1}{2})w_2^*}{\log(\frac{1-p}{p} \frac{q^*}{1-q^*})}$ .  $\lambda$  will have to fall within the range:

$$0 < \lambda < \frac{(1 - \frac{q^*}{2})w_1^* - (\frac{q^*+1}{2})w_2^*}{\log(\frac{1-p}{p} \frac{q^*}{1-q^*})} \quad (114)$$

where  $z_1 \geq w_1^*$  and  $z_2 \geq w_2^*$ ,  $w_1^*$  and  $w_2^*$  are determined by (39) and (40),  $q^*$  is determined by the system of equations (19), (17) and (18).

The condition on  $\lambda$  for existence of ASPE is stricter than  $\lambda$  required for fulfilling the (113), thus welfare is in general higher given ASPE exist. For instance, if workers behave close to pure coordination,  $\lambda$  required for ASPE to exist needs to be close to 0, but remain positive for (113). As we move from that to more symmetric strategies,  $\lambda$  for existence of ASPE is slightly relaxed, it would be necessary to verify

$$\frac{(1 - \frac{q^*}{2})w_1^* - (\frac{q^*+1}{2})w_2^*}{\log(\frac{1-p}{p} \frac{q^*}{1-q^*})} < \frac{(1 - q^*)^2 z_1 + q^{*2} z_2}{\log(\frac{1}{1-p}) + \log(\frac{1}{p}) - 2q^* \log(\frac{q^*}{p}) - 2(1 - q^*) \log(\frac{1 - q^*}{1 - p})} \quad (115)$$

If this holds, then welfare is higher in ASPE than SSPE as long as ASPE exist.

Suppose firms are homogeneous,  $z_1 = z_2 = z$ , from (115),

$$\frac{(\frac{1}{2} - q^*)w^*}{\log(\frac{1-p}{p} \frac{q^*}{1-q^*})} < \frac{((1 - q^*)^2 + q^{*2})z}{\log(\frac{1}{1-p}) + \log(\frac{1}{p}) - 2q^* \log(\frac{q^*}{p}) - 2(1 - q^*) \log(\frac{1 - q^*}{1 - p})} \quad (116)$$

For unbiased workers (i.e.  $p = \frac{1}{2}$ ),  $q^* = \frac{1}{2}$  when facing homogeneous firms, LHS of (116) tends to  $-0.25w^*$  (as per L'Hopital's Rule), and RHS tends to  $0.361z$ , the inequality always hold. For biased workers, equilibrium wages will be affected. For workers biased to firm 1 (i.e.  $p > \frac{1}{2}$ ),  $w_2^*$

increases relative to  $w_1^*$  in the equilibrium, LHS of (115) is lowered, RHS is always positive, and is higher the larger the  $q^*$ . The inequality will be satisfied. To illustrate an example,

$$\frac{(1 - \frac{q^*}{2})w_1^* - (\frac{q^*+1}{2})w_2^*}{\log(\frac{1-p}{p} \frac{q^*}{1-q^*})} < \frac{((1 - q^*)^2 + q^{*2})z}{\log(\frac{1}{1-p}) + \log(\frac{1}{p}) - 2q^* \log(\frac{q^*}{p}) - 2(1 - q^*) \log(\frac{1-q^*}{1-p})}$$

for  $z = 10$ ,  $p = 0.6$ ,  $w_1^* = 2.765$ ,  $w_2^* = 3.23$ ,  $q^* = 0.508$ ,  $LHS - RHS = 2.593 > 0$ .

The same analysis apply for workers biased in the other direction.

Suppose firms are heterogeneous, let  $z_1 = xz_2$ ,  $x \geq 1$ , firm 1 is  $x$  times more productive than firm 2, and  $w_1^* = w_2^* + \delta x$ ,  $\delta \in [0, 1]$ , indicating that wages differ as productivity difference grows, but wage differential increases only as a proportion of the increase in productivity. This generic simplification allows for easy comparison between the sides of the inequality:

$$\frac{(\frac{1}{2} - q^*)w_2^* + (1 - \frac{q^*}{2})\delta x}{\log(\frac{1-p}{p} \frac{q^*}{1-q^*})} < \frac{((1 - q^*)^2 + q^{*2})z_2 + (1 - q^*)^2 x}{\log(\frac{1}{1-p}) + \log(\frac{1}{p}) - 2q^* \log(\frac{q^*}{p}) - 2(1 - q^*) \log(\frac{1-q^*}{1-p})} \quad (117)$$

As  $x$  becomes substantially large, unbiased workers would apply with higher probability to firm 1,  $q^*$  approaches 1, LHS of (117) would be

$$\lim_{x \rightarrow \infty, q^* \rightarrow 1} \frac{(\frac{1}{2} - q^*)w_2^* + (1 - \frac{q^*}{2})\delta x}{\log(\frac{1-p}{p} \frac{q^*}{1-q^*})} = 0 \quad (118)$$

Denominator goes to infinity at a faster rate than numerator. As for the RHS,

$$\lim_{x \rightarrow \infty, q^* \rightarrow 1} \frac{((1 - q^*)^2 + q^{*2})z_2 + (1 - q^*)^2 x}{\log(\frac{1}{1-p}) + \log(\frac{1}{p}) - 2q^* \log(\frac{q^*}{p}) - 2(1 - q^*) \log(\frac{1-q^*}{1-p})} = \frac{z_2}{\log(\frac{p}{1-p})} \quad (119)$$

This shows that as  $x$  becomes larger and  $q^*$  approaches 1, the inequality always hold. Further, if adjustment in firm 1 wage is slow such that  $\delta$  is small, then (117) is always satisfied. In Figure 4a, for a given increase in  $z_1$ , the speed of change in firm 1 wage is slower for ASPE as compared to SSPE, which indicates higher  $x$  is associated with lower  $\delta$ , therefore, as productivity difference grow faster than wage differentials, social welfare will be higher when workers are applying to different firms than overcrowding at the same one.

Return to the Section 2.3.3.

## A.6 Existence of ASPE with Heterogeneous Default Strategies

To show that ASE exists under certain conditions when workers have heterogeneous defaults.

*Step 1:* I first prove there exists at least one ASE solution.

Based on workers' FOCs:

$$q_1 : (1 - q_1^{-i} + \frac{q_1^{-i}}{2})w_1 - (q_1^{-i} + \frac{1 - q_1^{-i}}{2})w_2 = \lambda \log(\frac{1 - p_1}{p_1} \frac{q_1}{1 - q_1}) \quad (120)$$

$$q_1^{-i} : (1 - q_1 + \frac{q_1}{2})w_1 - (q_1 + \frac{1 - q_1}{2})w_2 = \lambda \log(\frac{1 - p_1^{-i}}{p_1^{-i}} \frac{q_1^{-i}}{1 - q_1^{-i}}) \quad (121)$$

For a given pair of wages,  $(w_1, w_2)$ , solve for  $(q_1^*, q_1^{-i*})$  in the subgame:

$$\frac{p_1}{1 - p_1} \exp(\frac{(2 - q_1^{-i*})w_1 - (1 + q_1^{-i*})w_2}{2\lambda}) - \frac{q_1^*}{1 - q_1^*} = 0 \quad (122)$$



$$\frac{p_1^{-i}}{1 - p_1^{-i}} \exp\left(\frac{(2 - q_1^*)w_1 - (1 + q_1^*)w_2}{2\lambda}\right) - \frac{q_1^{-i*}}{1 - q_1^{-i}} = 0 \quad (123)$$

Set LHS of the equations to be  $f(\cdot)$  and  $g(\cdot)$  respectively.

$$f(\cdot) = \frac{p_1}{1 - p_1} \exp\left(\frac{(2 - q_1^{-i*})w_1 - (1 + q_1^{-i*})w_2}{2\lambda}\right) - \frac{q_1^*}{1 - q_1^*} \quad (124)$$

$$g(\cdot) = \frac{p_1^{-i}}{1 - p_1^{-i}} \exp\left(\frac{(2 - q_1^*)w_1 - (1 + q_1^*)w_2}{2\lambda}\right) - \frac{q_1^{-i*}}{1 - q_1^{-i}} \quad (125)$$

Given  $q_1^*$  and  $q_1^{-i*}$  are continuous and  $q_1^* \in (0, 1)$ ,  $q_1^{-i*} \in (0, 1)$ ,  $h(q_1^*, q_1^{-i*}) = (f(q_1^*, q_1^{-i*}), g(q_1^*, q_1^{-i*})) = f(\cdot) - g(\cdot)$  can map any combination of points from the  $q_1^*-q_1^{-i*}$  plane to the  $f$ - $g$  plane.

Knowing  $q_1^*$  and  $q_1^{-i*}$  go in opposing direction, to show at least one solution exist,

$$\lim_{q_1^* \rightarrow 0, q_1^{-i*} \rightarrow 1} h(q_1^*, q_1^{-i*}) = \frac{p_1}{1 - p_1} \exp\left(\frac{w_1 - 2w_2}{2\lambda}\right) - \frac{p_1^{-i}}{1 - p_1^{-i}} \exp\left(\frac{2w_1 - w_2}{2\lambda}\right) + \infty > 0$$

$$\lim_{q_1^* \rightarrow 1, q_1^{-i*} \rightarrow 0} h(q_1^*, q_1^{-i*}) = \frac{p_1}{1 - p_1} \exp\left(\frac{2w_1 - w_2}{2\lambda}\right) - \infty - \frac{p_1^{-i}}{1 - p_1^{-i}} \exp\left(\frac{w_1 - 2w_2}{2\lambda}\right) < 0$$

Since  $f(q_1^*, q_1^{-i*}) - g(q_1^*, q_1^{-i*})$  shows opposing signs from one extreme set of values to another, by intermediate value theorem, there would exist at least one point  $(q_1^*, q_1^{-i*})$  for any given  $(w_1, w_2)$  where  $h(q_1^*, q_1^{-i*}) = 0$ .

*Step 2:* Suppose  $\lambda \rightarrow 0$ , by the system of equations for ASPE (52), (53), (54) and (55),  $q_1^* = q_1^{-i*} = \frac{2w_1 - w_2}{w_1 + w_2}$ , workers will behave symmetrically. This showcase the outcome of complete information. However, in this model,  $\lambda$  may be negligible but remain positive,  $\lambda \rightarrow 0^+$ , workers would not behave symmetrically, but can be fairly close.

Then, suppose  $\lambda \rightarrow \infty$ , rewriting workers' FOCs (120) and (121), taking a limit on both sides,

$$\begin{aligned} \text{LHS: } \lim_{\lambda \rightarrow \infty} \frac{(1 - q_1^{-i*} + \frac{q_1^{-i*}}{2})w_1^* - (q_1^{-i*} + \frac{1 - q_1^{-i*}}{2})w_2^*}{\log(\frac{1 - p_1}{p_1} \frac{q_1^*}{1 - q_1^*})} &= \frac{(2 - q_1^{-i*})w_1^* - (1 + q_1^{-i*})w_2^*}{2 \log(\frac{1 - p_1}{p_1} \frac{q_1^*}{1 - q_1^*})} \\ \text{RHS: } \lim_{\lambda \rightarrow \infty} \lambda &= \infty \end{aligned}$$

For the equation to hold,  $q_1^*$  needs to converge to  $p_1$ ,

$$\begin{aligned} \lim_{q_1^* \rightarrow p_1^-} \frac{(2 - q_1^{-i*})w_1 - (1 + q_1^{-i*})w_2}{2 \log(\frac{1 - p_1}{p_1} \frac{q_1^*}{1 - q_1^*})} &= \infty \text{ if } (2 - q_1^{-i*})w_1 - (1 + q_1^{-i*})w_2 < 0 \\ \lim_{q_1^* \rightarrow p_1^+} \frac{(2 - q_1^{-i*})w_1 - (1 + q_1^{-i*})w_2}{2 \log(\frac{1 - p_1}{p_1} \frac{q_1^*}{1 - q_1^*})} &= \infty \text{ if } (2 - q_1^{-i*})w_1 - (1 + q_1^{-i*})w_2 > 0 \end{aligned}$$

The same applies for (121), as  $\lambda \rightarrow \infty$ , limit on both sides equalize only if  $q_1^{-i*} \rightarrow p_1^{-i}$ .

$$q_1^* \rightarrow p_1, q_1^{-i*} \rightarrow p_1^{-i} \text{ for } \lambda \rightarrow \infty$$

Workers converge to playing their default strategies. Since  $p_1 \neq p_1^{-i}$ , workers' choice probabilities in the equilibrium differ. As a result, there is at least one ASE for any  $\lambda > 0$ .

*Step 3:* Conditions for multiple equilibria.

Based on workers' FOCs, (120) and (121), find the limit as  $q_1^* \rightarrow 1, q_1^{-i*} \rightarrow 0$ . For (120),

$$\begin{aligned} \frac{p_1}{1-p_1} \exp\left(\frac{(2-q_1^{-i*})w_1 - (1+q_1^{-i*})w_2}{2\lambda}\right) &= \frac{q_1^*}{1-q_1^*} \\ \text{LHS: } \frac{p_1}{1-p_1} \lim_{q_1^* \rightarrow 1, q_1^{-i*} \rightarrow 0} \exp\left(\frac{(2-q_1^{-i*})w_1 - (1+q_1^{-i*})w_2}{2\lambda}\right) &= \frac{p_1}{1-p_1} \exp\left(\frac{2w_1 - w_2}{2\lambda}\right) \\ \text{RHS: } \lim_{q_1^* \rightarrow 1, q_1^{-i*} \rightarrow 0} \frac{q_1^*}{1-q_1^*} &= \infty \end{aligned}$$

For LHS=RHS,  $\lambda \rightarrow 0^+$  and  $2w_1 > w_2$  have to hold, such that  $\lim_{\lambda \rightarrow 0^+} \exp\left(\frac{2w_1 - w_2}{2\lambda}\right) = \infty$ . Otherwise, it can also be possible that if  $p_1 \rightarrow 1$ , then  $\lim_{p_1 \rightarrow 1} \frac{p_1}{1-p_1} \exp\left(\frac{2w_1 - w_2}{2\lambda}\right) = \infty$ .

Finding the limit for (121) ends up with similar conditions:  $\lambda \rightarrow 0^+$ ,  $2w_2 > w_1$ , or if  $p_1^{-i} \rightarrow 0$ .

For  $q_1^* \rightarrow 0, q_1^{-i*} \rightarrow 1$ , (120) shows

$$\begin{aligned} \frac{p_1}{1-p_1} \lim_{q_1^* \rightarrow 0, q_1^{-i*} \rightarrow 1} \exp\left(\frac{(2-q_1^{-i*})w_1 - (1+q_1^{-i*})w_2}{2\lambda}\right) &= \lim_{q_1^* \rightarrow 0, q_1^{-i*} \rightarrow 1} \frac{q_1^*}{1-q_1^*} \\ \text{LHS: } \frac{p_1}{1-p_1} \lim_{q_1^* \rightarrow 0, q_1^{-i*} \rightarrow 1} \exp\left(\frac{(2-q_1^{-i*})w_1 - (1+q_1^{-i*})w_2}{2\lambda}\right) &= \frac{p_1}{1-p_1} \exp\left(\frac{w_1 - 2w_2}{2\lambda}\right) \\ \text{RHS: } \lim_{q_1^* \rightarrow 0, q_1^{-i*} \rightarrow 1} \frac{q_1^*}{1-q_1^*} &= 0 \end{aligned}$$

For LHS=RHS,  $\lambda \rightarrow 0^+$  and  $w_1 < 2w_2$  need to hold, such that  $\lim_{\lambda \rightarrow 0^+} \exp\left(\frac{w_1 - 2w_2}{2\lambda}\right) = 0$ . Otherwise, if  $p_1 \rightarrow 0$ , then the equality is also satisfied, but I do not consider this case because the aforementioned assumption of  $p_1 > \frac{1}{2}$  and  $p_1^{-i} < \frac{1}{2}$  would be violated.

The same workings can be done for (121), these are not detailed.

For existence of ASE where workers apply more to different firms and almost purely coordinate, the conditions are similar to Lemma 2.2, where  $\lambda$  needs to be small relative to expected payoffs. But since defaults are allowed to vary, given the preset scenario of  $p_1 > \frac{1}{2}$  and  $p_1^{-i} < \frac{1}{2}$ , such ASE is also more likely to exist if workers are strongly biased to different firms (i.e.  $p_1 \rightarrow 1$  and  $p_1^{-i} \rightarrow 0$ ).

*Step 4:* I use the implicit function theorem to determine the presence of multiple solutions for the system of equations defined by (52), (53), (54) and (55). From (124) and (125), where  $f(\cdot)$  and  $g(\cdot)$  are continuously differentiable functions. Define the Jacobian matrix to

$$\text{be } J(f, g) = \begin{bmatrix} \frac{\partial f}{\partial q_1^*} & \frac{\partial f}{\partial q_1^{-i*}} \\ \frac{\partial g}{\partial q_1^*} & \frac{\partial g}{\partial q_1^{-i*}} \end{bmatrix}.$$

$$\det(J(f, g)(a, b)) = \frac{\partial f}{\partial q_1^*}(a, b) \cdot \frac{\partial g}{\partial q_1^{-i*}}(a, b) - \frac{\partial f}{\partial q_1^{-i*}}(a, b) \cdot \frac{\partial g}{\partial q_1^*}(a, b) \quad (126)$$

For any point  $(a, b)$ ,  $q_1^* = a$  and  $q_1^{-i*} = b$ , if the Jacobian evaluated at that point  $\det(J(f, g)(a, b)) \neq 0$ , the theorem guarantees existence of solution for the system of equations  $f(\cdot) = 0$  and  $g(\cdot) = 0$

in the neighborhood of each of these points. (Wolfram (2023)), (Royster (2023)) For possibility of multiple solutions, there is a need to check for many points  $(a, b)$  in the domain where  $q_1^* \in (0, 1)$  and  $q_1^{-i*} \in (0, 1)$ , and find  $\det(J(f, g)(a, b))$  that fulfills  $\det(J(f, g)(a, b)) \neq 0$  while also satisfying the relations  $f(\cdot) = 0$  and  $g(\cdot) = 0$ .

Evaluating for the general case of  $(a, b)$ ,

$$\begin{aligned}\frac{\partial f}{\partial q_1^*} &= \frac{p_1}{1-p_1} \frac{1}{2\lambda} \exp\left(\frac{(2-q_1^{-i*})w_1 - (1+q_1^{-i*})w_2}{2\lambda}\right) \left[(2-q_1^{-i*}) \frac{\partial w_1}{\partial q_1^*} - (1+q_1^{-i*}) \frac{\partial w_2}{\partial q_1^*}\right] - \frac{1}{(1-q_1^*)^2} \\ \frac{\partial f}{\partial q_1^{-i*}} &= \frac{p_1}{1-p_1} \frac{1}{2\lambda} \exp\left(\frac{(2-q_1^{-i*})w_1 - (1+q_1^{-i*})w_2}{2\lambda}\right) \left[(2-q_1^{-i*}) \frac{\partial w_1}{\partial q_1^{-i*}} - (1+q_1^{-i*}) \frac{\partial w_2}{\partial q_1^{-i*}} - w_1 - w_2\right] \\ \frac{\partial g}{\partial q_1^*} &= \frac{p_1^{-i}}{1-p_1^{-i}} \frac{1}{2\lambda} \exp\left(\frac{(2-q_1^*)w_1 - (1+q_1^*)w_2}{2\lambda}\right) \left[(2-q_1^*) \frac{\partial w_1}{\partial q_1^*} - (1+q_1^*) \frac{\partial w_2}{\partial q_1^*} - w_1 - w_2\right] \\ \frac{\partial g}{\partial q_1^{-i*}} &= \frac{p_1^{-i}}{1-p_1^{-i}} \frac{1}{2\lambda} \exp\left(\frac{(2-q_1^*)w_1 - (1+q_1^*)w_2}{2\lambda}\right) \left[(2-q_1^*) \frac{\partial w_1}{\partial q_1^{-i*}} - (1+q_1^*) \frac{\partial w_2}{\partial q_1^{-i*}}\right] - \frac{1}{(1-q_1^{-i*})^2}\end{aligned}$$

$\det(J(f, g))$  for  $(a, b)$  is dependent on relative belief distortion (as per definition (3.1)), cognitive costs and equilibrium wages.

Suppose  $\lambda \rightarrow \infty$ , a point that satisfies the relations  $f(\cdot) = 0$  and  $g(\cdot) = 0$  would be  $(q_1^*, q_1^{-i*}) = (p_1, p_1^{-i})$ . At this point,  $\det(J(f, g)(p_1, p_1^{-i})) = \frac{1}{(1-p_1)^2} \frac{1}{(1-p_1^{-i})^2} > 0$ . Since the determinant is non-zero within the region of interest, the theorem guarantees a solution in the neighborhood of the point.

For extreme values of  $q_1^*$  and  $q_1^{-i*}$ , such as  $q_1^* \rightarrow 0$  and  $q_1^{-i*}$ ,

$$\begin{aligned}\det(J(f, g)) &= \left(\frac{p_1}{1-p_1} \frac{1}{2\lambda} \exp\left(\frac{w_1 - 2w_2}{2\lambda}\right) \left(\frac{\partial w_1}{\partial q_1} - 2 \frac{\partial w_2}{\partial q_1}\right) - 1\right) \\ &\quad \left(\frac{p_1^{-i}}{1-p_1^{-i}} \frac{1}{2\lambda} \exp\left(\frac{2w_1 - w_2}{2\lambda}\right) \left(2 \frac{\partial w_1}{\partial q_1^{-i*}} - \frac{\partial w_2}{\partial q_1^{-i*}}\right) - \infty\right) - \frac{p_1}{1-p_1} \frac{p_1^{-i}}{1-p_1^{-i}} \left(\frac{1}{2\lambda}\right)^2 \\ &\quad \exp\left(\frac{w_1 - 2w_2}{2\lambda}\right) \exp\left(\frac{2w_1 - w_2}{2\lambda}\right) \left(\frac{\partial w_1}{\partial q_1^{-i*}} - 2 \frac{\partial w_2}{\partial q_1^{-i*}} - w_1 - w_2\right) \left(2 \frac{\partial w_1}{\partial q_1^*} - 2 \frac{\partial w_2}{\partial q_1^*} - w_1 - w_2\right) > 0\end{aligned}$$

It was found that  $\det(J(f, g)) > 0$  for non-zero and reasonably small  $\lambda$  relative to expected payoffs. The same applies for  $q_1^* \rightarrow 1$  and  $q_1^{-i*} \rightarrow 0$ . Since the neighborhood of the two points do not overlap, it is likely there are two distinct solutions in the vicinity of these two points.

Default strategies also play a role. If  $\frac{p_1}{1-p_1} \frac{p_1^{-i}}{1-p_1^{-i}} = 1$ , there are no relative distortion, the impact of  $p_1$  and  $p_1^{-i}$  on  $\det(J(f, g))$  is likely to be small. If there exist relative distortion, such that  $\frac{p_1}{1-p_1} \frac{p_1^{-i}}{1-p_1^{-i}} > 1$  (for stronger worker 1 distortion) or  $< 1$  (for stronger worker 2 distortion), while  $\det(J(f, g))$  remain positive, increasing or decreasing  $\det(J(f, g))$  could lead to heightened complexities in the analysis of multiple equilibria, where it opens up the possibility that the current set of solutions may not fully encompass the range of equilibria that are known to potentially exist within the system. There could be an extended set of equilibria, which will not be characterized completely in this paper given their significance to the research question is not substantial relative to their analytical complexity.

Return to the Section 3.1.

## A.7 Heterogeneous Defaults on Equilibrium Search Behaviour

Given  $\lambda$ ,  $z_1$  and  $z_2$ , and the equilibrium solutions  $(q_1^{-i*}, q_1^*, w_1^*, w_2^*)$ , using workers' FOCs to find  $\frac{dq_1^{-i*}}{dp_1}$ ,  $\frac{dq_1^*}{dp_1}$ ,  $\frac{dq_1^{-i*}}{dp_1^{-i}}$  and  $\frac{dq_1^*}{dp_1^{-i}}$ , which demonstrate the impact of heterogeneous defaults on equilibrium strategies. Let  $a = \frac{2w_1^* - w_2^*}{w_1^* + w_2^*}$  and  $b = \frac{2\lambda}{w_1^* + w_2^*}$ ,

$$\begin{aligned}\frac{dq_1^{-i*}}{dp_1} &= \frac{\frac{da}{dp_1}(1 - \frac{b}{q_1^*(1-q_1^*)}) - \frac{db}{dp_1}[\log(\frac{1-p_1}{p_1} \frac{q_1^*}{1-q_1^*}) - \frac{b}{q_1^*(1-q_1^*)} \log(\frac{1-p_1^{-i}}{p_1^{-i}} \frac{q_1^{-i*}}{1-q_1^{-i*}})] + \frac{b}{p_1(1-p_1)}}{1 - \frac{b^2}{q_1^*(1-q_1^*)q_1^{-i*}(1-q_1^{-i*})}} \\ \frac{dq_1^*}{dp_1} &= \frac{\frac{da}{dp_1}(1 - \frac{b}{q_1^{-i*}(1-q_1^{-i*})}) - \frac{db}{dp_1}[\log(\frac{1-p_1}{p_1} \frac{q_1^{-i*}}{1-q_1^{-i*}}) - \frac{b}{q_1^{-i*}(1-q_1^{-i*})} \log(\frac{1-p_1}{p_1} \frac{q_1^*}{1-q_1^*})] - \frac{b^2}{p_1(1-p_1)q_1^{-i*}(1-q_1^{-i*})}}{1 - \frac{b^2}{q_1^*(1-q_1^*)q_1^{-i*}(1-q_1^{-i*})}} \\ \frac{dq_1^{-i*}}{dp_1^{-i}} &= \frac{\frac{da}{dp_1^{-i}}(1 - \frac{b}{q_1^*(1-q_1^*)}) - \frac{db}{dp_1^{-i}}[\log(\frac{1-p_1}{p_1} \frac{q_1^*}{1-q_1^*}) - \frac{b}{q_1^*(1-q_1^*)} \log(\frac{1-p_1^{-i}}{p_1^{-i}} \frac{q_1^{-i*}}{1-q_1^{-i*}})] - \frac{b^2}{p_1^{-i}(1-p_1^{-i})q_1^*(1-q_1^*)}}{1 - \frac{b^2}{q_1^*(1-q_1^*)q_1^{-i*}(1-q_1^{-i*})}} \\ \frac{dq_1^*}{dp_1^{-i}} &= \frac{\frac{da}{dp_1^{-i}}(1 - \frac{b}{q_1^{-i*}(1-q_1^{-i*})}) - \frac{db}{dp_1^{-i}}[\log(\frac{1-p_1^{-i}}{p_1^{-i}} \frac{q_1^{-i*}}{1-q_1^{-i*}}) - \frac{b}{q_1^{-i*}(1-q_1^{-i*})} \log(\frac{1-p_1}{p_1} \frac{q_1^*}{1-q_1^*})] + \frac{b}{p_1^{-i}(1-p_1^{-i})}}{1 - \frac{b^2}{q_1^*(1-q_1^*)q_1^{-i*}(1-q_1^{-i*})}}\end{aligned}$$

While the sign can be difficult to determine, which is not surprising given the possibility of equilibrium multiplicity, it can be observed from the derivatives that when  $\lambda$  is low,  $b$  would be low, so  $\frac{dq_1^{-i*}}{dp_1} > \frac{dq_1^*}{dp_1}$ , and  $\frac{dq_1^{-i*}}{dp_1^{-i}} < \frac{dq_1^*}{dp_1^{-i}}$ . This highlights when it is cheap to process information, workers are more responsive to changes in the opponent's default strategy than their own. On the other hand, when  $\lambda$  is high,  $b$  would be high, so  $\frac{dq_1^{-i*}}{dp_1} < \frac{dq_1^*}{dp_1}$  and  $\frac{dq_1^{-i*}}{dp_1^{-i}} > \frac{dq_1^*}{dp_1^{-i}}$ . This suggests when it is expensive to process information, workers are more sensitive to changes in their own default strategies than the opponent's.

Return to the Section 3.1.

## A.8 Heterogeneous Cognitive Costs on Equilibrium Wages in SSPE.

To show the impact of heterogeneous cognitive costs on equilibrium wages in SSPE. Based on equation (59),

$$\frac{dw^*}{d\lambda_1} = \frac{-64\lambda_2 + 8(z - w^*)}{-16w^* + 8\lambda_1 + 8\lambda_2 + 4z} \quad (127)$$

$$\frac{dw^*}{d\lambda_2} = \frac{-64\lambda_1 + 8(z - w^*)}{-16w^* + 8\lambda_1 + 8\lambda_2 + 4z} \quad (128)$$

As evident, assuming  $\lambda_1 > \lambda_2$ ,  $\frac{dw^*}{d\lambda_1} > \frac{dw^*}{d\lambda_2}$ .

Return to the Section 3.2.

## A.9 Existence of ASPE with Heterogeneous Cognitive Costs

To show ASE exists under certain conditions when workers have heterogeneous cognitive costs.

*Step 1.* I first prove there exists at least one ASE solution.

Rewriting (60) and (61),

$$\exp\left(\frac{(2 - q_1^{-i*})w_1 - (q_1^{-i*} + 1)w_2}{2\lambda_1}\right) - \frac{q_1^*}{1 - q_1^*} = 0 \quad (129)$$

$$\exp\left(\frac{(2 - q_1^*)w_1 - (q_1^* + 1)w_2}{2\lambda_2}\right) - \frac{q_1^{-i*}}{1 - q_1^{-i*}} = 0 \quad (130)$$

Set LHS of the equations to be  $f(\cdot)$  and  $g(\cdot)$  respectively.

$$f(\cdot) = \exp\left(\frac{(2 - q_1^{-i*})w_1 - (q_1^{-i*} + 1)w_2}{2\lambda_1}\right) - \frac{q_1^*}{1 - q_1^*} \quad (131)$$

$$g(\cdot) = \exp\left(\frac{(2 - q_1^*)w_1 - (q_1^* + 1)w_2}{2\lambda_2}\right) - \frac{q_1^{-i*}}{1 - q_1^{-i*}} \quad (132)$$

Given continuous  $q_1^* \in (0, 1)$  and  $q_1^{-i*} \in (0, 1)$ ,  $h(q_1^*, q_1^{-i*}) = (f(q_1^*, q_1^{-i*}), g(q_1^*, q_1^{-i*}))$  can map any combination of points from  $q_1^*-q_1^{-i*}$  to the  $f$ - $g$  plane. Firstly, check for  $f(q_1^*, 0) - g(q_1^*, 0)$  as  $q_1^*$  changes value.

For  $q_1^* = 0$ ,

$$f(\cdot) - g(\cdot) = \exp\left(\frac{2w_1 - w_2}{2\lambda_1}\right) - \exp\left(\frac{2w_1 - w_2}{2\lambda_2}\right) < 0$$

For  $q_1^* = 1$ ,

$$f(\cdot) - g(\cdot) = \exp\left(\frac{2w_1 - w_2}{2\lambda_1}\right) - \infty - \exp\left(\frac{w_1 - 2w_2}{2\lambda_2}\right) < 0$$

Assuming  $\lambda_1 > \lambda_2$ , for low  $\lambda$ s relative to expected payoff and  $2w_1 > w_2$ , both equations above are negative.

Secondly, check for  $f(q_1^*, 1) - g(q_1^*, 1)$  as  $q_1^*$  changes value. For  $q_1^* = 0$ ,

$$f(\cdot) - g(\cdot) = \exp\left(\frac{w_1 - 2w_2}{2\lambda_1}\right) - \exp\left(\frac{2w_1 - w_2}{2\lambda_2}\right) + \infty > 0$$

For  $q_1^* = 1$ ,

$$f(\cdot) - g(\cdot) = \exp\left(\frac{w_1 - 2w_2}{2\lambda_1}\right) - \infty - \exp\left(\frac{w_1 - 2w_2}{2\lambda_2}\right) + \infty > 0$$

Similarly, given  $\lambda_1 > \lambda_2$ , for low  $\lambda$ s relative to expected payoff and  $w_1 < 2w_2$ , both equations would be positive.

These shows that for small  $\lambda$ s relative to expected payoff and wage condition  $2w_2 > w_1 > \frac{w_2}{2}$ , since  $f(q_1^*, 0) - g(q_1^*, 0)$  is always negative and  $f(q_1^*, 1) - g(q_1^*, 1)$  is always positive, the function  $f(\cdot) - g(\cdot)$  must change signs at some point in the interior of  $[0, 1] \times [0, 1]$ , such that it takes negative value at one boundary and positive value at the other. As a result, by intermediate value theorem, there exists at least one solution to the system of equations.

*Step 2.* For more extreme asymmetric workers' strategies exist. Using (60) and (61) and finding the limit as  $q_1^* \rightarrow 1$ , and  $q_1^{-i*} \rightarrow 0$ . For (60),

$$\exp\left(\frac{(2 - q_1^{-i*})w_1 - (q_1^{-i*} + 1)w_2}{2\lambda_1}\right) = \frac{q_1^*}{1 - q_1^*}$$

$$\text{LHS: } \lim_{q_1^* \rightarrow 1, q_1^{-i*} \rightarrow 0} \exp\left(\frac{(2 - q_1^{-i*})w_1 - (q_1^{-i*} + 1)w_2}{2\lambda_1}\right) = \exp\left(\frac{2w_1 - w_2}{2\lambda_1}\right)$$

$$\text{RHS: } \lim_{q_1^* \rightarrow 1, q_1^{-i*} \rightarrow 0} \frac{q_1^*}{1 - q_1^*} = \infty$$

To satisfy LHS=RHS,  $\lambda_1 \rightarrow 0^+$ ,  $2w_1^* > w_2^*$ .

Similarly, finding the limit for (61) as  $q_1^* \rightarrow 1$ , and  $q_1^{-i*} \rightarrow 0$ ,

$$\exp\left(\frac{(2 - q_1^*)w_1 - (q_1^* + 1)w_2}{2\lambda_2}\right) = \frac{q_1^{-i*}}{1 - q_1^{-i*}}$$

$$\text{LHS: } \lim_{q_1^* \rightarrow 1, q_1^{-i*} \rightarrow 0} \exp\left(\frac{(2 - q_1^*)w_1 - (q_1^* + 1)w_2}{2\lambda_2}\right) = \exp\left(\frac{w_1 - 2w_2}{2\lambda_2}\right)$$

$$\text{RHS: } \lim_{q_1^* \rightarrow 1, q_1^{-i*} \rightarrow 0} \frac{q_1^{-i*}}{1 - q_1^{-i*}} = 0$$

To satisfy LHS=RHS,  $\lambda_2 \rightarrow 0^+$ ,  $2w_2^* > w_1^*$ .

Furthermore, finding the limit as  $q_1^* \rightarrow 0$ , and  $q_1^{-i*} \rightarrow 1$ . For (60),

$$\text{LHS: } \lim_{q_1^* \rightarrow 0, q_1^{-i*} \rightarrow 1} \exp\left(\frac{(2 - q_1^{-i*})w_1 - (q_1^{-i*} + 1)w_2}{2\lambda_1}\right) = \exp\left(\frac{w_1 - 2w_2}{2\lambda_1}\right)$$

$$\text{RHS: } \lim_{q_1^* \rightarrow 0, q_1^{-i*} \rightarrow 1} \frac{q_1^*}{1 - q_1^*} = 0$$

To satisfy LHS=RHS,  $\lambda_1 \rightarrow 0^+$ ,  $2w_2 > w_1$ . For (61),

$$\text{LHS: } \lim_{q_1^* \rightarrow 0, q_1^{-i*} \rightarrow 1} \exp\left(\frac{(2 - q_1^*)w_1 - (q_1^* + 1)w_2}{2\lambda_2}\right) = \exp\left(\frac{2w_1 - w_2}{2\lambda_2}\right)$$

$$\text{RHS: } \lim_{q_1^* \rightarrow 0, q_1^{-i*} \rightarrow 1} \frac{q_1^{-i*}}{1 - q_1^{-i*}} = \infty$$

To satisfy LHS=RHS,  $\lambda_2 \rightarrow 0^+$ ,  $2w_1 > w_2$ .

The conditions for ASPE, where workers almost purely coordinate, to exist, would be  $\lambda_1 \rightarrow 0^+$ ,  $\lambda_2 \rightarrow 0^+$  and  $2w_2^* > w_1^* > \frac{w_2^*}{2}$ .

*Step 3.* The implicit function theorem can be used to determine the presence of multiple solutions. Using (131) and (A.9), I can find the Jacobian matrix and the corresponding determinant,  $\det(J(f, g)(a, b))$ . For Jacobian evaluated at the point that satisfies the relations  $f(\cdot) = 0$  and  $g(\cdot) = 0$ , if  $\det(J(f, g)(a, b)) \neq 0$ , the theorem would guarantee existence of solution in the neighborhood of the point. Since the steps are very similar to the past section of the paper, the details are omitted. The main point to note is that magnitude of  $\lambda_1$  and  $\lambda_2$  are independently and also jointly important in influencing the equilibrium multiplicity.

Return to the Section 3.2.

## A.10 Heterogeneous Cognitive Costs on Equilibrium Strategies in ASPE

To demonstrate heterogeneous cognitive costs on equilibrium strategies, I first start with (60) and (61),

$$\exp\left(\frac{(2 - q_1^{-i*})w_1^* - (q_1^{-i*} + 1)w_2^*}{2\lambda_1}\right) - \frac{q_1^*}{1 - q_1^*} = 0 \quad (133)$$

$$\exp\left(\frac{(2 - q_1^*)w_1^* - (q_1^* + 1)w_2^*}{2\lambda_2}\right) - \frac{q_1^{-i*}}{1 - q_1^{-i*}} = 0 \quad (134)$$

Set LHS of (133) and (134) to be  $f(\cdot)$  and  $g(\cdot)$  respectively.

The SSPE solution for heterogeneous cognitive costs is  $q_1^* = q_1^{-i*} = \frac{1}{2}$  and  $w_1^* = w_2^*$ . For  $q_1^* = \frac{1}{2}$ ,

$$f(\cdot) = \exp\left(\frac{(2 - q_1^{-i*})w_1^* - (q_1^{-i*} + 1)w_2^*}{2\lambda_1}\right) - 1$$

$f(\cdot) = 0$  holds for  $q_1^{-i*} = \frac{1}{2}$  and  $w_1^* = w_2^*$ . Assuming existence of ASPE,  $q_1^* \neq q_1^{-i*}$ . Suppose  $q_1^{-i*} > \frac{1}{2}$ ,  $f(\cdot)$  will decrease. In order to fulfill  $f(\cdot) = 0$ ,  $\exp(\cdot)$  has to increase, through increasing  $w_1$  and/or decreasing  $w_2$ , assuming  $z_1 > z_2$ . This implies if ASPE exist, equilibrium wages have to be further apart. Correspondingly, for  $g(\cdot)$ , when  $q_1^* = \frac{1}{2}$ ,

$$g(\cdot) = \exp\left(\frac{\frac{3}{2}(w_1^* - w_2^*)}{2\lambda_2}\right) - \frac{q_1^{-i*}}{1 - q_1^{-i*}}$$

For  $q_1^{-i*} > \frac{1}{2}$ ,  $g(\cdot)$  will decrease. To maintain  $g(\cdot) = 0$ ,  $\exp(\cdot)$  has to increase, implying  $w_1$  increases and/or  $w_2$  decreases. The changes are matched for both equations. This indicates that for worker 2 to apply more to firm 1 (the more productive firm) than worker 1, equilibrium wages should be further apart, in which case, gathering more precise information is more attractive and the more attentive worker would want to reap the benefit from the lower information costs.

On the other hand, if  $q_1^{-i*} < \frac{1}{2}$ ,  $f(\cdot)$  will increase, and to fulfill  $f(\cdot) = 0$ ,  $\exp(\cdot)$  has to decrease, which can be met if  $w_1$  decrease and/or  $w_2$  increase, assuming  $z_1 > z_2$ . Correspondingly, as  $q_1^{-i*}$  decreases,  $g(\cdot)$  increases, and to maintain  $g(\cdot) = 0$ ,  $\exp(\cdot)$  has to decrease, which can be met if  $w_1$  decreases and/or  $w_2$  increases. The changes are matched for both equations again. This implies for worker 2 to apply less to the more productive firm 1 than worker 1, equilibrium wages should be closer together. The benefit from higher matching probability by applying to the slightly less productive firm is more attractive.

Return to the Section 3.2.

## A.11 Inefficient SSPE

Herein, I study the efficiency of workers' SSPE strategy,  $q^*$ , as compared to the socially optimal strategy,  $\hat{q}$ .

*Step 1:* I first evaluate SSPE solution at extreme values of cognitive costs. When  $\lambda \rightarrow 0^+$ , from (71), the socially optimal  $\hat{q}$  is

$$\lim_{\lambda \rightarrow 0^+} (z_1 - \hat{q}(z_1 + z_2) - \lambda \log\left(\frac{1-p}{p} \frac{\hat{q}}{1-\hat{q}}\right)) = 0$$

$$\lim_{\lambda \rightarrow 0^+} \hat{q} = \frac{z_1}{z_1 + z_2} \quad (135)$$

Based on (12), the equilibrium  $q^*$  is

$$\lim_{\lambda \rightarrow 0^+} (2w_1^* - w_2^* - q^*(w_1^* + w_2^*) - 2\lambda \log(\frac{1-p}{p} \frac{q^*}{1-q^*})) = 0$$

$$\lim_{\lambda \rightarrow 0^+} q^* = \frac{2w_1^* - w_2^*}{w_1^* + w_2^*}, \text{ where } q^* \in (0, 1) \quad (136)$$

where  $w_1^*$  and  $w_2^*$  are determined by (17) and (18).

Suppose prior beliefs are accurate, such that  $p = \frac{z_1}{z_1+z_2}$ . Then for the simplest case of homogeneous firms,  $z_1 = z_2$ ,  $w_1^* = w_2^*$ , if  $\lambda \rightarrow 0^+$ ,  $q^* = \frac{1}{2}$  and  $\hat{q} = \frac{1}{2}$ , the equilibrium strategy would be efficient.

If  $\lambda \rightarrow \infty$ ,

$$\lim_{\lambda \rightarrow \infty} \hat{q} = p, \lim_{\lambda \rightarrow \infty} q^* = p \quad (137)$$

Both the equilibrium strategy and the socially optimal strategy converge to the default strategy, the equilibrium strategy is efficient.

Suppose prior beliefs are incorrect, such that  $p \neq \frac{z_1}{z_1+z_2}$ . When facing homogeneous firms, the same results hold for  $\lambda \rightarrow 0^+$  and  $\lambda \rightarrow \infty$ . This suggests that given firms are homogeneous,  $q^* = \hat{q}$  regardless of prior beliefs being correct or incorrect at extreme  $\lambda$  values.

*Step 2:* Relaxing the assumption on extreme values of  $\lambda$ , and evaluating for intermediary range of  $\lambda$ s.

If firms remain homogeneous ( $z_1 = z_2 = z$ ,  $w_1^* = w_2^* = w^*$ ), and workers have the correct prior beliefs ( $p = \frac{1}{2}$ ), then by (71) and (12), workers' equilibrium and socially optimal strategies remain the same,  $q^* = \hat{q} = \frac{1}{2}$ , the equilibrium outcome is efficient. However, if workers have incorrect prior beliefs, default strategies would affect both  $q^*$  and  $\hat{q}$ .

Suppose the incorrect prior information leads to a lower application probability to firm 1 by default (i.e.  $p < \frac{1}{2}$ ). When  $\lambda \rightarrow 0^+$ ,  $\hat{q} = \frac{1}{2}$  and  $q^* = \frac{1}{2}$ . As  $\lambda$  increases, both  $q^*$  and  $\hat{q}$  converge to  $p$ , which lays below  $\frac{1}{2}$ , indicating downward sloping  $q^*$  and  $\hat{q}$  curves.

Finding the change in  $\hat{q}$  and  $q^*$  given a change in  $\lambda$ ,  $\frac{d\hat{q}}{d\lambda}$  and  $\frac{dq^*}{d\lambda}$ ,

$$\frac{d\hat{q}}{d\lambda} = -\frac{\log(\frac{1-p}{p} \frac{\hat{q}}{1-\hat{q}})}{2z + \lambda \frac{1}{\hat{q}(1-\hat{q})}} \quad (138)$$

$$\frac{dq^*}{d\lambda} = \frac{(2-q^*) \frac{2q^*}{1-q^*} + (1+q^*) \frac{2(1-q^*)}{q^*} - 2\log(\frac{1-p}{p} \frac{q^*}{1-q^*})}{4(1-2q^*)z + (w_1^* + w_2^*) + \frac{4q^{*3}-8q^{*2}+2}{(q^*-1)^2q^{*2}}\lambda} \quad (139)$$

$q^*$ ,  $\hat{q}$  will converge at different rate.

When  $\lambda$  is small,  $|\frac{d\hat{q}}{d\lambda}| > |\frac{dq^*}{d\lambda}|$ , the speed of convergence is faster for  $\hat{q}$  as compared to  $q^*$ . As  $\lambda$  increases, the relationship could reverse. Since  $p < \frac{1}{2}$  and  $q^*$  and  $\hat{q}$  are downward sloping, these indicate that  $q^* > \hat{q}$  for small  $\lambda$ , workers should apply less to firm 1 in equilibrium, and at a point, as  $\lambda$  grows,  $\hat{q} > q^*$ . SSPE is not efficient in general, but there can be a threshold point of crossing,  $q^* = \hat{q}$ , where social welfare is the same.

This point is determined as,

$$q^* = \hat{q} = \frac{2(z - w_1^*) + w_2^*}{4z - (w_1^* + w_2^*)} = \frac{z + (z - w_1^*) - (w_1^* - w_2^*)}{2z + (z - w_1^*) + (z - w_2^*)} \quad (140)$$



It is dependent on the difference in equilibrium wages offered by the firms. Holding all else equal, if the wage differential is high, the threshold of crossing is expected to be low. This implies there can be a larger range of low  $\lambda$ s where  $q^* > \hat{q}$ , workers should have applied less to firm 1.

Next, suppose incorrect prior information induces a higher application probability to firm 1 by default (i.e.  $p > \frac{1}{2}$ ). For  $\lambda \rightarrow 0^+$ ,  $\hat{q} = \frac{1}{2}$  and  $q^* = \frac{1}{2}$ . As  $\lambda$  increases,  $\hat{q}$  and  $q^*$  converge to  $p$ ,  $\hat{q}$  and  $q^*$  are upward sloping curves.

Based on (138) and (139), when  $\lambda$  is small, for  $q^*$  and  $\hat{q}$  slightly higher than  $\frac{1}{2}$ ,  $|\frac{d\hat{q}}{d\lambda}| > |\frac{dq^*}{d\lambda}|$ . With larger  $\lambda$ , the relationship reverse. Since the curves are upward sloping, this indicates that at low  $\lambda$ ,  $\hat{q} > q^*$ , workers should have applied more to firm 1, and as  $\lambda$  increases,  $q^* > \hat{q}$ , workers should have applied less to firm 1. SSPE is inefficient in general, except at the threshold point of crossing, similarly characterized by (140), where social welfare will be the same. Higher wage differential would lead to crossing happening at lower  $q^*$  and  $\hat{q}$ , implying there could be a smaller range of low  $\lambda$ s where  $\hat{q} > q^*$ , where workers should have applied more to firm 1.

*Step 3: Relaxing the assumption on firm homogeneity.*

If firms are heterogeneous,  $z_1 \neq z_2$ , then as  $\lambda \rightarrow 0^+$ ,  $\hat{q} \rightarrow \frac{z_1}{z_1+z_2}$ , and  $q^* \rightarrow \frac{2w_1^*-w_2^*}{w_1^*+w_2^*}$ , with the wage determination equations,

$$\lim_{\lambda \rightarrow 0^+} q^* = \frac{2w_1^* - w_2^*}{w_1^* + w_2^*} = \frac{(3 - 2q^* - q^{*2})z_1 - (q^{*2} + 2q^*)z_2}{(2q^* - 2q^{*2})(z_1 + z_2)} \quad (141)$$

Assuming  $z_1 = xz_2$ ,  $x > 1$ , firm 1 is  $x$  times more productive than firm 2,

$$\lim_{\lambda \rightarrow 0^+} q^* = \frac{2w_1^* - w_2^*}{w_1^* + w_2^*} = \frac{(3 - 2q^* - q^{*2})x - (q^{*2} + 2q^*)}{(2q^* - 2q^{*2})(x + 1)} \quad (142)$$

and  $\lim_{\lambda \rightarrow 0^+} \hat{q} = \frac{z_1}{z_1+z_2} = \frac{x}{1+x}$ . For  $q^* \geq \frac{1}{2}$ ,  $\hat{q} > q^*$ , SSPE is inefficient.

As  $\lambda \rightarrow \infty$ , both  $\hat{q}$  and  $q^*$  would converge to the default strategy, then the equilibrium strategy would be efficient.

For intermediary range of  $\lambda$ s. If firms are heterogeneous, and workers have the correct prior beliefs,  $p = \frac{z_1}{z_1+z_2}$ . Finding the change in  $\hat{q}$  and  $q^*$  given a change in  $\lambda$ , but  $z_1 \neq z_2$ :

$$\frac{d\hat{q}}{d\lambda} = -\frac{\log(\frac{1-p}{p} \frac{\hat{q}}{1-\hat{q}})}{z_1 + z_2 + \lambda \frac{1}{\hat{q}(1-\hat{q})}} \quad (143)$$

$$\frac{dq^*}{d\lambda} = \frac{(2 - q^*) \frac{2q^*}{1-q^*} + (1 + q^*) \frac{2(1-q^*)}{q^*} - 2 \log(\frac{1-p}{p} \frac{q^*}{1-q^*})}{2(1 - 2q^*)(z_1 + z_2) + (w_1^* + w_2^*) + \frac{4q^{*3} - 8q^{*2} + 2}{(q^* - 1)^2 q^{*2}} \lambda} \quad (144)$$

If prior beliefs are correct,  $\frac{d\hat{q}}{d\lambda} = 0$ , and  $\frac{dq^*}{d\lambda} \neq 0$ , since  $\hat{q} > q^*$  holds even when  $\lambda \rightarrow 0^+$ , then for any intermediary  $\lambda$ , workers should be applying to firm 1 with higher probability.

Suppose prior beliefs are incorrect and a lower application probability to firm 1 by default is expected,  $p < \frac{z_1}{z_1+z_2}$ . Both  $\hat{q}$  and  $q^*$  are downward sloping, implying  $\hat{q} > q^*$  for any intermediary range of  $\lambda$ , workers should be applying to firm 1 with higher probability.

On the other hand, if prior beliefs are incorrect and a higher application probability to firm 1 by default is expected,  $p > \frac{z_1}{z_1 + z_2}$ . Since  $\lambda \rightarrow 0^+$ ,  $\hat{q} > q^*$ , and as  $\lambda$  increases,  $q^*$  and  $\hat{q}$  would converge to  $p$ , which lay above the initial starting point, this indicates that  $\hat{q}$  and  $q^*$  are upward sloping.

$$\text{Cases for } p > \frac{z_1}{z_1 + z_2} = \begin{cases} \hat{q} > q^*, \text{ given small } \lambda \\ q^* = \hat{q}, \text{ threshold crossing at } \frac{z_1 + (z_1 - w_1^*) - (w_1^* - w_2^*)}{z_1 + z_2 + (z_1 - w_1^*) + (z_2 - w_2^*)} \\ \hat{q} < q^*, \text{ given large } \lambda \end{cases} \quad (145)$$

Threshold  $\bar{\lambda}$ , where  $q^* = \hat{q}$ , for any given  $p$ , can be formulated as

$$\bar{\lambda} = \frac{z_1 - \frac{z_1 + (z_1 - w_1^*) - (w_1^* - w_2^*)}{z_1 + z_2 + (z_1 - w_1^*) + (z_2 - w_2^*)}(z_1 + z_2)}{\log\left(\frac{1-p}{p} \frac{z_2 + (z_2 - w_2^*) + (w_1^* - w_2^*)}{z_1 + z_2 + (z_1 - w_1^*) + (z_2 - w_2^*)}\right)} \quad (146)$$

The main difference between homogeneous and heterogeneous firms is that their productivity difference will also affect the speed of convergence and threshold crossing. Obstructing away from the threshold crossing point as well as when  $\lambda$  is sufficiently large such that both  $q^*$  and  $\hat{q}$  converge to default strategy, SSPE is generally inefficient. When facing homogeneous firms, SSPE is inefficient if prior beliefs are incorrect. When facing heterogeneous firms, SSPE is inefficient even if prior beliefs are accurate.

Return to the Section 4.1.

## A.12 Heterogeneous Defaults on Socially Optimal Strategy

To explore the impact of heterogeneous defaults on socially optimal solution,  $\hat{q}$ , based on (80), find  $\frac{d\hat{q}}{dp_1}$  and  $\frac{d\hat{q}}{dp_1^{-i}}$ :

$$\frac{d\hat{q}}{dp_1} = \frac{\frac{\lambda}{(1-p_1)p_1}}{2(z_1 + z_2 + \frac{\lambda}{(1-\hat{q})\hat{q}})} > 0, \frac{d\hat{q}}{dp_1^{-i}} = \frac{\frac{\lambda}{(1-p_1^{-i})p_1^{-i}}}{2(z_1 + z_2 + \frac{\lambda}{(1-\hat{q})\hat{q}})} > 0 \quad (147)$$

Express (80) in terms of  $\hat{q}$  on the RHS and  $p_1$  and  $p_1^{-i}$  on the LHS,

$$2z_1 - \lambda \log\left(\frac{1-p_1}{p_1} \frac{1-p_1^{-i}}{p_1^{-i}}\right) = 2\lambda \log\left(\frac{\hat{q}}{1-\hat{q}}\right) + 2\hat{q}(z_1 + z_2) \quad (148)$$

Given exogeneously determined value for  $z_1$ ,  $z_2$  and  $\lambda$ , when  $p_1 > \frac{1}{2}$  and  $p_1^{-i} < \frac{1}{2}$ ,

- if  $p_1 = 1 - p_1^{-i}$ , belief distortion is the same for the workers:  $\hat{q}$  will be independent of defaults.
- if  $p_1 > 1 - p_1^{-i}$ , worker 1 has a stronger belief distortion than worker 2: LHS of (148) becomes larger as relative asymmetry in default strategies (i.e.  $\frac{1-p_1}{p_1} \frac{1-p_1^{-i}}{p_1^{-i}}$ ) lowers,  $\hat{q}$  has to increase.
- if  $p_1 < 1 - p_1^{-i}$ , worker 2 has stronger belief distortion than worker 1: LHS of (148) becomes smaller as relative asymmetry in default strategies (i.e.  $\frac{1-p_1}{p_1} \frac{1-p_1^{-i}}{p_1^{-i}}$ ) increases,  $\hat{q}$  has to decrease.

As a result, socially optimal  $\hat{q}$  is pulled towards the worker with stronger belief distortion.

Return to Section 4.1.

### A.13 Heterogeneous Cognitive Costs on Socially Optimal Strategy

To determine the impact of heterogeneous cognitive costs on socially optimal solution,  $\hat{q}$ , based on (80), find  $\frac{d\hat{q}}{d\lambda_1}$  and  $\frac{d\hat{q}}{d\lambda_2}$ :

$$\frac{d\hat{q}}{d\lambda_1} = \frac{-\log(\frac{\hat{q}}{1-\hat{q}}) - \lambda_1(\frac{1}{\hat{q}} + \frac{1}{1-\hat{q}})}{2z_1 + 2z_2} \quad (149)$$

$$\frac{d\hat{q}}{d\lambda_2} = \frac{-\log(\frac{\hat{q}}{1-\hat{q}}) - \lambda_2(\frac{1}{\hat{q}} + \frac{1}{1-\hat{q}})}{2z_1 + 2z_2} \quad (150)$$

Given  $\lambda_1 > \lambda_2$ , change in  $\lambda_1$  by the same amount as  $\lambda_2$  will lead to a larger decrease in  $\hat{q}$ .

Return to Section 4.1.

### A.14 Social Welfare for Heterogeneous Workers

*Equilibrium Efficiency for Workers with Heterogeneous Defaults.* Since there are no SSPE when  $\lambda > 0$  for workers with heterogeneous defaults, I compare  $\hat{q}$  with ASPE solution,  $(q_1^*, q_1^{-i*})$  (Lemma 3.1). For social welfare to be higher under ASPE than socially optimal level,  $\mathbb{W}(q_1^*, q_1^{-i*}) > \mathbb{W}(\hat{q})$ ,

$$\begin{aligned} & ((1 - \hat{q})^2 - (1 - q_1^*)(1 - q_1^{-i*}))z_1 + (\hat{q}^2 - q_1^*q_1^{-i*})z_2 \\ & > \lambda \left[ \sum_{j=1}^2 q_j^* \log \frac{q_j^*}{p_j} + \sum_{j=1}^2 q_j^{-i*} \log \frac{q_j^{-i*}}{p_j^{-i}} - \sum_{j=1}^2 \hat{q} \log \frac{\hat{q}}{p_j} - \sum_{j=1}^2 \hat{q} \log \frac{\hat{q}}{p_j^{-i}} \right] \end{aligned} \quad (151)$$

Given possibility of multiple equilibria, it is difficult to evaluate this inequality. It weighs the difference in probability of firms meeting a worker on the LHS and the difference in attention costs incurred in the equilibrium and in the socially optimal scenario on the RHS. In general, if the expected return from probability of meeting a worker for the firms is substantially higher (with a greater weight assign to the more productive firm), and the difference in attention costs is minimal, then ASPE will be more efficient.

*Equilibrium Efficiency for Workers with Heterogeneous Cognitive Costs.* When workers have heterogeneous cognitive costs, the only possible SSPE is when  $q^* = \frac{1}{2}$ , and in which case, the socially optimal  $\hat{q}$  is the same, SSPE is efficient in this special case. As for comparison between  $\hat{q}$  and the ASPE solution,  $(q_1^*, q_1^{-i*})$ . Suppose  $p = \frac{1}{2}$ , then  $\mathbb{W}(q_1^*, q_1^{-i*}) > \mathbb{W}(\hat{q})$  if

$$\begin{aligned} & ((1 - \hat{q})^2 - (1 - q_1^*)(1 - q_1^{-i*})) \frac{z_1}{\lambda_1 + \lambda_2} + (\hat{q}^2 - q_1^*q_1^{-i*}) \frac{z_2}{\lambda_1 + \lambda_2} > \frac{\lambda_1}{\lambda_1 + \lambda_2} \left( \sum_{j=1}^2 q_j^* \log \frac{q_j^*}{\frac{1}{2}} \right. \\ & \left. - \hat{q} \log \frac{\hat{q}}{\frac{1}{2}} - (1 - \hat{q}) \log \frac{1 - \hat{q}}{\frac{1}{2}} \right) + \frac{\lambda_2}{\lambda_1 + \lambda_2} \left( \sum_{j=1}^2 q_j^{-i*} \log \frac{q_j^{-i*}}{\frac{1}{2}} - \hat{q} \log \frac{\hat{q}}{\frac{1}{2}} - (1 - \hat{q}) \log \frac{1 - \hat{q}}{\frac{1}{2}} \right) \end{aligned} \quad (152)$$

Similarly, given the possibility of multiple equilibria, it is difficult to evaluate this inequality. When firms' productivity are high and sum of cognitive costs is low, it would be more likely for social welfare to be higher under ASPE than under the socially optimal solution. But degree of heterogeneity matters. Larger weight is placed on the difference in probability of meeting a worker in the equilibrium vs. under socially optimal solution for the more productive firm, and the higher cognitive cost worker would have greater impact on whether asymmetric strategy would lead to better outcome than socially optimal solution. For instance, if  $z_1$  is substantially higher than  $z_2$ , and  $\frac{z_1}{\lambda_1 + \lambda_2}$  is larger, the difference in probability of firm 1 meeting a worker is

more important in determining if welfare is to be higher in ASPE than with socially optimal strategy. Also, the weight allocated to the difference in attention costs for different workers is determined by the proportion of their respective cognitive cost against the sum of cognitive costs,  $\frac{\lambda_i}{\lambda_i + \lambda_{-i}}$ ). If  $\lambda_1$  is high, then the weight placed on worker 1's attention cost difference in these two scenarios would be larger. While total productivity and aggregate cognitive costs contribute to fulfilling the condition, but with large  $z_1$  and/or large  $\lambda_1$  that imply huge disparity between firms and/or workers, the upholding of the condition could be subjected to circumstances of the more productive firm and the more cognitive constrained worker.

Return to Section 4.1

## B Online Appendix: Examples

### B.1 Job Search Process

The main narrative of workers' decision process:

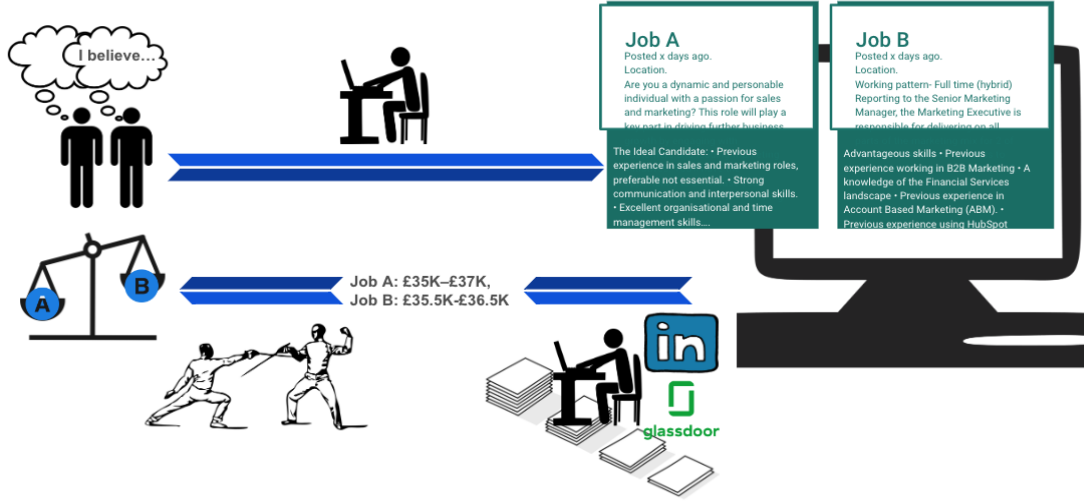


Figure 16: Workers' Decision Process

**Acquiring and processing wage information.** Wage information are not explicitly displayed. The information acquisition process comprises of workers attempting to reduce uncertainty about the wage distribution. To learn the wage information with greater precision, one would need to put in more efforts and do more research. On the broader scale, one can research on the type of business and sector that the firms are operating in, and the average wage that are typically offered. For more in-depth research, one can cross-check with other platforms that may consist of different advertising details of the same positions, or looking at platforms that comprise of reviews from interviewees, past or current employees. One can also connect with those who could hold more precise wage information about the firms and the vacancies via social media. These would provide more precise estimates of the wages, but are also assumed to be more costly.

Return to Section 2.1.

### B.2 Impact of Default Strategies on Equilibrium Strategies and Wages in SSPE

To visualize the impact of default strategies on equilibrium application strategies and wages in SSPE, I plot the changes in  $p$  on  $q^*$ , given some exogenous values of  $\lambda$ ,  $z_1$  and  $z_2$ .

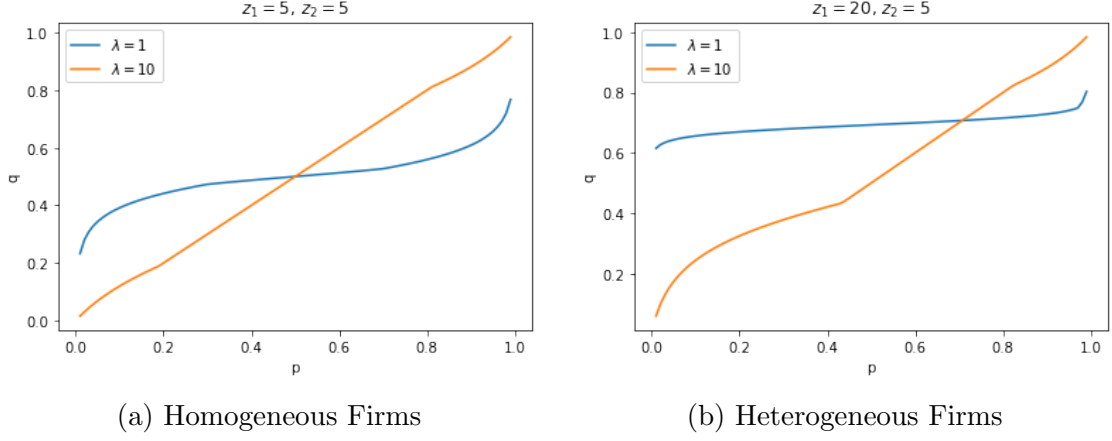


Figure 17: Effects of Default Strategies on Equilibrium Application Probabilities

Figure 17 shows that a higher default application probability towards firm 1 translates into higher equilibrium application probability towards firm 1, as proven in Proposition 1.

Relating to Proposition 2. For low cognitive costs, the equilibrium application probability is represented by a nearly horizontal line that revolves around  $\frac{1}{2}$  when firms are homogeneous, and assign higher application probability towards the more productive firm 1 across all  $p$  values when firms are heterogeneous. For large cognitive costs, there is an almost one to one transformation between changes in default strategies and changes in equilibrium strategies, represented by the approximately 45° line for both graphs.

For changes in default strategies on equilibrium wages,  $(w_1^*, w_2^*)$ :

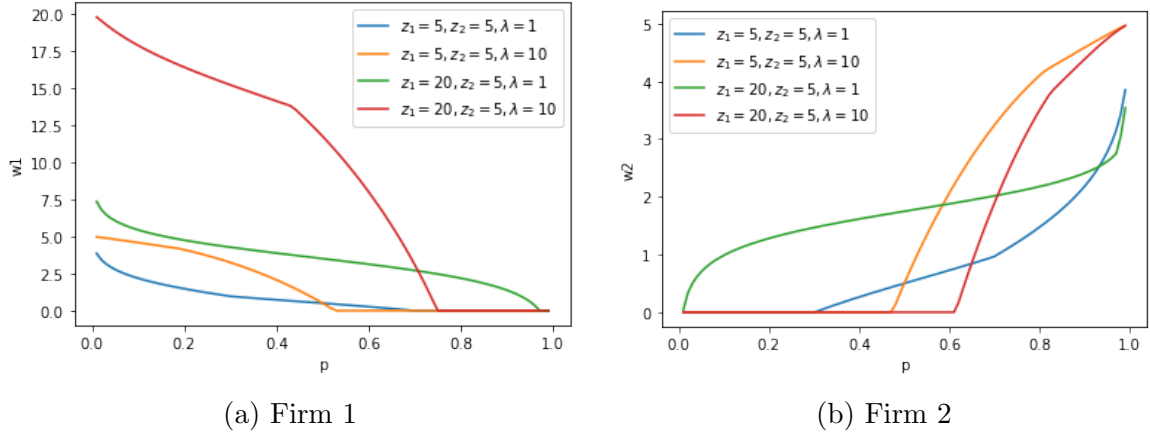


Figure 18: Effects of Default Strategies on Equilibrium Wages

Figure 18 shows a clear decreasing trend for  $w_1^*$  and increasing trend for  $w_2^*$  when  $p$  increases, complying with proof for Proposition 1, whereby in the case of skewed defaults, the firm with lower expected probability of meeting a worker will set a higher wage in order to attract applications.

Return to the Section 2.3.1.

### B.3 Examples of Equilibrium Solutions

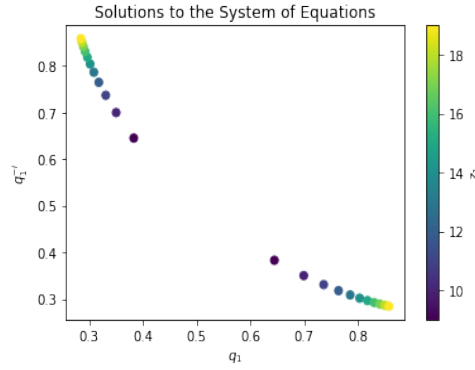
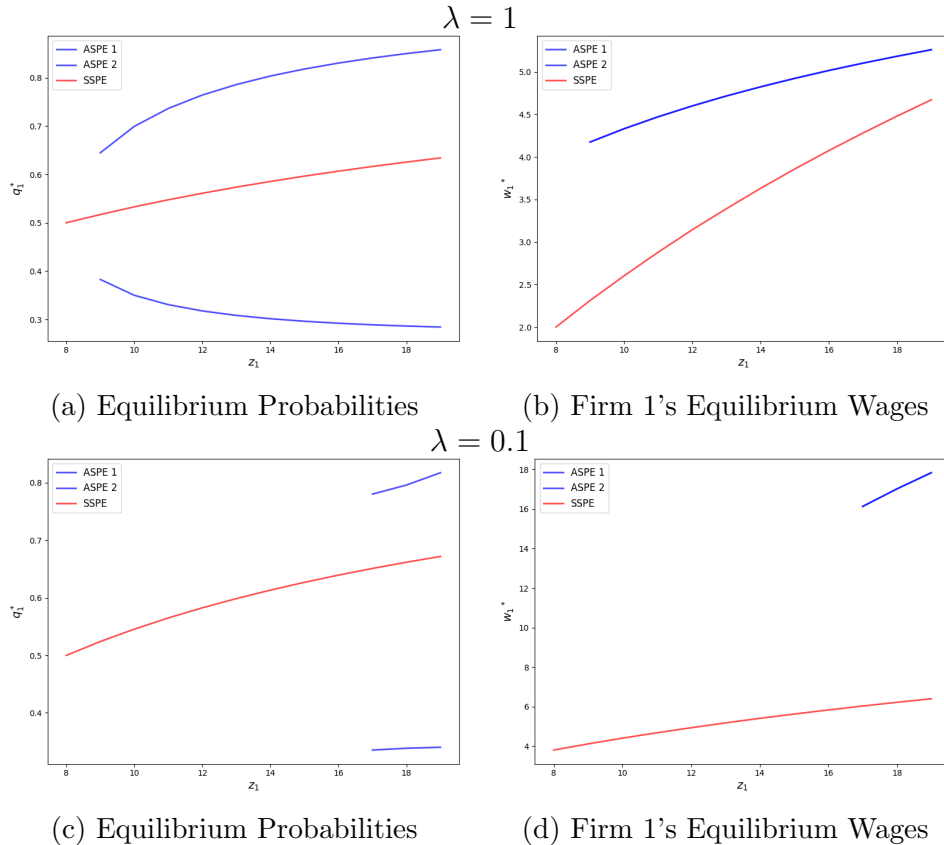


Figure shows ASPE solutions when  $\lambda = 1$ ,  $z_1 \in [8, 20]$ ,  $z_2 = 8$ ,  $p_1 = p_1^{-i} = \frac{1}{2}$ . The colour spectrum displays a range of  $z_1$  while fixing  $z_2$  to show the impact of heterogeneous productivity on workers' asymmetric equilibrium behaviour.

Figure 19: Workers' Asymmetric Application Probability to Firm 1

*Possible ASPE solutions.* Prior to proving the existence of ASPE, I compute some numerical examples to paint a better picture of what to expect. Figure 19 shows numerically that ASPE solutions, if exist, comprise of  $q_1^{-i*}$  and  $q_1^*$  going in opposite directions.



Figures show bifurcation graphs with  $q_1^*$  and  $w_1^*$  plot against changes in  $z_1$ ,  $z_1 \in [8, 20]$ ,  $z_2 = 8$ ,  $p_1 = p_1^{-i} = \frac{1}{2}$ .

Figure 20: Effects of Varying Productivity on Equilibrium Probabilities and Wages

*Equilibrium Solutions with different  $\lambda$ .* Figure 20 demonstrates: (1) There could be three possible equilibria, SSPE always exist (red) and ASPE only exist when all necessary conditions are

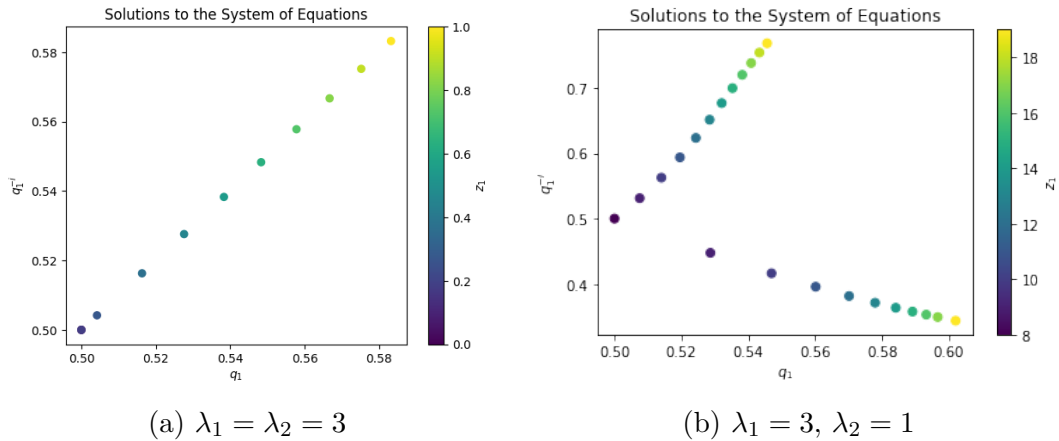
fulfilled (blue) – lowering  $\lambda$  does not necessarily imply existence of ASPE as wage condition may not be fulfilled.<sup>9</sup> (2) The equilibrium wages are lower when  $\lambda$  is high, demonstrating that firms have greater power in penalizing the workers when information costs are higher. In between ASPE and SSPE, ASPE could comprise of higher equilibrium wages than SSPE.

Return to the Section 2.3.1.

## B.4 Workers' Equilibrium Strategies Given Homogeneous and Heterogeneous Cognitive Costs

The independent cognitive cost of each worker as well as their combined cognitive costs would influence the existence of equilibrium multiplicity. The specifics depends on the actual values of the exogenous parameters. For the social welfare comparison, I used a few numerical examples, where welfare computation is based on workers' equilibrium strategies, so below I show workers' behaviour for different cognitive costs.

*High combined cognitive costs.* It is possible that only SSPE exists when workers have homogeneous and high cognitive costs. By reducing one of the workers' cognitive costs, effectively increasing worker heterogeneity, could render ASPE possible.



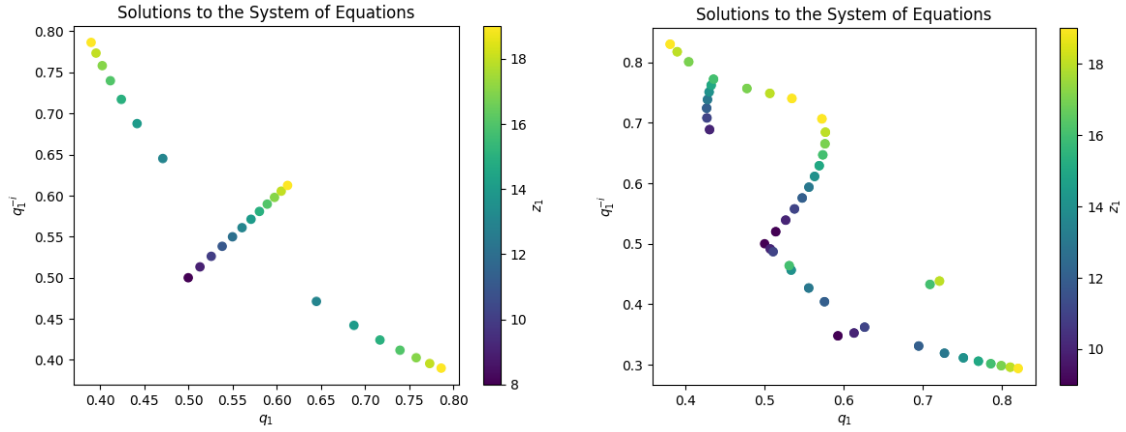
For  $z_1 \in [8, 20]$ ,  $z_2 = 8$ ,  $p_1 = p_1^{-i} = 0.5$ .

Figure 21: Equilibrium Solutions with High Cognitive Costs

*Low combined cognitive costs.* When cognitive costs are low, both SSPE and ASPE exist when workers are homogeneous. By reducing one of the workers' cognitive costs, more equilibria might emerge.

<sup>9</sup>To illustrate the example of  $\lambda = 0.1$ : As  $\lambda$  is lowered, workers know almost precisely which firm is offering a higher wage, they will apply to that firm with high probability. There would be fiercer competition between the firms, resulting in low wage differentials. In anticipation of wage setting behaviour,  $\lambda$  may no longer be small relative to expected payoff, thus the wage condition for ASE to exist is not fulfilled, workers do not apply asymmetrically in the equilibrium (Lemma 2.2).





(a)  $\lambda_1 = \lambda_2 = 1.2$

(b)  $\lambda_1 = 1.2, \lambda_2 = 1$

For  $z_1 \in [8, 20]$ ,  $z_2 = 8$ ,  $p_1 = p_1^{-i} = 0.5$ .

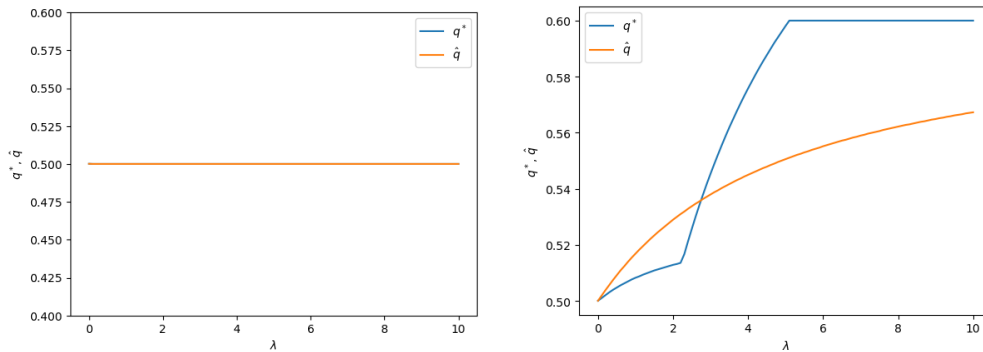
Figure 22: Equilibrium Solutions with Low Cognitive Costs

Return to the Section 3.2.

## B.5 Comparison of SSPE Solution to Socially Optimal Solution

In the figures below, I compare numerically the socially optimal solution,  $\hat{q}$ , against the SSPE solution,  $q^*$ , for a small range of  $\lambda$ s.

Figure 23a shows for homogeneous firms and correct prior beliefs ( $p = \frac{1}{2}$ ). Figure 23b shows for homogeneous firms incorrect prior beliefs, biased towards firm 1 by default ( $p > \frac{1}{2}$ ), and Figure 24a shows the other case of biased towards firm 2 by default ( $p < \frac{1}{2}$ ). The rest of the figures allow the firms to be heterogeneous, and check for the cases where prior beliefs are correct or incorrect. In general, the examples illustrate that SSPE is efficient when firms are homogeneous and workers hold the correct prior beliefs. Any other variations could lead to inefficiency, defined as deviation of  $q^*$  away from  $\hat{q}$ .

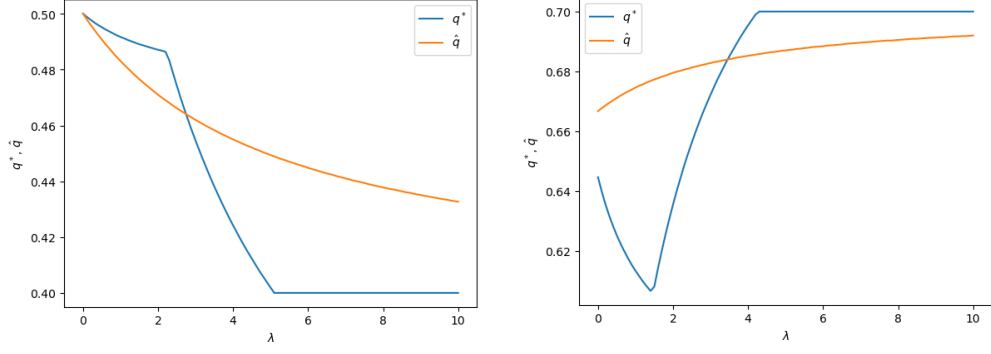


(a) Homogeneous Firms &  $p = \frac{1}{2}$

(b) Homogeneous Firms &  $p > \frac{1}{2}$

Figures show for  $\lambda \in [0, 10]$ ,  $z_1 = 10$ ,  $z_2 = 10$ , (a)  $p = 0.5$ , (b)  $p = 0.6$ .

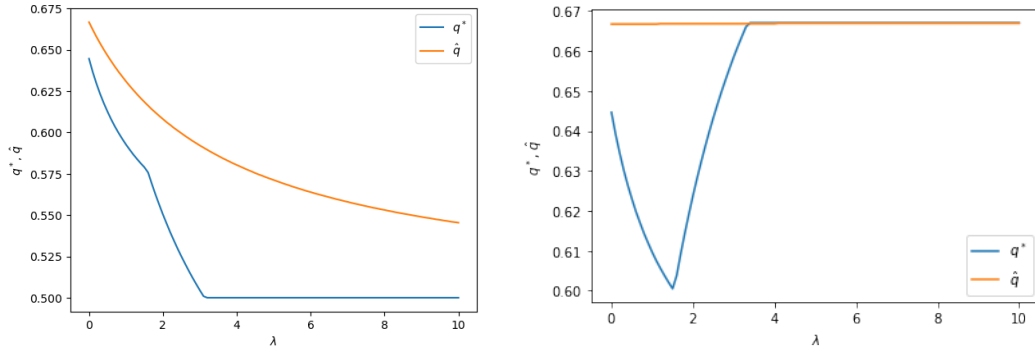
Figure 23: SSPE vs. Socially Optimal Application Probabilities (Homogeneous Firms)



(a) Homogeneous Firms &  $p < 0.5$       (b) Heterogeneous Firms &  $p > \frac{z_1}{z_1+z_2}$

For  $\lambda \in [0, 10]$ , (a)  $z_1 = 10$ ,  $z_2 = 10$  (Homogeneous Firms),  $p = 0.4$ ; and (b)  $z_1 = 10$ ,  $z_2 = 5$  (Heterogeneous Firms),  $p = 0.7$ .

Figure 24: Comparison between Socially Optimal Solution  $\hat{q}$  and SSPE solution  $q^*$ .



(a) Heterogeneous Firms &  $p < \frac{z_1}{z_1+z_2}$       (b) Heterogeneous Firms &  $p = \frac{z_1}{z_1+z_2}$

Figures show for  $\lambda \in [0, 10]$ ,  $z_1 = 20$ ,  $z_2 = 10$ , (a)  $p = 0.5$ , (b)  $p = 0.667$ .

Figure 25: SSPE vs. Socially Optimal Application Probabilities (Heterogeneous Firms)

Return to the Section 4.1.

## C Online Appendix: Extensions

### C.1 Unconstrained Socially Optimal Strategy and ASPE

The FOCs of unconstrained social planner problem (70) would provide the socially optimal solution  $(\hat{q}_1, \hat{q}_1^{-i})$ , where  $\hat{q}_1$  can differ from  $\hat{q}_1^{-i}$ :

$$(1 - q_1^{-i})z_1 - q_1^{-i}z_2 = \lambda_1 \log\left(\frac{1-p_1}{p_1} \frac{q_1}{1-q_1}\right) \quad (153)$$

$$(1 - q_1)z_1 - q_1z_2 = \lambda_2 \log\left(\frac{1-p_1^{-i}}{p_1^{-i}} \frac{q_1^{-i}}{1-q_1^{-i}}\right) \quad (154)$$

The resulting  $(\hat{q}_1, \hat{q}_1^{-i})$  would be,

$$\hat{q}_1 = \frac{1}{1 + \frac{1-p_1}{p_1} \exp\left(-\left(\frac{(1-\hat{q}_1^{-i})z_1 - \hat{q}_1^{-i}z_2}{\lambda_1}\right)\right)} \quad (155)$$

$$\hat{q}_1^{-i} = \frac{1}{1 + \frac{1-p_1^{-i}}{p_1^{-i}} \exp\left(-\left(\frac{(1-\hat{q}_1)z_1 - \hat{q}_1z_2}{\lambda_2}\right)\right)} \quad (156)$$

Within the interval  $(0,1)$ , there is one solution pair where  $\hat{q}_1 = \hat{q}_1^{-i} = \hat{q}$ , which is also the solution for the constrained social planner problem. However, if  $\hat{q}_1 \neq \hat{q}_1^{-i}$ , then  $\hat{q}_1$  and  $\hat{q}_1^{-i}$  are going in opposite direction, and there could exist 2 asymmetric socially optimal strategies where workers apply more to different firms. This may not exist all the time and have to depend on certain conditions that is similar to that of existence of ASPE. Hypothetically, when comparing the asymmetric socially optimal strategies (155), (156), and ASPE strategies (37) and (38), to improve efficiency, workers should behave in a more directed or asymmetric manner by applying with higher probability to different firms.

For instance, suppose workers are homogeneous with  $p = \frac{1}{2}$  and  $\lambda_1 = \lambda_2 = \lambda$ , and firms are homogeneous as well,  $z_1 = z_2 = z$ , such that  $w_1 = w_2 = w^*$ , the equilibrium probability of applying to firm 1 would be,

$$q_1^* = \frac{1}{1 + \exp\left(-\frac{(1-2q^{-i*})w^*}{2\lambda}\right)} \quad (157)$$

and the socially optimal probability of applying to firm 1 would be,

$$\hat{q}_1 = \frac{1}{1 + \exp\left(-\frac{(1-2\hat{q}^{-i})z}{\lambda}\right)} \quad (158)$$

If the choice probabilities are the same and ASPE is efficient, then  $q_1^* = \hat{q}_1 = \tilde{q}_1$ ,  $q_1^{-i*} = \hat{q}_1^{-i} = \tilde{q}_1^{-i}$ , the two equations (157) and (158) combine to form,

$$z = \frac{w^*}{2}$$

This cannot hold since  $z \geq w^*$ . As a result, choice probabilities cannot be the same. Using (157) and (158),

$$\hat{q}_1 - q_1^* = \frac{\exp\left(-\frac{(1-2q^{-i*})w^*}{2\lambda}\right) - \exp\left(-\frac{(1-2\hat{q}^{-i})z}{\lambda}\right)}{(1 + \exp\left(-\frac{(1-2\hat{q}^{-i})z}{\lambda}\right))(1 + \exp\left(-\frac{(1-2q^{-i*})w^*}{2\lambda}\right))} \quad (159)$$

$\hat{q}_1 - q_1^* > 0$  if  $(1 - 2q^{-i*})\frac{w^*}{2} < (1 - 2\hat{q}^{-i})z$ , and  $\hat{q}_1 - q_1^* < 0$  if  $(1 - 2q^{-i*})\frac{w^*}{2} > (1 - 2\hat{q}^{-i})z$ , these always hold if socially optimal strategy is more asymmetric than equilibrium strategy.

The unconstrained social planner problem allows the flexibility of having different  $\hat{q}_j^i$  for different workers. By evaluating the equilibrium outcomes against the  $\hat{q}_1$  and  $\hat{q}_1^{-i}$ -benchmark, this would improve the accuracy of studying the prevailing market efficiency. However, using such benchmark for policy implementation on a larger scale maybe challenging. When there are more workers and firms, it can be very costly to consider the possibility of them having different strategies and evaluate how inefficient each strategy is against the socially optimal benchmarks. There may be multiple equilibria and multiple socially optimal benchmarks. Worker heterogeneity could further complicate the issue. The discussion of unconstrained social planner problem will not be detailed in this paper, but it remain important to acknowledge the plausibility of computer algorithms to identify workers' strategies and compare against different benchmarks on a larger scale.

Return to the Section 4.