Costly Job Search with Inattentive Workers*

Siting Estee Lu†

October 14, 2025

[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. Equilibria 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

*I am indebted to my supervisors, Ed Hopkins and Axel Gottfries, for their invaluable suggestions and continuous guidance. I would also like to express my sincere gratitude to Michael Woodford, Mark Dean, Hassan Afrouzi and others at Cognition and Decision Lab at Columbia University for their comments and advices. I also hope to thank Tatiana Kornienko, Rava Azeredo da Silveira, Kohei Kawamura, Ennio Bilancini, Pawel Gola, Larbi Alaoui, Dilip Soman, audiences at EEA Congress 2025, FUR Conference 2024, Cognitive Foundations of Decision-Making 2024, 33rd Advanced School in Economic Theory, SABE-IAREP Conference 2023, SGPE 2021, and QMUL Economics and Finance Workshop, as well as Scottish Economic Society for its support. Older versions of the paper can be accessed here. All errors are my own.

†School of Economics, The University of Edinburgh. Contact: estee.l828@gmail.com

1 Introduction

Job search behaviour is inherently complex. Despite the abundance of information provided by online job search platforms, a fundamental information bottleneck remains, where workers have limited attention and may not be able to process all available information. This cognitive limitation could shape workers' imperfect search behaviour and lead to mismatch in the labour market. While 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 broad search pattern contrasts with empirical findings that suggest inattentive workers exhibit narrow, biased search behaviours (Belot et al. (2019)). This discrepancy may be attributed to stringent assumptions of the model, limiting its explanatory power. Workers are typically assumed to apply symmetrically due to potential coordination difficulties with asymmetric strategies (Wright et al. (2021)), but such simplification could suppress the strategic diversity seen in reality and may overlook existence of more efficient equilibria. Moreover, like other decision problems, workers' search behaviour can also be influenced by past experience (Erev et al. (2010)), assuming a default to random search neglects the role of prior knowledge on workers' application decisions. These can lead to inaccurate predictions of search efficiency and implications for welfare analysis. The literature therefore lacks a generalized, information-theoretic framework that fully captures workers' imperfect cognition on labour market outcomes, and my work addresses this gap by developing a model that yields richer behavioural predictions and offers clearer guidance for policies aimed at mitigating information frictions.

In this paper, I propose a Generalized Partially Directed Search (GPDS) model featuring two firms and two attention-constrained workers. Wages are not directly observable, but workers can learn about the wages at a precision level of their choice by incurring a cost proportional to the Kullback-Leibler divergence between the default and chosen strategy, following the approach in Wu (2024). My model builds on Wu (2024)'s existing partially directed search model and relaxes two restrictive assumptions: first, that workers must employ symmetric application strategies; and second, that their default search strategy is random.

By allowing for asymmetric application strategies, this work yields a key theoretical result. I prove the existence of multiple equilibria, showing that asymmetric strategies can emerge when information is sufficiently inexpensive relative to expected payoffs. Even when workers are homogeneous in attention costs and face firms of identical productivity, they may endogenously coordinate on applying with higher probability to different firms, generating targeted search beyond what wage differentials alone would predict. The refinement also captures strategic diversity that has been overlooked in standard directed search formulations and carries implications for welfare analysis, as different equilibria can have varying levels of market efficiency.

To further enrich the behavioural dimension, I incorporate informative priors that represent workers' implicit knowledge about the wage distribution, thereby relax the assumption of random default strategies. This allows workers to adopt exogenous, non-random default search strategies prior to information acquisition. This behavioural extension introduces imperfect search behaviour that arises from workers' prior biases, which can exacerbate information frictions and contribute to market mismatch.

Information frictions play a non-trivial role in the matching problem and can substantially affect social welfare. When comparing equilibrium outcomes in presence of multiple equilibria, I found social welfare is higher when workers adopt differentiated application strategies rather than applying symmetrically. Since wages differ across equilibria, there are also variations in monopsony power, revealing a trade-off between firms' and workers' welfare. This result could be interesting for policymakers as some equilibrium may be deemed more "desirable" than others depending on policy objectives, thus serving as basis for further exploration of equilibrium selection measures. I further extend the model to investigate a constrained social planner problem, where the socially optimal strategy can serve as a benchmark to evaluate the efficiency level of equilibrium outcomes. Equilibrium where workers use identical application strategies is found to be generally inefficient, which supports policies to improve efficiency by reducing information frictions.

Finally, I examine a less conventional form of horizontal differentiation in attention costs, arising from heterogeneity in workers' default strategies or cognitive costs. This provides a novel angle in analysing imperfect search driven by differences in cognitive capacity, rather than the traditional focus on disparities in skills or productivity (Shi (2002); Wright et al. (2021)). I found that when attention costs differ across workers, they tend to adopt asymmetric application strategies, providing a behavioural explanation for the empirically observed narrow search patterns where workers assign higher application probabilities to certain firms. Incorporating heterogeneity in default strategies further reveals that prior bias is not necessarily detrimental to market outcomes, instead welfare depends on the structure of biases across the population, and complementary biases can serve as an effective coordination device that improves efficiency.

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 2×2 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.[1] 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 ΔW is the set of Borel probability distributions on W.
- (*Decision Problem*) One chooses an information structure, $F^i(\mathbf{s}|\mathbf{w})$, and a search strategy given signals, $q^i_j(\mathbf{s})$.
- (*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.

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

  λ_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* $\{G^e(\boldsymbol{w})^i, F^e(\boldsymbol{s}|\boldsymbol{w})^i, \{q^e_j(\boldsymbol{s})^i\}_j, G^e(\boldsymbol{w})^{-i}, F^e(\boldsymbol{s}|\boldsymbol{w})^{-i}, \{q^e_j(\boldsymbol{s})^{-i}\}_j, \{w^i_j(z|\boldsymbol{z})\}_j\}$,

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

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

$$\max_{\{F^e(\boldsymbol{s}|\boldsymbol{w})^i, \{q^e_j(\boldsymbol{s})^i\}_j\}} \int_{\boldsymbol{w}}\int_{\boldsymbol{s}} [q^i_1(\boldsymbol{s})(1 - q_1^{-i}(\boldsymbol{s}) + \frac{q_1^{-i}(\boldsymbol{s})}{2})w_1 \\ + q^i_2(\boldsymbol{s})(1 - q_2^{-i}(\boldsymbol{s}) + \frac{q_2^{-i}(\boldsymbol{s})}{2})w_2]F^i(d\boldsymbol{s}|\boldsymbol{w})G^{-i}(d\boldsymbol{w}) - \lambda_i[\mathbb{H}(G^{-i}(\boldsymbol{w})) - \mathbb{E}_{\boldsymbol{s}}\mathbb{H}(F^i_{\boldsymbol{w}|\boldsymbol{s}})] \tag{2}$$

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

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

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))

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

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 λ_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.

$$\max_{q_1^i, q_2^i \in [0,1]} q_1^i(1 - q_1^{-i} + \frac{q_1^{-i}}{2})w_1 + q_2^i(1 - q_2^{-i} + \frac{q_2^{-i}}{2})w_2 - \lambda_i(q_1^i \log \frac{q_1^i}{p_1^i} + q_2^i \log \frac{q_2^i}{p_2^i}) \tag{4}$$

$$\text{s.t. } q_1^i + q_2^i = 1, \quad q_1^{-i} + q_2^{-i} = 1, w_1 \geq 0, w_2 \geq 0$$

(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:

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

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

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

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

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:

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

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 $[1-(1-q_j^i)(1-q_j^{-i})]$. 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 increases 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.

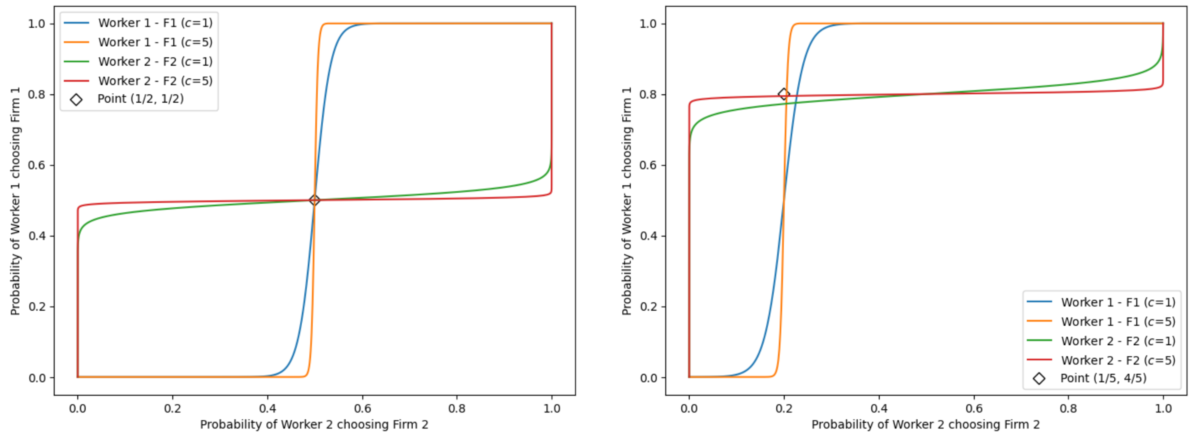
2.2.2 Relation to Quantal Response Functions

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 in literature such as McKelvey and Palfrey (1995) and Goeree et al. (2016). Modelling workers' search behaviours using QRFs instead of RIFs implies extending the rational expectation framework by allowing for errors in workers' choices rather than them having attention constraint. Their strategies would be captured by smoothed-out response functions:

$$q^i_{F1} = \frac{\exp(c\pi_{F1}(q^{-i}_{F1}))}{\exp(c\pi_{F1}(q^{-i}_{F1})) + \exp(c\pi_{F2}(q^{-i}_{F1}))} = \frac{1}{1 + \exp(-c(\pi_{F1}(q^{-i}_{F1}) - \pi_{F2}(q^{-i}_{F1})))} \tag{10}$$

$$q^{-i}_{F1} = \frac{\exp(c\pi_{F1}(q^{i}_{F1}))}{\exp(c\pi_{F1}(q^{i}_{F1})) + \exp(c\pi_{F2}(q^{i}_{F1}))} = \frac{1}{1 + \exp(-c(\pi_{F1}(q^{i}_{F1}) - \pi_{F2}(q^{i}_{F1})))} \tag{11}$$

where q^i_{F1} and q^{-i}_{F1} are the probabilities of applying to firm 1, π 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 $c \to \infty$ suggests complete rationality. It can also be interpreted as sensitivity to expected payoff. 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, and the equilibrium solutions can be more diffused and diverge further away from standard NE under complete information.



(a) Homogeneous Firms ($w_1^* = w_2^* = 60$).

(b) Heterogeneous Firms ($w_1^* = 60, w_2^* = 40$).

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

There are mathematical similarities between the QRFs (10), (11) and the RIFs (7), (8). For instance, cognitive cost in RIFs is the inverse of precision in QRFs (i.e. $c = \frac{1}{\lambda_i}$), which indicates that higher unit cost of processing information makes it harder to learn precise wages or workers experience lower sensitivity to expected payoff. In RIFs, workers' sensitivity to wages is additionally weighted by skewness in default.[2] These parallels have not been explicitly established, but Goeree et al. (2020) has highlighted the potential link.

Despite the equivalence, the source of noise modelled by the two concepts and their implications are slightly different. Understanding their differences is vital for policy informativeness. In QRE, workers observe the wages precisely and better respond to beliefs about randomness in opponent's strategies and the expected payoff. In the Rationally Inattentive Equilibrium (RIE) characterized by RIFs, wages are not directly observed, workers can learn about the wages by acquiring a costly signal. They may choose to remain partially informed, their beliefs about the wages and the strategies of their opponent are shaped by the amount of information they have acquired and processed, so they are best responding given a noisy environment. RIE not only rides on top of QRE's advantage as a statistical model, where one can infer the learning precision by fitting the model to empirical data, it can also be perceived as an econometric model, where cognitive costs can be approximated by IQ scores, and the relationship between IQ and search decisions can be evaluated. Furthermore, since the source of randomness in workers' strategies in the QRE is attributed to error, it could be difficult to decipher, and thus its policy impact is lacking. On the other hand, the backdrop to RIE is one of information-theoretic, the source of noise comes from endogenous decision of information acquisition and processing, which leaves room for policy tools to improve attention allocation.

[2]Figure 1 can also be used to illustrate the case of workers with homogeneous attention costs with random search default strategies, and cognitive costs of either 1 or $\frac{1}{5}$.

2.2.3 Deterministic Search Behaviour

Applying to Both Firms. There is a misconception about applications being almost costless and one should apply to all firms deterministically, with probability 1. On the very contrary, both tangible and intangible cost associated with each application decision prevent this and instead propel one to make deliberate choice between the firms. 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 judgments tests, psychometric tests, HireVue assessments, Leetcode training); and social capital costs (i.e. job referees that constitute as favour asking and 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. Although they do not affect which firm a worker choose to apply for, but given limited resources, one will be making a selection instead of apply-to-all.

Finite Cost for Precise Information. In this model, it is possible for one to incur finite attention cost in acquiring precise information, would their search behaviour be deterministic as a result? 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. The same assumption can be imposed in this model such that as long as information cost is positive, one should optimally decide to tolerate some imprecision and behave probabilistically. However, even without this assumption, equilibrium strategy could comprise of probabilistic choices. Given bounded defaults, $p_j^i \in (0,1)$, if the expected payoff ($\mathbb{E}$) is 0 or that λ_i and λ_{-i} are sufficiently high, then there are no incentive to acquire information, equilibrium strategy would be the same as default, which is inherently probabilistic.

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

For non-zero expected payoff, where workers 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.

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

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.

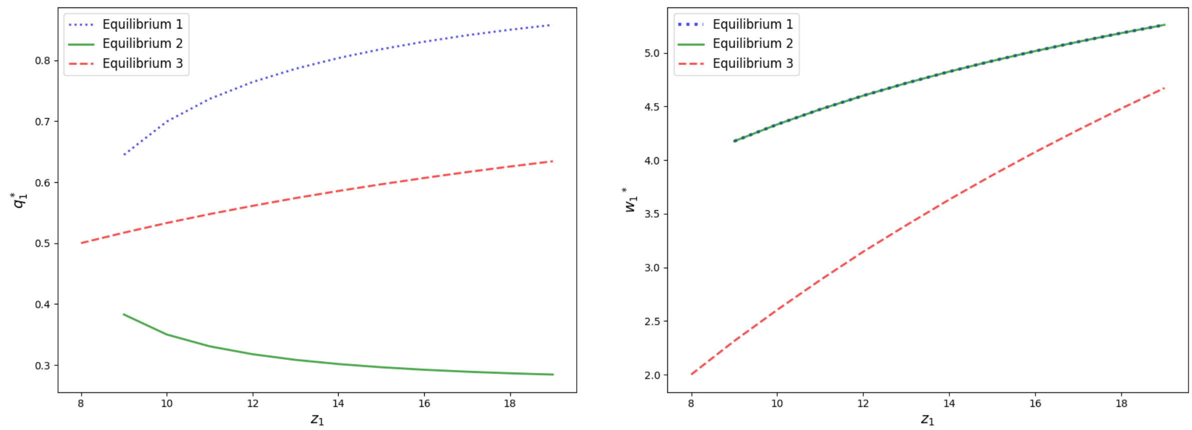
Deterministic Default Strategies. Workers may posses strong initial bias and apply with conviction to one of the firms by default. They can choose to acquire information and deviate away from this deterministic choice, but in the case of high information costs, workers would not acquire any information, deterministic default would thus translate into deterministic equilibrium behaviour. This will contribute to severe overcrowding problem if both workers apply to the same firm. It is unlikely for firms to have such a strong goodwill, therefore, this model assumes bounded defaults (i.e. $p_j^i \in (0,1)$), which rules out this possibility.

2.3 Subgame Perfect Equilibrium

I have now specified the GPDS framework and explained the second-stage game, where workers take wages as given. In the full game, workers' application strategies 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,* λ_i*, and wages,* w_1, w_2*.*

1. $q^i_j(w_1, w_2; p^i_j, \lambda_i)$ *maximizes worker i's payoff, take as given* w_1, w_2 *and* $q^{-i}_j(w_1, w_2; p^{-i}_j, \lambda_{-i})$.
2. w^e_j *maximizes firm j's payoff, take as given* $q^i_j(w_1, w_2; p^i_j, \lambda_i), q^{-i}_j(w_1, w_2; p^{-i}_j, \lambda_{-i})$ *and* w^e_{-j}.



(a) Worker 1's Equilibrium Probabilities (b) Firm 1's Equilibrium Wages

Figures show bifurcation graphs of q_1^* (referring to worker 1) and w_1^* against z_1; $z_1 \in [8, 20]$, $z_2 = 8$, $p^i_1 = p^{-i}_1 = \frac{1}{2}$, $\lambda = 1$.

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

For a given set of exogenous λ, z_1, z_2, p^i_1, p^{-i}_1, Figure 2 shows three possible sets of equilibrium solutions.[3] I will first analyze workers' search behaviour, followed by firms' wage setting.

2.3.1 Workers' Equilibrium Search Behaviours

I define the types of SPE based on the symmetric or asymmetric nature of workers' application strategies. In Figure 3, workers' equilibrium search behaviour constitute of two possible types, one where both workers use the same application strategy and apply more randomly. As firm 1 becomes more productive, such that z_1 grows against z_2, workers would apply with higher probability to firm 1. This equilibrium is defined as the symmetric subgame perfect equilibrium (SSPE) and was addressed in Wu (2024). The other equilibrium type consists of workers using different application strategies and apply with higher probability to different firms. Growing firm heterogeneity is coupled with growing asymmetry in workers' equilibrium strategies, shown by them applying with even higher probability to different firms. This is defined as the asymmetric subgame perfect equilibrium (ASPE), which was not accounted for in the existing literature as it is assumed to be more difficult for workers to coordinate on (Wright et al. (2021)). I will first discuss the SSPE.

Definition 2.3 (Symmetric Subgame Perfect Equilibrium (SSPE))**.** *Symmetric subgame perfect equilibrium is defined to be a case of SPE that* $q^i_j(w_1, w_2; p^i_j, \lambda_i) = q^{-i}_j(w_1, w_2; p^{-i}_j)$*. 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.*[4]

Symmetric Subgame Perfect Equilibrium (SSPE). I retain the original assumptions from Wu (2024) where workers are homogeneous in attention costs, $\hat{c}_1 = \hat{c}_2$, such that $\lambda_i = \lambda_{-i} = \lambda$, $p_1 = p^{-i}_1 = p$; and they adopt symmetric application strategies, $q_1 = q^{-i}_1 = q$. However, I relax

[3]Henceforth, for notation simplicity, I omit the superscript i for worker i's choice probabilities in the main text, except in definitions.

[4]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, do not go into the equilibrium definition, which emphasizes the focus on the symmetric or asymmetric nature of workers' strategies.

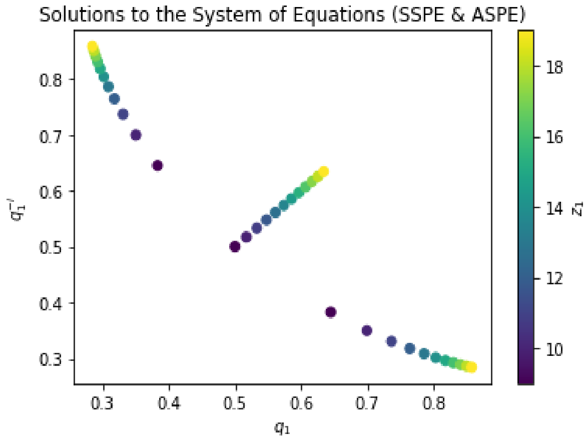


Figure shows workers' equilibrium solutions in both SSPE and ASPE when $\lambda = 1$, $z_1 \in [8, 20]$, $z_2 = 8$, $p_1 = p_1^{-i} = \frac{1}{2}$. The colour spectrum displays a range of z_1 while fixing z_2 to demonstrate the impact of heterogeneous productivity.

Figure 3: Workers' Application Probability

the assumption on random search default strategies, allowing them to vary but stay bounded (i.e. $p \in (0, 1)$). From workers' search problem (4), I derive the first-order condition (FOC):

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

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

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

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

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

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

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

SSPE System of Equations. The SSPE consists of a system of simultaneous equations, (12), (15), (16) that 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 the firms.

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

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

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

The more expensive the unit cost of processing information, the less likely workers will acquire and process them, and thus are less sensitive to wage changes. As workers' strategies become less dependent on wages, firms are more likely to offer low wages. Therefore, higher λ, higher wage markdowns. 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. I propose instead that workers may have biased initial beliefs about wage distribution, so 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 (*$p_j \neq \frac{1}{2}$*) is denoted as belief distortion resulting from informative priors.*

Definition 2.5 (Correct Prior Beliefs)**.** *Prior beliefs are defined to be correct if workers' default choice probabilities match the firms' productivity share (i.e.* $p_j = \frac{z_j}{z_j + z_{-j}}$*), such that the more productive firm is rightfully allocated higher selection probability and is more likely to produce.*

The potential for belief distortion in the GPDS model indicates that if workers receive a signal that is in conflict with their prior beliefs, the cost incurred to deviate away from the corresponding default strategy could have an additional interpretation of psychological cost arising from "cognitive dissonance", where one's belief clashes with new information (Festinger (1962)). Workers are not only updating their beliefs with the acquired information, but also rectifying their past biases. Therefore, changes in default strategies could provide insights on how innate beliefs may bias search behaviours and affect wage-setting.

Proposition 1 (Belief Distortion on Equilibrium Application Probability)**.** *For homogeneous workers applying symmetrically, biased defaults lead to biased search. Equilibrium application probability to firm 1 is strictly increasing in default choice probability of firm 1,* $\frac{dq^*}{dp} > 0$*.*

Proof. Based on wage-setting equations, (17) and (18), I find the indirect impact of default application probability on equilibrium wages:

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

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

From workers' FOC (12),

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

Combining (22),(20) and (21), I show the effect of changing default probabilities on equilibrium probabilities:

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

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

As for the equilibrium wages, since $\frac{dq^*}{dp} > 0$, $\frac{dw_1^*}{dp} < 0$ from (20) and $\frac{dw_2^*}{dp} > 0$ from (21). Firms would set a lower wage than their rival if there is higher probability of meeting a worker.

Proposition 2 (Cognitive Costs on Equilibrium Application Probability)**.** *Workers are reliant on their default search strategies, they stick to it in the limit of infinite cognitive costs. Formally:* $\lim_{\lambda \to \infty} q^* = p$ *and* $\lim_{\lambda \to \infty} \frac{dq^*}{dp} = 1$

Proof. From workers' equilibrium choice (19), define the corresponding expected payoff to be $\mathbb{E}$,

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

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

Workers are sensitive to $\mathbb{E}$ in absence of cognitive costs, they deviate away from defaults for $\mathbb{E} \neq 0$. For extremely high cognitive costs, workers are not sensitive to $\mathbb{E}$. Equilibrium strategy converges to default strategy in the limit. Based on (22) and (23), considering that $\lambda \to \infty$, $w_1^* \to 0$ and $w_2^* \to 0$ from (17) and (18), and $\frac{dw_1^*}{dp} \to 0$, $\frac{dw_2^*}{dp} \to 0$,

$$\lim_{\lambda \to 0^+} \frac{dq^*}{dp} = 0, \lim_{\lambda \to \infty} \frac{dq^*}{dp} = 1 \tag{26}$$

Equilibrium application probabilities are more reliant on defaults as cognitive costs increase. □

By allowing workers to have diverse priors, I demonstrate that innate biases could be carried forward into the equilibrium strategies when information is costly. This provides new insights on labour market mismatch. If the default search strategies is biased, imperfect search could result from a combination of cognitive limitation and bias. One of the firms may receive higher application probability than intended given the market condition, thus leading to higher chance of overcrowding at a single firm, and higher likelihood of mismatch. (Appendix B.2)

Existence of Multiple Equilibria. One of the key contributions in this paper is that I explore the possibility of multiple equilibria. Workers are often assumed in search models to behave symmetrically, presence of equilibrium multiplicity, particularly the equilibrium with asymmetric strategies, has largely been discarded to prevent potential coordination problems. However, there are no definite indication that symmetric equilibrium is easier to achieve than asymmetric ones. Furthermore, considering for the existence of multiple equilibrium allows for evaluation of social welfare under each scenario, which opens up the possibility of equilibrium selection if efficiency differs.

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). I assume workers remain homogeneous in attention costs. From workers' search problem (4), workers' FOCs:

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

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

From firms' maximization problem (9), firms' FOCs:

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

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

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

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

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

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

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

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

$$\frac{1-(1-q_1)(1-q_1{}^{-i})}{z_1-w_1} = \frac{\begin{array}{l}(w_1+w_2)(1-q_1)q_1(1-q_1^{-i})q_1^{-i}[(2-q_1^{-i})(1-q_1)+(2-q_1)(1-q_1^{-i})]\\ -2\lambda(1-q_1^{-i})(1-q_1)[q_1^{-i}(2-q_1)+q_1(2-q_1^{-i})]\end{array}}{(w_1+w_2)^2 q_1(1-q_1)q_1^{-i}(1-q_1^{-i}) - 4\lambda^2} \tag{35}$$

$$\frac{1-q_1q_1^{-i}}{z_2-w_2} = -\{\frac{\begin{array}{l}(w_1+w_2)(1-q_1^{-i})q_1^{-i}(1-q_1)q_1[(2-q_1)q_1^{-i}+q_1(2-q_1^{-i})]\\ -2\lambda q_1 q_1^{-i}[(1-q_1)(2-q_1^{-i})+(1-q_1^{-i})(2-q_1)]\end{array}}{(w_1+w_2)^2 q_1(1-q_1)q_1^{-i}(1-q_1^{-i}) - 4\lambda^2}\} \tag{36}$$

<u>*ASPE System of Equations.*</u> The ASPE consists of a system of simultaneous equations that can be solved for $(q_1^*, q_1^{-i^*}, w_1^*, w_2^*)$.

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

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

$$w_1^* = \max\{z_1 - \lambda \frac{(q_1^*+q_1^{-i^*}-q_1^*q_1^{-i^*})[1-(\frac{1}{q_1^*-q_1^{-i^*}}\log(\frac{q_1^*}{1-q_1^*}\frac{1-q_1^{-i^*}}{q_1^{-i^*}}))^2 q_1^* q_1^{-i^*}(1-q_1^*)(1-q_1^{-i^*})]}{\begin{array}{l}(1-q_1^{-i^*})(1-q_1^*)(q_1^*+q_1^{-i^*}-q_1^*q_1^{-i^*}) - \frac{1}{2(q_1^*-q_1^{-i^*})}\log(\frac{q_1^*}{1-q_1^*}\frac{1-q_1^{-i^*}}{q_1^{-i^*}})\\ q_1^*q_1^{-i^*}(1-q_1^*)(1-q_1^{-i^*})(2q_1^*q_1^{-i^*}-3q_1^*-3q_1^{-i^*}+4)\end{array}}, 0\} \tag{39}$$

$$w_2^* = \max\{z_2 - \lambda \frac{(1-q_1^*q_1^{-i^*})[1-q_1^*q_1^{-i^*}(1-q_1^*)(1-q_1^{-i^*})(\frac{1}{q_1^*-q_1^{-i^*}}\log(\frac{q_1^*}{1-q_1^*}\frac{1-q_1^{-i^*}}{q_1^{-i^*}}))^2]}{\begin{array}{l}q_1^*q_1^{-i^*}(1-q_1^*q_1^{-i^*}) - q_1^*q_1^{-i^*}(1-q_1^*)(1-q_1^{-i^*})(q_1^*+q_1^{-i^*}+2q_1^*q_1^{-i^*})\\ \frac{1}{2(q_1^*-q_1^{-i^*})}\log(\frac{q_1^*}{1-q_1^*}\frac{1-q_1^{-i^*}}{q_1^{-i^*}})\end{array}}, 0\} \tag{40}$$

While small cognitive costs unambiguously have positive impact on wages, it is less intuitive what would be the effect of large cognitive costs on equilibrium wages in ASPE. When workers apply more to different firms, the second term of (39) and (40) becomes large and pushes the wages down, firms in anticipation of meeting at least one worker with high probability would reduce wages, particularly when cognitive costs are high and workers are less sensitive to wage changes. Same as the symmetric case, increasing cognitive costs would lower equilibrium wages. However, as wages are lowered, workers may lack incentive to acquire any information, thus they would not adjust away from defaults if expected payoff are too low, and ASPE will cease to exist. Unlike SSPE, whose existence and uniqueness is not dependent on any specific values of wages. The existence of asymmetric strategies in the subgame is conditional on the wage announcements. Therefore, it is necessary for me to characterize the condition for asymmetric strategies to prevail in the subgame equilibrium, and if they exist for a set of wages, then there will be mapping to the productivity, and ASPE would exist as a result.

In the subgame, if workers adopt asymmetric strategies, (q_1, q_1^{-i}), where $q_1 \neq q_1^{-i}$, their application probabilities would go in opposite direction. 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 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, I show the derivative with respect to opponent's choice probability of firm 1.

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

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

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 (numerical examples in 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 sufficiently small λ relative to expected payoff, which is determined by the wage levels (for $w_1 = w_2 = w$), or wage differences (for $w_1 \neq w_2$, where $2w_1 > w_2 > \frac{w_1}{2}$ is fulfilled).*

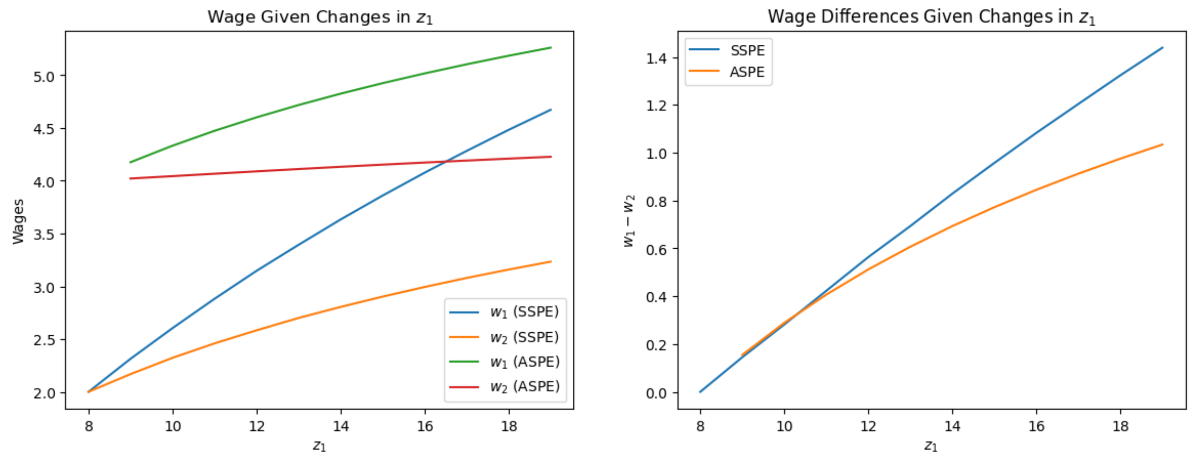
Proof. See Appendix A.3. □

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 with higher probability to different firms, but cognitive costs have to be relatively low as compared to the expected payoff. The threshold of λ is an implicit expression of w_1, w_2, q^*, p. In the full game, the wage condition in Lemma 2.2 implies λ in turn needs to be small relative to productivities or that productivity difference has to be higher than the difference in markdowns.

Strategic Stability. More generally, to analyze all the possible equilibrium. I consider the case where workers commit to an application strategy, firms in view of that would decrease their wages to maximize profits, and workers would be made worse off. Such scenario is unstable as workers could deter this wage reduction by credibly threaten to acquire some information, thus preventing firms from choosing off-the-equilibrium-path wages. This credible threat of information acquisition and processing help to support the stability of equilibrium on the path.

Impact of Belief Distortion. 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" of workers towards firm 1 in the equilibrium (see Appendix A.6). Therefore, belief distortion can still contribute to biased search and high likelihood of overcrowding.

2.3.2 Firms' Equilibrium Wage Setting



(a) Equilibrium Wages (b) Wage Differentials

Figures show wages in SSPE and ASPE when $\lambda = 1$, $z_1 \in [8, 20]$, $z_2 = 8$, $p_1 = p_1^{-i} = \frac{1}{2}$.

Figure 4: Effects of Varying Productivity Differences on Equilibrium Wages

Moving onto the firms' side, Figure 4a demonstrates equilibrium wages in SSPE and ASPE. There are two counter-intuitive points: (1) Previously in Figure 3, growing firm heterogeneity is coupled with growing asymmetry in workers' equilibrium strategies for ASPE, where workers display greater "loyalty" towards different firms. Despite firm 2 being less productive and have increasingly lower ability to match firm 1's wage offer, shown by the wage gap as z_1 raises against z_2, it can still receive higher application 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 firms have incentive to set high wages. However, this is not observed in Figure 4a when firms have similar productivities and wages are higher in ASPE than SSPE.

For the first point, since increasing productivity differences has a positive, linear impact on wage differences for SSPE, shown in Figure 4b, the benefit of applying with higher probability to firm 1 consistently outweighs gain from higher matching probability if one were to direct their search more to firm 2. However, in ASPE, wage differentials are increasing at a decreasing rate with growing productivity disparity. The benefit of applying for a job with a slightly lower wage but higher chances of being matched could outweigh that of heavily competing for a job with slightly higher wage, as a result, contributing to increasing asymmetry in workers' strategies. Nonetheless, higher firm 1 productivity can still attract both workers to some extent when they are applying asymmetrically, thus the meeting rate faced by the firms is asymmetric.

The most intriguing aspect is the wage levels. Based on the wage determination equations (17) and (39), I denote the equilibrium wage for firm 1 under SSPE and ASPE as w_1^{SSPE} and w_1^{ASPE}

respectively. For $w_1^{\mathrm{ASPE}} \geq w_1^{\mathrm{SSPE}}$:

$$\frac{z_1+z_2}{\lambda} \geq \frac{(q_1+q_1^{-i}-q_1q_1^{-i})[(\frac{1}{q_1-q_1^{-i}}\log(\frac{q_1}{1-q_1}\frac{1-q_1^{-i}}{q_1^{-i}}))^2 q_1q_1^{-i}(1-q_1)(1-q_1^{-i})-1]}{\begin{array}{l} q^2[(\frac{1}{2(q_1-q_1^{-i})}\log(\frac{q_1}{1-q_1}\frac{1-q_1^{-i}}{q_1^{-i}}))q_1q_1^{-i}(1-q_1)(1-q_1^{-i}) \\ (2q_1q_1^{-i}-3q_1-3q_1^{-i}+4)-(1-q_1^{-i})(1-q_1)(q_1+q_1^{-i}-q_1q_1^{-i})]\end{array}} - \frac{2}{q(1-q)} \tag{43}$$

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

Lemma 2.3 (Wages given Feasible Asymmetric Subgame Strategies)**.** *For a feasible set of (q_1^*, q_1^{-i*}), where wages satisfy the necessary condition for existence of asymmetric workers' strategies in the second-stage game, as in Lemma 2.2:*

i. For extreme asymmetric strategies, $\lim_{q_1^*, q_1^{-i*}} \in (1,0),(0,1)$, $w_j^{ASPE} \geq w_j^{SSPE}$ breaks down in the limit.

ii. For close to symmetric strategies, $\lim_{q_1^*, q_1^{-i*}} \in (q^{*+}, q^{*-}),(q^{*-}, q^{*+})$, $w_j^{ASPE} \geq w_j^{SSPE}$ holds if $\frac{z_1+z_2}{\lambda} \geq \frac{1-2q^*+2q^{*2}}{q^{*2}(1-q^*)^2}$, where q^* depends on λ and p, and indirectly on z_1 and z_2.

iii. As q_1^* and q_1^{-i*} become more symmetric, w_j^{ASPE} increases and could exceed w_j^{SSPE}.

Proof. See Appendix A.4. □

Proposition 3 (Lower Monopsony Power in ASPE)**.** *Consider a case of homogeneous firms, $z_1 = z_2 = z$, $w_1^* = w_2^* = w^*$, and workers have random search as default strategy, $p = 0.5$.*

(1) Symmetric case: $q^* = 0.5$, $w^{SSPE} = \frac{z}{2} - 2\lambda$. As $\lambda \to 0$, $w^{SSPE} \to \frac{z}{2}$.

(2) Asymmetric case: each worker applying to different firm in similar balance of probability, $q_1^* = 1-\epsilon$ and $q_1^{-i*} = \epsilon$, where $\epsilon \in (0,1)$.

Under Lemma 2.2, ASPE cease to exist as $\lambda \to 0$ and $\lambda \to \infty$. For λ close to 0, w^{ASPE} can remain small and close to 0; and as λ increases, $\epsilon \to 0.5$, $w^{ASPE} \to z - 4\lambda$. $w^{ASPE} \geq w^{SSPE}$ if $\frac{z}{\lambda} \geq 4$, implying productivity (z) has to be at least 4 times higher than cognitive costs (λ) in the limit, consistent with Lemma 2.3.

As λ increases, there exist a point where $w^{ASPE} \geq w^{SSPE}$, the threshold is:

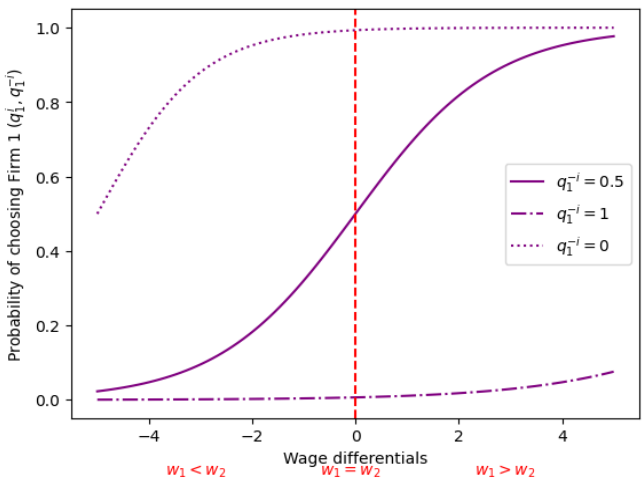
$$\lambda \geq \frac{z}{2}\frac{\epsilon(1-\epsilon)}{1+2\epsilon(1-\epsilon)} \tag{44}$$

where ϵ is an implicit solution to the function:

$$-\frac{2\lambda}{1-2\epsilon}\log(\frac{\epsilon}{1-\epsilon}) = z - \lambda\frac{(1-\epsilon+\epsilon^2)[1-(\frac{1}{1-2\epsilon}\log(\frac{1-\epsilon}{\epsilon}\frac{1-\epsilon}{\epsilon}))^2(1-\epsilon)^2\epsilon^2]}{\begin{array}{l}(1-\epsilon)\epsilon(1-\epsilon+\epsilon^2) - \frac{1}{2(1-2\epsilon)}\log(\frac{1-\epsilon}{\epsilon}\frac{1-\epsilon}{\epsilon}) \\ (1-\epsilon)^2\epsilon^2(1+2\epsilon-2\epsilon^2)\end{array}} \tag{45}$$

Conditional on their existence, ASPE could yield higher equilibrium wages and lower monopsony power compared to SSPE for some intermediate values of λ.

Proof. See Appendix A.5. □



$\lambda = 1$, $p = 0.5$, $w_1 \in [5, 15]$, $w_2 = 10$, opponent's choice is fixed at 0, 0.5 and 1

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

Labour Supply and Wage Sensitivity. Figure 5 shows an increase in workers' application probability to firm 1 as w_1 increases against w_2, which also represents increase in labour supply with growing wages. The graphs correspond to 3 cases where opponent's application rate to firm 1 is set at 0, 0.5 and 1 respectively. While higher wage leads to higher application rate, the speed of change differs when opponents' behaviour varies, which can be attributed to difference in potential overcrowding. For instance, as worker 2 applies more to firm 1, the increase in worker 1's application rate to firm 1 is slower as w_1 grows due to higher risk of overcrowding.

To understand why wages are sometimes higher in ASPE than SSPE, I focus 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 low wage sensitivity region to high wage sensitivity region as w_1 increases. At the margin of $w_1 = w_2$, a small change in w_1 would have a large impact on q_1, thus firms have strong incentive to outbid each other, leading to intense wage competition. In **(2)**, worker 1 starts from high wage sensitivity region, but as w_1 increases, q_1 rapidly converge to 1 and reach the low wage sensitivity region. Any further increase in w_1 generates no additional benefits, thus leading to less intensive wage competition at the point where wages equalize. Considering wage adjustments towards the equilibrium, more intense competition may imply higher equilibrium wages, but this depends on the initial wages set by the firms. From firm 1's perspective, it faces 50% chance of meeting worker 2 in **(1)** and 0% chance in **(2)**. Therefore, it has incentive to set a high initial wage to secure worker 1 in **(2)**. Although reducing w_1 can improve profits at the margin of $w_1 = w_2$, there is risk of firm 1 losing both workers at small cost savings if firm 2 chooses 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 wage, worker 2 would also be attracted to firm 1 and firm 2 risk losing both workers, thus it is motivated to match the wage. The risk faced by firms varies by workers' application behaviour. When workers apply asymmetrically, the risk is higher because failing to attract their "expected" worker, firms can be left with no worker at all; and when workers apply symmetrically, the risk is lower as there is lesser chance of firms being left worker-less even if they post a low initial wage and workers resort to default. Therefore, the competition is in a slightly different manner with more aggressive initial wage setting rather than contesting at the margin, which may result in higher overall wages.

This pattern also depends on the ability of firms to compete. If firm 1 becomes more productive than firm 2, firm 2 will experience less ability to match 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 when workers apply asymmetrically than symmetrically as firms' heterogeneity grows.

Wage sensitivity is also affected by cognitive costs, λ (see Appendix, Figure 20):

From (13),(31), (32), and for a given set of (w_1, w_2), as $\lambda \to 0^+$,

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

Conditional on the existence of ASE and knowing q_1 and q_1^{-i} go in opposing direction, workers would have different sensitivity to wages in ASE.

As information becomes more expensive, in SSE, (13) and (14), $\frac{dq}{dw_1} > 0$, $\frac{dq}{dw_2} < 0$. For $\lambda \to \infty$, $\frac{dq}{dw_1} \to 0$, $\frac{dq}{dw_2} \to 0$. 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 $\lambda \to 0^+$; and as $\lambda \to \infty$, $\frac{dq_1}{dw_1} \to 0$, $\frac{dq_1^{-i}}{dw_1} \to 0$, $\frac{dq_1^*}{dw_2} \to 0$, $\frac{dq_1^{-i}}{dw_2} \to 0$.

Labour supply naturally become less responsive to wage changes. ASE may also cease to exist.

Relating to Section 2.2.1. In the second-stage game, when workers face complete wage information, they are only affected by wage changes when adopting symmetric application strategies, pure strategies are independent of wages. With inattentive workers, close to pure strategies can still prevail in the equilibrium, but these are influenced by wages. Particularly, workers would earn a wage of 0 when using pure strategies given complete information as firms meet at least one worker for certain, but inattentive worker would not acquire any information if wages are 0, firms would set positive wages to motivate workers to apply with higher probability to them, which benefits the workers. This shows that a small perturbation in the system could lead to more wage competition regardless of workers' behaving symmetrically or asymmetrically. Furthermore, the difference in equilibrium wages under ASPE and SSPE indicates variation in monopsony power across equilibrium types, which implies the potential implications from equilibrium selection.

2.3.3 Social Welfare Comparison

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

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

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

$$\mathbb{W}'(q^*) = 2(1-q^*)z_1 - 2q^* z_2 - 2\lambda \log(\frac{1-p}{p}\frac{q^*}{1-q^*}) \tag{47}$$

$$\mathbb{W}''(q^*) = -2(z_1+z_2) - 2\lambda(\frac{1}{q^*(1-q^*)}) < 0 \tag{48}$$

$\mathbb{W}'(q^*)$ 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, λ and p.

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

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

Social welfare is maximized when equilibrium strategy, q^*, matches the true underlying productivity distribution, $\frac{z_1}{z_1+z_2}$. For $\lambda \to 0^+$, local maximum welfare is achieved at $q^* = \frac{z_1}{z_1+z_2}$. Default strategy does not play a role when workers are fully attentive. As $\lambda \to \infty$, workers are completely inattentive, q^* converges to p, social welfare is only maximized if workers have the correct default strategy. Otherwise, there will be persistent deviation away from maximum welfare, suggesting higher information cost is detrimental. For intermediary inattentiveness, q^* varies from p. From (47), $\mathbb{W}'(q^*) > 0$ 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 examples where application to firm 1 is lower than needed, workers should be more targeted in order to improve welfare.

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

$$\frac{d\mathbb{W}(q^*)}{dp} = \frac{dq^*}{dp}[2(1-q^*)z_1 - 2q^*z_2 - 2\lambda \log(\frac{q^*}{p}) - 2\lambda \log(\frac{1-q^*}{1-p})] - 2\lambda(\frac{1-q^*}{1-p} - \frac{q^*}{p}) \tag{49}$$

$\frac{d\mathbb{W}(q^*)}{dp} = 0$ when information is costless, as well as when information is infinitely expensive but workers have correct default. For intermediary $\lambda \in (0, \infty)$, from (23), $\frac{dq^*}{dp} > 0$, $\frac{d\mathbb{W}(q^*)}{dp}$ is likely to deviate from 0. This indicates that biased defaults could adversely affect welfare, suggesting additional loss when workers attempt to "correct" their initial bias.

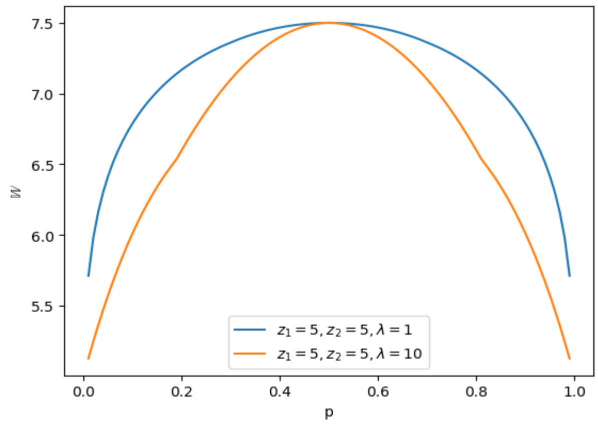


Figure 6: Changing Default Application Probability to Firm 1 on Social Welfare

Figure 6 shows the impact of varying default choice probability on social welfare. In this example, firms are assumed to be homogeneous, and social welfare can be maximized at $p = 0.5$. Lower cognitive costs is generally associated with higher social welfare, and belief distortion could lead to a greater dip in welfare.

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

$$\begin{aligned} \mathbb{W}(q_1^*, q_1^{-i*}) &= (1-(1-q_1^*)(1-q_1^{-i*}))z_1 + (1-q_1^*q_1^{-i*})z_2 \\ &\quad - \lambda[q_1^* \log(\frac{q_1^*}{p}) + (1-q_1^*)\log(\frac{1-q_1^*}{1-p}) + q_1^{-i*}\log(\frac{q_1^{-i*}}{p}) + (1-q_1^{-i*})\log(\frac{1-q_1^{-i*}}{1-p})] \end{aligned} \tag{50}$$

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

Proof. Shown in Appendix A.7.

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.

$$
\begin{aligned}
&((1-q^*)^2 - (1-q_1^*)(1-q_1^{-i^*}))z_1 + (q^{*2} - q_1^* q_1^{-i^*})z_2 \\
&\qquad > \lambda[\sum_{j=1}^{2} q_j^* \log \frac{q_j^*}{p_j} + \sum_{j=1}^{2} q_j^{-i^*} \log \frac{q_j^{-i^*}}{p_j} - 2q^* \log(\frac{q^*}{p}) - 2(1-q^*) \log(\frac{1-q^*}{1-p})] \quad (51)
\end{aligned}
$$

In the proof, I first demonstrate that welfare gain is higher under ASPE than SSPE (i.e. LHS of (51)> 0)). I then study the relative attention costs (i.e. RHS of (51)). More extreme asymmetric strategies (e.g. $q_1^* \to 1, q_1^{-i^*} \to 0$) tend to generate higher attention cost than symmetric strategies, particularly when workers are biased. However, by Lemma 2.2, asymmetric strategies emerge only when λ is relatively low, thus comparing the λ requirement for (51) to hold and λ condition formulated based on workers' FOCs, (37) and (38), I verify that existence condition is stricter than (51). As a result, social welfare is higher under ASPE than SSPE. $\square$

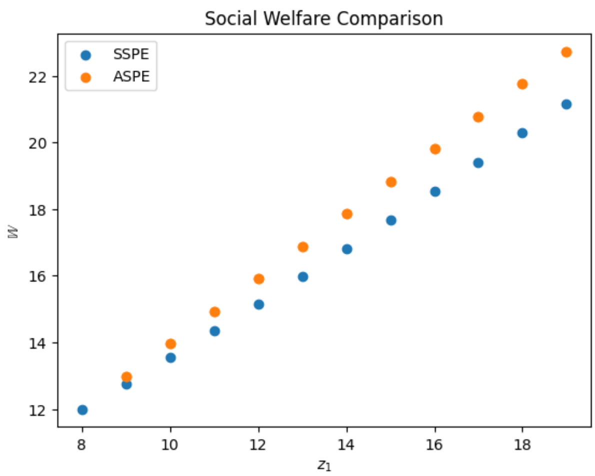


Figure shows welfare comparison for SSPE and ASPE when $\lambda = 1$, $z_1 \in [8, 20]$, $z_2 = 8$, $p_1 = p_1^{-i} = \frac{1}{2}$.

Figure 7: Effects of Varying Productivity Differences on Social Welfare

Figure 7 shows some examples where social welfare under ASPE is higher than SSPE. Intuitively, there is 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.

Concluding Remarks. In this section, I lay down the fundamental building blocks for the GPDS model. To fully characterize workers' search behaviours, I devote a part to analyze the second-stage game. I show that biased defaults could contribute to biased search in the equilibrium. I also highlight the possibility of multiple equilibria, where workers could adopt symmetric or asymmetric equilibrium application strategies, leading to the possibility of observed randomized or narrow search behaviour. For the full game, Proposition 3 and 4 are strong results. While firms have the power to extract high markdowns when workers' attention costs are high, Proposition 3 shows that the level of monospony power differs and workers may be better off in ASPE than SSPE for some intermediate range of inattentiveness. Furthermore, unlike search under complete information, where pure strategies would lead to an equilibrium wage of 0, having some attention costs that permits the existence of asymmetric strategies that are close to the pure ones could allow workers to earn positive wage. On top of this, Proposition 4 indicates that ASPE could generate higher social welfare than SSPE. These highlight that ASPE can be both socially efficient and welfare-maximizing for workers under some condition, which give rise to important policy implications. Though the mechanism is not discussed in the current paper, but this model establishes the potential to dive into equilibrium selection as a policy measure in presence of multiple equilibria depending on the policy objectives.

3 Variations in Default Strategies and Cognitive Costs

Using the GPDS model, I further explore the impact of ex-ante worker heterogeneity in attention costs on equilibrium outcomes. Worker disparity are often analyzed along the lines of skills, differences in attention costs have yet to be analyzed, but they can also contribute to labour market mismatch. If we consider workers of different skill level, high-skilled workers could incur lower information processing cost than low-skilled workers when studying the same environment, they may also possess better social connections and more insider information about the wage distribution. While these can be accommodated by the model, I do not include skills heterogeneity because such distinction might pose complications about how firms treat the workers. They could discriminate between workers, choosing high-skilled workers due to their superior capabilities instead of their access to information. By assuming workers have the same skill level but different attention constraints, I can disentangle the two impact and provide an alternative policy tool centered on information costs. Therefore, I investigate 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 about the wage distribution due to insider information or firms having asymmetric reputation influence over the workers, thereby resulting in workers adopting different default strategies. Hereby, I explore how heterogeneous defaults could affect workers' equilibrium application 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, which can be due to 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 $p_1^{-i} < \frac{1}{2}$. Worker 1 choose firm 1 with higher probability and worker 2 choose firm 2 with higher probability by default.

Definition 3.1 (Strength of Belief Distortion)**.** *Building on Definition 2.4, I define strength of belief distortion (δ) to be the degree of deviation away from random search default.*
E.g. For $p_1 > \frac{1}{2}$ and $p_1^{-i} < \frac{1}{2}$, let $p_1 = \frac{1}{2} + \delta_1$ and $p_1^{-i} = \frac{1}{2} - \delta_2$, where $\delta_1, \delta_2 \in (0, \frac{1}{2})$. Large δ_1 and δ_2 indicate strong belief distortion, while small δ_1 and δ_2 indicate weak belief distortion.

Relative strength of belief distortion is a comparison of the asymmetry in default strategies between the workers:

$$\textit{Cases} = \begin{cases} p_1 = 1 - p_1^{-i}, \delta_1 = \delta_2, \textit{No relative distortion} \\ p_1 > 1 - p_1^{-i}, \delta_1 > \delta_2, \textit{Stronger worker 1 distortion} \\ p_1 < 1 - p_1^{-i}, \delta_1 < \delta_2, \textit{Stronger worker 2 distortion} \end{cases}, \textit{for } p_1 \in (\frac{1}{2}, 1), p_1^{-i} \in (0, \frac{1}{2}) \tag{52}$$

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 λ. As 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, one has incentive to acquire less information and adopt a different application strategy. In the special case of firms having the same productivity and set the same equilibrium wages. For any positive λ, workers with heterogeneous defaults but no relative distortion could incur the same information cost by choosing a random search strategy. However, since wages do not differ between firms, the expected payoff from random search would be 0, yet workers would incur positive information costs, so they would fare better by not acquiring any information and use their default strategies instead. Therefore, in contrast to the outcome with homogeneous workers, there are no SSPE as long as there are information costs.

ASPE. Both workers and firms' FOCs are modified to include heterogeneous default strategies. The ASPE solution, $(q_1^*, q_1^{-i^*}, w_1^*, w_2^*)$, can be determined with the following system of equations:

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

$$q_1^{-i^*} = \frac{1}{1 + \frac{1-p_1^{-i}}{p_1^{-i}}\exp(-(\frac{(2-q_1{}^*)w_1^* - (1+q_1{}^*)w_2^*}{2\lambda}))} \tag{54}$$

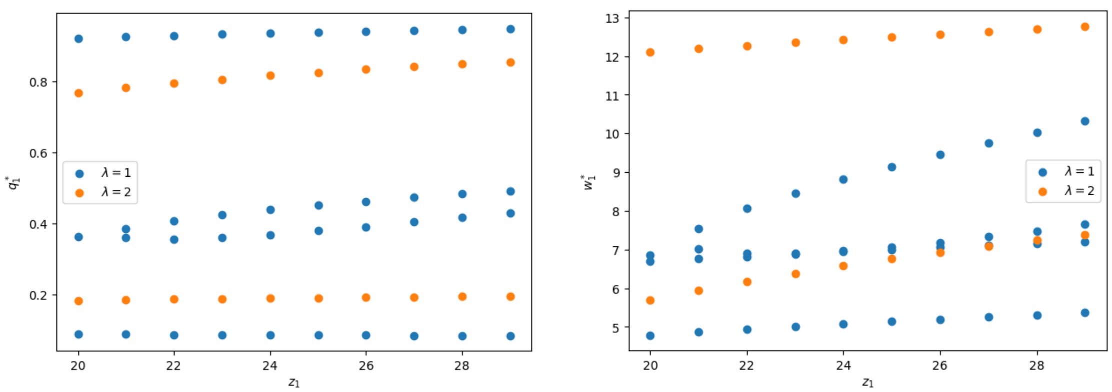
$$w_1^* = \max\{z_1 - \frac{\lambda(q_1^* + q_1^{-i^*} - q_1^* q_1^{-i^*})[1 - q_1^* q_1^{-i^*}(1-q_1^*)(1-q_1^{-i^*})(\frac{1}{q_1^* - q_1^{-i^*}}\log(\frac{1-p_1}{p_1}\frac{p_1^{-i}}{1-p_1^{-i}}\frac{q_1^*}{1-q_1^*}\frac{1-q_1^{-i^*}}{q_1^{-i^*}}))^2]}{(1-q_1^{-i^*})(1-q_1^*)(q_1^* + q_1^{-i^*} - q_1^* q_1^{-i^*}) - q_1^* q_1^{-i^*}(1-q_1^*)(1-q_1^{-i^*})(2q_1^* q_1^{-i^*} - 3q_1^* - 3q_1^{-i^*} + 4)\frac{1}{2(q_1^* - q_1^{-i^*})}\log(\frac{1-p_1}{p_1}\frac{p_1^{-i}}{1-p_1^{-i}}\frac{q_1^*}{1-q_1^*}\frac{1-q_1^{-i^*}}{q_1^{-i^*}})}, 0\} \tag{55}$$

$$w_2^* = \max\{z_2 - \frac{\lambda(1 - q_1^* q_1^{-i^*})[1 - q_1^* q_1^{-i^*}(1-q_1^*)(1-q_1^{-i^*})(\frac{1}{q_1^* - q_1^{-i^*}}\log(\frac{1-p_1}{p_1}\frac{p_1^{-i}}{1-p_1^{-i}}\frac{q_1^*}{1-q_1^*}\frac{1-q_1^{-i^*}}{q_1^{-i^*}}))^2]}{q_1^* q_1^{-i^*}(1 - q_1^* q_1^{-i^*}) - q_1^* q_1^{-i^*}(1-q_1^*)(1-q_1^{-i^*})(q_1^* + q_1^{-i^*} + 2q_1^* q_1^{-i^*})\frac{1}{2(q_1^* - q_1^{-i^*})}\log(\frac{1-p_1}{p_1}\frac{p_1^{-i}}{1-p_1^{-i}}\frac{q_1^*}{1-q_1^*}\frac{1-q_1^{-i^*}}{q_1^{-i^*}})}, 0\} \tag{56}$$

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 $\lambda > 0$.*

Proof. Shown in Appendix A.8.

I prove that workers will only apply symmetrically when $\lambda \to 0$, but as long as $\lambda > 0$, they apply asymmetrically. Using implicit function theorem, I also show that heterogeneous defaults could affect equilibrium multiplicity, particularly when there is relative belief distortion, which makes it analytically more complex to capture the entire range of equilibria. □



(a) Equilibrium Application Probabilities (b) Equilibrium Wages

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

Figure 8 illustrates for certain $(z_1, z_2, \lambda, p_1, p_1^{-i})$, there can be several $(q_1^*, q_1^{-i^*}, w_1^*, w_2^*)$. Multiple equilibria may arise when information costs are low. While the number of equilibria are not of particular interest, it highlights that heterogeneous defaults could be a potential reason behind greater variability in application strategies. Since SPE is absent when workers are inattentive and have heterogeneous defaults, it is necessary to consider ASE in order to accommodate worker heterogeneity and model its impact.

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.*

$$\frac{dq_1^{-i*}}{dp_1} > \frac{dq_1{}^*}{dp_1}, \frac{dq_1^{-i*}}{dp_1^{-i}} < \frac{dq_1{}^*}{dp_1^{-i}}$$

At high cognitive costs, the reverse happens, workers are more influenced by their own default.

$$\frac{dq_1^{-i*}}{dp_1} < \frac{dq_1{}^*}{dp_1}, \frac{dq_1^{-i*}}{dp_1^{-i}} > \frac{dq_1{}^*}{dp_1^{-i}}$$

Proof. Shown in Appendix A.9. □

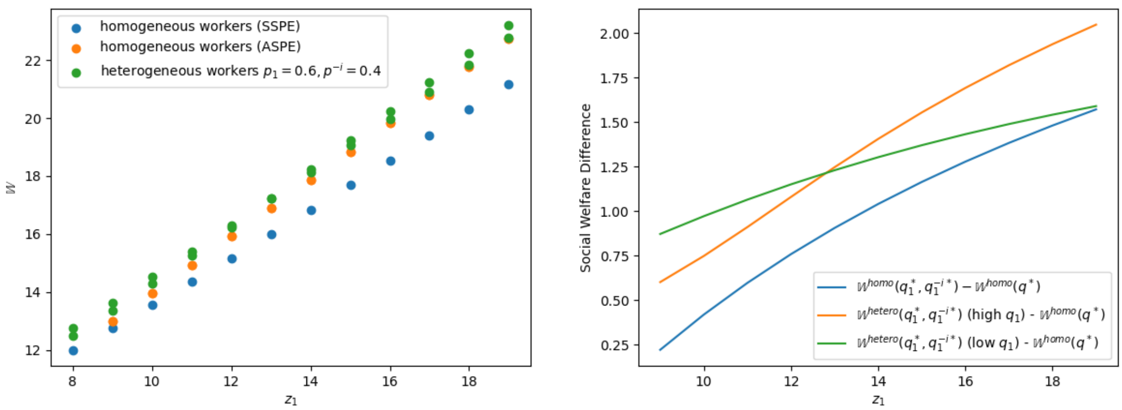
Workers factor in their opponent's biases when it is inexpensive to acquire information, and are swayed more by their own innate biases when information costs increase. In reality, as highlighted by Archer (2016), workers may hold deep-seated beliefs about the type of position they would go for. Briel et al. (2022) also suggests such beliefs may be gender-specific. These create strong defaults and could lead to workers hardly switch in the type of firms they apply to. However, having strong opposite beliefs could pay off as it may serve as a natural mechanism in facilitating better coordination and mitigate overcrowding to some extent. For instance, higher worker heterogeneity, reflected via extremely biased defaults (i.e. $p_1 \to 1$, $p_1^{-i} \to 0$) could lead to workers applying more to different firms in the equilibrium.

The additional behavioural implication from heterogeneous defaults could inevitably affect and complicate firms' wage setting behaviours. For the workers' side, the derivatives with respect to w_1, w_2 (i.e. (31), (32), (33), (34)) are not directly affected p_1, p_1^{-i}, so the general pattern of wage sensitivity remain the same, but workers would have different responsiveness to expected payoff, as characterized by the term $\frac{1-p_j^i}{p_j^i}$ in (53) and (54). As a result, on the firms' side, wages will be affected indirectly by defaults through workers' search behaviours. The wage setting equations (55) and (56) also imply that the relative strength of belief distortion plays a direct role in influencing wages.

Social Welfare with Heterogeneous Defaults.

$$\mathbb{W}^{hetero}(q_1^*, q_1^{-i*}) = (1-(1-q_1^*)(1-q_1^{-i*}))z_1 + (1-q_1{}^* q_1^{-i*})z_2 - \lambda[\sum_{j=1}^{2} q_j^* \log \frac{q_j^*}{p_j} + \sum_{j=1}^{2} q_j^{-i*} \log \frac{q_j^{-i*}}{p_j^{-i}}] \tag{57}$$

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. 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 depends on firm heterogeneity and most importantly, on equilibrium selection. In Figure 9b, even though social welfare is higher for all other equilibria as compared to symmetric equilibrium with homogeneous workers, $\mathbb{W}^{homo}(q^*)$, the difference



(a) Homogeneous vs. Heterogeneous

(b) Efficiency Compare to SSPE

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

grow smaller for ASPE with low q_1^* (i.e. green line) as firm heterogeneity increases. Given the trajectory of change, $\mathbb{W}^{homo}(q_1^*, q_1^{-i*})$ could surpass $\mathbb{W}^{hetero}(q_1^*, q_1^{-i*})$, indicating that heterogeneous defaults could also be a source of efficiency loss.

Benefit of Default Heterogeneity in Costly Search. When information is costless (i.e. $\lambda \to 0^+$), the game converge to one with complete information, there will not be any difference between the outcome of homogeneous workers and heterogeneous workers. However, as λ increases, defaults play a part in workers' search strategy. In which case, having some degree of heterogeneity in defaults might even be beneficial in reducing overcrowding 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 infinitely costly search (i.e. $\lambda \to \infty$), workers do not acquire any information, social welfare would depend entirely on the primitive default value and could be higher than homogenous default case as long as p_1 and p_1^{-i} are sufficiently different, such that workers apply more to different firms by default (see Appendix, Figure 21). Although this can be less interesting as workers' equilibrium application strategies are completely determined by the pre-defined default parameter, but it highlights the possibility that efficiency can be completely exogeneous in high cognitive costs environment. Policy measures in tackling prior belief formation could lead to better social outcome.

3.2 Worker Heterogeneity in Cognitive Costs

Workers could also differ in information processing capability or ability to acquire information, which can be described as variations in cognitive costs. Therefore, in this subsection, I explore how worker heterogeneity in cognitive costs affect equilibrium outcomes.

Model Assumption. I assume workers still have homogeneous, random search default strategy that follows Wu (2024), $p_1 = p_2 = \frac{1}{2}$, $p_1^{-i} = p_2^{-i} = \frac{1}{2}$. However, I relax the assumption on cognitive costs, such that $\lambda_i = [\lambda_1, \lambda_2]$, $\lambda \in \Lambda$, and λ_i can take different values for different workers. Henceforth, I suppose $\lambda_1 > \lambda_2$, worker 1 has higher cognitive cost than worker 2.

SSPE. Based on workers' FOCs, (5) and (6), if workers use 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 λs differ, the only possible equilibrium strategies would be when $q^* = \frac{1}{2}$ and $w_1 = w_2 = w^*$.

Firms' FOCs:

$$w_1 = z_1 - \frac{(w_1+w_2)^2 q^2(1-q) - 4\lambda_1\lambda_2\frac{1}{1-q}}{2(1-q)^2 q(w_1+w_2) - 2(1-q)(\lambda_1+\lambda_2)} \tag{58}$$

$$w_2 = z_2 - \frac{(w_1+w_2)^2(1-q)^2 q - 4\lambda_1\lambda_2\frac{1}{q}}{2q^2(1-q)(w_1+w_2) - 2q(\lambda_1+\lambda_2)} \tag{59}$$

and given $q^* = \frac{1}{2}$ and $w_1 = w_2 = w^*$,

$$w^* = z_1 - \frac{(2w^*)^2 - 64\lambda_1\lambda_2}{2(2w^*) - 8(\lambda_1+\lambda_2)}, w^* = z_2 - \frac{(2w^*)^2 - 64\lambda_1\lambda_2}{2(2w^*) - 8(\lambda_1+\lambda_2)} \tag{60}$$

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 wages than the other worker, thus workers would not behave symmetrically. Therefore, the only possible SSPE solution (q^*, w^*) is

$$q^* = \frac{1}{2}, w^* = \max[\frac{(\lambda_1+\lambda_2+0.5z) \pm \sqrt{(\lambda_1+\lambda_2+0.5z)^2 + 32\lambda_1\lambda_2 - 4\lambda_1 z - 4\lambda_2 z}}{2}, 0] \tag{61}$$

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

Even though in SSPE, q^* is independent of λ_1 and λ_2, w^* is affected by them. Higher the cognitive costs, lower the equilibrium wage. Firms can extract larger markdowns if the combined cognitive costs are high, as a result, the less cognitively constrained worker will be penalized even if one 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. Specifically, assuming* $\lambda_1 > \lambda_2$, $\frac{dw^*}{d\lambda_1} > \frac{dw^*}{d\lambda_2}$.

Proof. Shown in Appendix A.10. □

ASPE. Upon incorporating different cognitive costs, the ASPE solution, $(q_1^*, q_1^{-i^*}, w_1^*, w_2^*)$, can be determined with the following system of equations:

$$q_1^* = \frac{1}{1+\exp(-(\frac{(2-q_1^{-i^*})w_1^* - (1+q_1^{-i^*})w_2^*}{2\lambda_1}))} \tag{62}$$

$$q_1^{-i^*} = \frac{1}{1+\exp(-(\frac{(2-q_1{}^*)w_1^* - (1+q_1{}^*)w_2^*}{2\lambda_2}))} \tag{63}$$

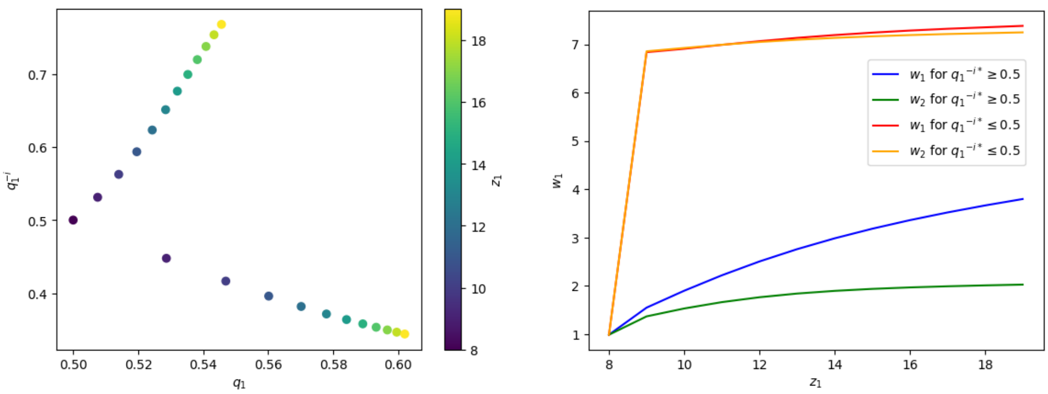
$$w_1^* = \max\{z_1 - \frac{\begin{array}{l}(q_1^* + q_1^{-i^*} - q_1^* q_1^{-i^*})[\lambda_1\lambda_2 \\ \quad - (\frac{1}{q_1^* - q_1^{-i^*}}(\lambda_1 \log(\frac{q_1^*}{1-q_1^*}) - \lambda_2\log(\frac{q_1^{-i^*}}{1-q_1^{-i^*}})))^2 q_1^* q_1^{-i^*}(1-q_1^*)(1-q_1^{-i^*})]\end{array}}{\begin{array}{l}(1-q_1^{-i^*})(1-q_1^*)(q_1^*+q_1^{-i^*} - q_1^* q_1^{-i^*}) - \frac{1}{2(q_1^*-q_1^{-i^*})}(\lambda_1\log(\frac{q_1^*}{1-q_1^*}) \\ \quad - \lambda_2\log(\frac{q_1^{-i^*}}{1-q_1^{-i^*}}))q_1^* q_1^{-i^*}(1-q_1^*)(1-q_1^{-i^*})(2q_1^* q_1^{-i^*} - 3q_1^* - 3q_1^{-i^*} + 4)\end{array}}, 0\} \tag{64}$$

$$
\begin{aligned}
w_2^* = \max\Big\{z_2 - \frac{\begin{aligned}&(1-q_1^* q_1^{-i^*})[\lambda_1\lambda_2 \\ &- q_1^* q_1^{-i^*}(1-q_1^*)(1-q_1^{-i^*})(\frac{1}{q_1^*-q_1^{-i^*}}(\lambda_1 \log(\frac{q_1^*}{1-q_1^*}) - \lambda_2 \log(\frac{q_1^{-i^*}}{1-q_1^{-i^*}})))^2]\end{aligned}}{q_1^* q_1^{-i^*}(1-q_1^* q_1^{-i^*}) - q_1^* q_1^{-i^*}(1-q_1^*)(1-q_1^{-i^*})(q_1^* + q_1^{-i^*} + 2q_1^* q_1^{-i^*})} \\
\frac{1}{2(q_1^*-q_1^{-i^*})}(\lambda_1 \log(\frac{q_1^*}{1-q_1^*}) - \lambda_2 \log(\frac{q_1^{-i^*}}{1-q_1^{-i^*}})), 0\Big\}
\end{aligned}
\tag{65}
$$

Lemma 3.2 (Existence of 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 more asymmetric subgame equilibrium (ASE), $(q_1^*, q_1^{-i^*}) \to (0,1)$ or $(q_1^*, q_1^{-i^*}) \to (1,0)$, to exist is small λ_1 and λ_2 relative to expected payoff, which is determined by the wage levels, and $2w_2 > w_1 > \frac{w_2}{2}$ needs to be fulfilled.*

Proof. Shown in Appendix A.11. □

Cognitive costs can also influence the number of equilibria. Decreasing the aggregate λs or lowering one of the λs could lead to equilibrium multiplicity, and the more extreme asymmetric strategies, where workers apply more to different firms, are also more likely to exist.



(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 10: Effects of Varying Productivity on Equilibrium Probabilities and Wages

In the example shown in Figure 10, 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 probability to firm 1 and worker 2 to firm 2. The first is fairly intuitive. Worker 2 takes advantage of the lower cost in acquiring and processing information, and apply with higher probability to the more productive firm. The second maybe slightly counter-intuitive, but could demonstrate that the more attentive worker accounts for the higher w_1 and also the higher possibility of overcrowding at firm 1. Figure 10b shows smaller wage differentials that correspond to this type of equilibrium, which indicates the benefit of "avoiding" the other worker could be larger than the expected gain from competing for the higher wage. Therefore, worker 2, who has lower cognitive cost, can be perceived as more strategic in choosing to apply with higher probability to the lower wage firm.

Proposition 7 (Heterogeneous Cognitive Cost on Equilibrium Strategies)**.** *Given firm 1 is more productive than firm 2 ($z_1 > z_2$), and worker 1 has higher cognitive cost than worker 2 ($\lambda_1 > \lambda_2$):*

- *For the more attentive worker 2 to avoid overcrowding by applying more to the less productive firm 2 ($q_1^{-i^*} < \frac{1}{2}$), the wage differential must be small.*
- *For worker 2 to leverage on the lower information cost by applying more to the more productive firm 1 ($q_1^{-i^*} > \frac{1}{2}$), the wage differential must be large.*

Proof. Shown in Appendix A.12. □

The changes in workers' search behaviour could similarly affect firms' wage setting. Heterogeneity in λ affects the wage sensitivities. The derivatives (31), (32), (33, (34) are reformulated:

$$\frac{dq_1}{dw_1} = \frac{(w_1+w_2)(2-q_1)q_1^{-i}(1-q_1^{-i})q_1(1-q_1) - 2\lambda_2(2-q_1^{-i})(1-q_1)q_1}{(w_1+w_2)^2 q_1^{-i}(1-q_1^{-i})q_1(1-q_1) - 4\lambda_1\lambda_2} \tag{66}$$

$$\frac{dq_1^{-i}}{dw_1} = \frac{(2-q_1^{-i})(1-q_1)q_1(1-q_1^{-i})q_1^{-i}(w_1+w_2) - 2\lambda_1 q_1^{-i}(1-q_1^{-i})(2-q_1)}{(w_1+w_2)^2 q_1^{-i}(1-q_1^{-i})q_1(1-q_1) - 4\lambda_1\lambda_2} \tag{67}$$

$$\frac{dq_1}{dw_2} = \frac{2\lambda_2(1+q_1^{-i})q_1(1-q_1) - q_1 q_1^{-i}(1-q_1)(1-q_1^{-i})(1+q_1)(w_1+w_2)}{q_1 q_1^{-i}(1-q_1)(1-q_1^{-i})(w_1+w_2)^2 - 4\lambda_1\lambda_2} \tag{68}$$

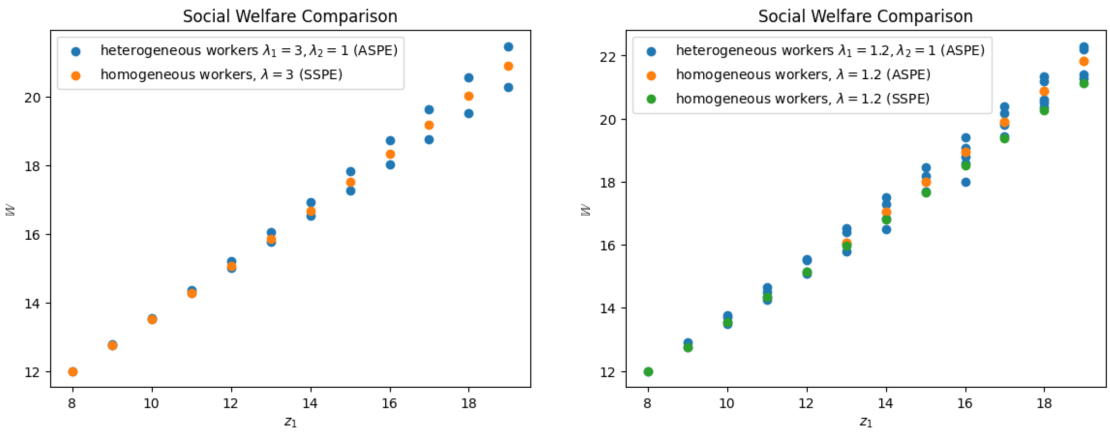
$$\frac{dq_1^{-i}}{dw_2} = \frac{2\lambda_1(1+q_1)q_1^{-i}(1-q_1^{-i}) - q_1 q_1^{-i}(1-q_1)(1-q_1^{-i})(1+q_1^{-i})(w_1+w_2)}{q_1 q_1^{-i}(1-q_1)(1-q_1^{-i})(w_1+w_2)^2 - 4\lambda_1\lambda_2} \tag{69}$$

If λs are substantially different, and $\lambda_2 \to 0^+$, 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 has lower cognitive cost, is less reactive to wage changes than worker 1. It is also possible for $\frac{dq_1^{-i}}{dw_1} < 0$ and $\frac{dq_1^{-i}}{dw_2} > 0$, which implies that worker 2 could be applying less to firm 1 despite increase in firm 1's wage and more to firm 1 when firm 2's wage increases. These demonstrate that workers have different wage senstivities, and worker 2 is more strategic in trading off between likelihood of being matched and gain from competing for the more productive firm.

In Figure 10b, equilibrium wages differ across ASPEs, and the equilibrium with workers applying more to different firms is associated with higher wages (i.e. red and orange lines). Similar to previous analysis, the reason behind this pattern is still attributed to the more aggressive initial wage setting behaviour, which may lead to higher overall wages, 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. In response to the looming overcrowding impact, worker 2 may switch firm in response, and to prevent this, firm 2 may set an even higher wage to retain attraction of worker 2. Firm 1 would also match up to avoid losing both workers and stay competitive. This is in contrast to workers applying closer to random, the initial wage do not need to be set as high as 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:

$$\mathbb{W}^{hetero}(q_1^*, q_1^{-i^*}) = (1-(1-q_1^*)(1-q_1^{-i^*}))z_1 + (1-q_1{}^* q_1^{-i^*})z_2 - \lambda_1 \sum_{j=1}^{2} q_j^* \log\frac{q_j^*}{p_j} - \lambda_2 \sum_{j=1}^{2} q_j^{-i^*} \log\frac{q_j^{-i^*}}{p_j} \tag{70}$$



(a) High combined λs & High Heterogeneity (b) Low combined λs & 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 11: Welfare Comparison between Homogeneous and Heterogeneous Workers

Figure 11 compares social welfare between homogeneous and heterogeneous workers. Figure 11a shows that when λs are high and disparate, social welfare given heterogeneous workers may be higher or lower than that of homogeneous workers, depending on equilibrium selection. In this example, the asymmetric 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 to direct workers to different firms, which may have outweighed the benefit from some degree of coordination. The pattern is similar when λs are low and close together. In Figure 11b, while SSPE is always less efficient than ASPE when workers are homogeneous, incorporation of heterogeneity could complicate the issue. Social welfare may 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 some degree of coordination attractive even after accounting for the cost of directing. Although worker heterogeneity makes efficiency analysis more nuanced, it is noticed 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 it is necessary to explore equilibrium selection mechanisms to nudge individuals towards a more efficient equilibrium. (Appendix B.5.)

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.*

$$\max_{q\in[0,1]} (1-(1-q)^2)z_1+(1-q^2)z_2-\lambda_1(q\log\frac{q}{p_1}+(1-q)\log\frac{1-q}{1-p_1})-\lambda_2(q\log\frac{q}{p_1^{-i}}+(1-q)\log\frac{1-q}{1-p_1^{-i}}) \tag{71}$$

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. There are no restriction on workers to use the same or different strategies.*

$$\begin{aligned}\max_{q_1, q_1^{-i}\in[0,1]} & (1-(1-q_1)(1-q_1^{-i}))z_1 + (1-q_1q_1^{-i})z_2 \\ & -\lambda_1[q_1\log\frac{q_1}{p_1} + (1-q_1)\log\frac{1-q_1}{1-p_1}] - \lambda_2[q_1^{-i}\log\frac{q_1^{-i}}{p_1^{-i}} + (1-q_1^{-i})\log\frac{1-q_1^{-i}}{1-p_1^{-i}}]\end{aligned} \tag{72}$$

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 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 λ and p), based on (71), the FOC is:

$$(1-q)z_1 - qz_2 - \lambda\log(\frac{1-p}{p}\frac{q}{1-q}) = 0 \tag{73}$$

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

$$\hat{q} = \frac{1}{1+\frac{1-p}{p}\exp\left(-(\frac{(1-\hat{q})z_1-\hat{q}z_2}{\lambda})\right)} \tag{74}$$

$$\frac{d\hat{q}}{dp} = \frac{\frac{\lambda}{(1-p)p}}{z_1+z_2+\frac{\lambda}{\hat{q}(1-\hat{q})}} > 0 \text{ for } \lambda > 0, p\in(0,1) \tag{75}$$

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 the belief distortion towards firm 1, higher the socially optimal choice probability of firm 1.

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

$$\mathbb{W}(\hat{q}) = (1-(1-\hat{q})^2)z_1 + (1-\hat{q}^2)z_2 - 2\lambda(\hat{q}\log(\frac{\hat{q}}{p}) + (1-\hat{q})\log(\frac{1-\hat{q}}{1-p})) \tag{76}$$

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 have heterogeneous productivity ($z_1 \neq z_2$).*

Proof. Shown in Appendix A.13.

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

$$\lim_{\lambda \to 0^+} \hat{q} = \frac{z_1}{z_1 + z_2} \tag{77}$$

$$\lim_{\lambda \to 0^+} q^* = \frac{2w_1^* - w_2^*}{w_1^* + w_2^*}, \text{ where } q^* \in (0, 1) \tag{78}$$

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

$$\lim_{\lambda \to \infty} \hat{q} = p, \lim_{\lambda \to \infty} q^* = p \tag{79}$$

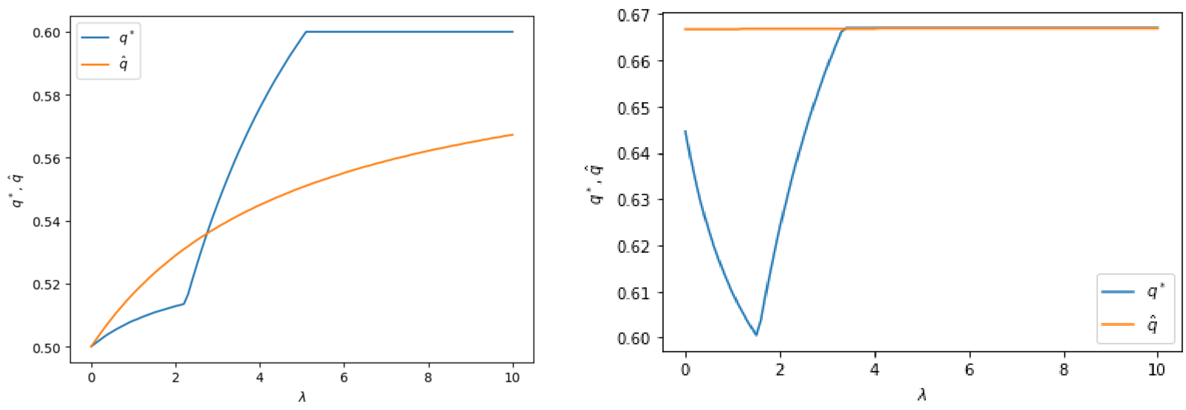
If prior beliefs are accurate and firms are homogeneous, then $q^* = \hat{q} = \frac{1}{2}$, SSPE is efficient. However, any other variations could lead to inefficiency. If prior beliefs are incorrect, then as λ increases, q^* and $\hat{q}$ converge to p at different rate, SSPE is inefficient for positive and affordable information cost. When facing heterogeneous firms, even if prior beliefs are correct, SSPE can still be inefficient. This applies even when $\lambda \to 0^+$.

Since $\hat{q}$ and q^* converge at different rate to p as λ increases, there may be a threshold crossing point, $\bar{\lambda}$, where $q^* = \hat{q}$, depending on the direction of convergence.

$$\bar{\lambda} = \frac{z_1 - \frac{z_1 + (z_1 - w_1^*) - (w_1^* - w_2^*)}{z_1 + z_2 + (z_1 - w_1^*) + (z_2 - w_2^*)}(z_1 + z_2)}{\log\left(\frac{1-p}{p} \frac{z_2 + (z_2 - w_2^*) + (w_1^* - w_2^*)}{z_1 + z_2 + (z_1 - w_1^*) + (z_2 - w_2^*)}\right)} \tag{80}$$

It can be affected by default choice probability and firm heterogeneity. □

By social planner standard, workers are generally not learning the optimal wage precision, leading to inefficiency. To illustrate 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 12: Comparison between Socially Optimal Solution $\hat{q}$ and SSPE solution q^*

Figure 12a shows while workers could start off searching efficiently when information is costless, both equilibrium strategy (q^*) and socially optimal strategy ($\hat{q}$) are affected by incorrect defaults as λ increases, but at different rates, and q^* systematically diverge away from $\hat{q}$, except at the crossing point. While inaccurate defaults provide a compelling case for inefficiency as workers incur 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 12b 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 cognitive cost 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, leading to the same wage-productivity ratio (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 wage information instead of productivities, to deviate away from the optimal strategy. This form of inefficiency is inherit in the nature of the 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 wage information accordingly. Only when information becomes sufficiently costly that workers no longer acquire any information. This results in them sticking to their defaults, which correctly correspond to productivity distribution, and the outcome would be efficient. This mechanism highlights the additional source of inefficiency caused by inattention. (Appendix B.6.)

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

$$((1-\hat{q})^2 - (1-q_1^*)(1-{q_1}^{-i*}))z_1 + (\hat{q}^2 - q_1^* {q_1}^{-i*})z_2 \\ > \lambda[\sum_{j=1}^{2} q_j^* \log\frac{q_j^*}{p_j} + \sum_{j=1}^{2} q_j^{-i*} \log\frac{q_j^{-i*}}{p_j} - 2\hat{q}\log(\frac{\hat{q}}{p}) - 2(1-\hat{q})\log(\frac{1-\hat{q}}{1-p})] \tag{81}$$

Although the conditions for λ are generally more binding for ASPE to exist, thus more likely to satisfy (81). The analysis is complex and subjected to specific values of exogenous parameters.

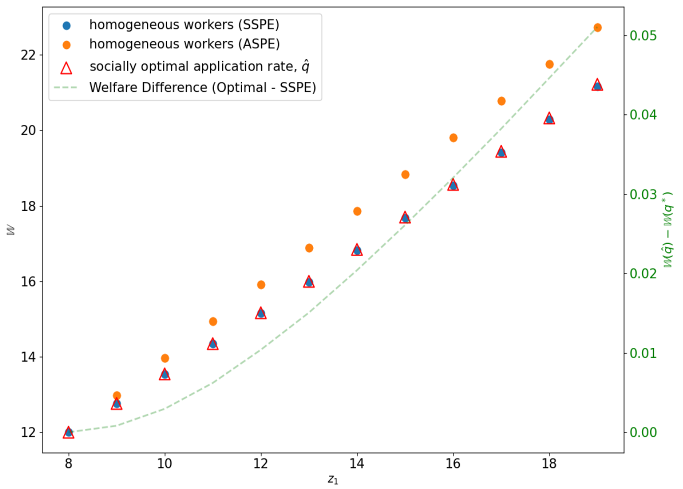


Figure show for homogeneous workers having $\lambda = 1$, figure shows z_1 ranging from 8 to 20, $z_2 = 8$, $p_1 = p_1^{-i} = 0.5$.

Figure 13: Comparison between Socially Optimal, ASPE and SSPE Outcomes

Figure 13 illustrates an example where constrained social planner is not as efficient as ASPE. Furthermore, even though welfare outcomes under SSPE appear close to socially optimal, the difference is positive (as shown by the green line). SSPE is inefficient and yields the least welfare. This example also highlights that focusing solely on comparison between SSPE and socially optimal outcome, overlooking the possibility of ASPE, could be detrimental to 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 (71), the FOC to determine socially optimal strategy, $\hat{q}$, considering for *heterogeneous defaults* is:

$$2(1-\hat{q})z_1 - 2\hat{q}z_2 - \lambda[\log(\frac{1-p_1}{p_1}\frac{\hat{q}}{1-\hat{q}}) + \log(\frac{1-p_1^{-i}}{p_1^{-i}}\frac{\hat{q}}{1-\hat{q}})] = 0 \tag{82}$$

where $p_1 \neq p_1^{-i}$, and as before, assume $p_1 > \frac{1}{2}$ and $p_1^{-i} < \frac{1}{2}$. Expressing (82) 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 $\hat{q}$.

$$2z_1 - \lambda\log(\frac{1-p_1}{p_1}\frac{1-p_1^{-i}}{p_1^{-i}}) = 2\lambda\log(\frac{\hat{q}}{1-\hat{q}}) + 2\hat{q}(z_1+z_2) \tag{83}$$

Proposition 9 (Heterogeneous Defaults on Socially Optimal Strategy)**.** *Socially optimal solution, $\hat{q}$, is positively affected by default probabilities, p_1 and p_1^{-i}. Worker with stronger belief distortion can influence the socially optimal strategy in one's favoured direction. Specifically, for $p_1 > \frac{1}{2}$ and $p_1^{-i} < \frac{1}{2}$, as p_1 increases against $(1-p_1^{-i})$, $\hat{q}$ increases.*

Proof. Shown in Appendix A.14. □

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

$$2(1-\hat{q})z_1 - 2\hat{q}z_2 - (\lambda_1+\lambda_2)\log(\frac{\hat{q}}{1-\hat{q}}) = 0 \tag{84}$$

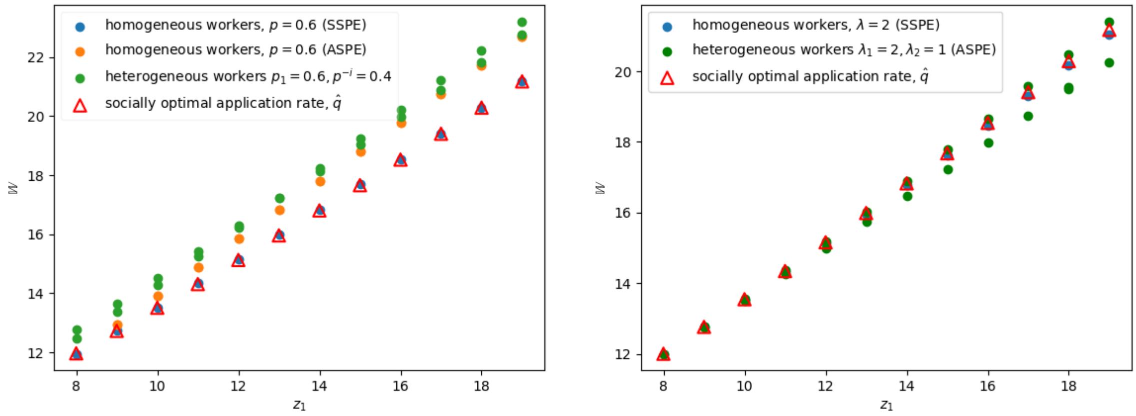
$$\hat{q} = \frac{\exp\frac{2(1-\hat{q})z_1-2\hat{q}z_2}{\lambda_1+\lambda_2}}{1+\exp\frac{2(1-\hat{q})z_1-2\hat{q}z_2}{\lambda_1+\lambda_2}} \tag{85}$$

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, $\frac{d\hat{q}}{d\lambda_1} < \frac{d\hat{q}}{d\lambda_2} < 0$.*

Proof. Shown in Appendix A.15. □

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 14 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 14a, equilibrium choices with heterogeneous workers could



(a) Heterogeneous p

(b) Heterogeneous λ

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

Figure 14: Comparison between Equilibrium and Socially Optimal Outcomes

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 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 14b, 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.16)

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 workers' narrow search behaviour that is demonstrated in Belot et al. (2019)'s survey results. This divergence can be explained by the GPDS model. For instance, in the simplest case where workers are assumed to be homogeneous, the GPDS model would predict that in presence of multiple equilibria, workers could display more targeted search behaviour and assign different choice probability to the firms. Since workers in the survey are not pre-specified to follow the same application strategy, the GPDS model provides a more realistic portrayal of the possible equilibrium outcomes. Furthermore, by implementing a search tool intervention, Wu (2024) suggests workers would use a more directed search strategy, whereas Belot et al. (2019) shows a broadening of search, and only if the original strategy is relatively broad, workers might apply narrower than before. This discrepancy could be explained by workers being nudged by the measure to switch away from the asymmetric or symmetric application strategies to a socially optimal one that can either be less or more targeted. Therefore, through accounting for the possibility of equilibrium multiplicity, the GPDS model can accommodate both outcomes. However, it may be argued that workers that use job centre services have high attention costs and they would only behave symmetrically, thus simply considering for multiple equilibrium may not suffice in explaining for the presence of asymmetric strategies. The GPDS model has the additional benefit of accounting for workers' heterogeneity in attention costs, in which case, workers could behave asymmetrically even at high cognitive costs. As a result, any recommendations to switch away from such equilibrium strategies to a socially optimal one may imply maneuver away from more targeted search to broadening of scope. In view of these, the GPDS model is shown to be more generalizable, and may provide slightly more accurate predictions of the labour market outcomes.

Beauty of Small Set-up. Although the GPDS model consists of a simplified 2×2 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, and extrapolating to a multi-agent setting simply 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 fits particularly well with the idea of consideration sets postulated by Hauser and Wernerfelt (1990) and Caplin et al. (2019), where workers may only consider a subset of the firms when making their application decisions. They could be placing 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. It can be hypothesized that the less the consideration sets overlap, the more likely workers can matched up with a firm. This may stimulate discussions on mechanisms 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. Therefore, as long as the consideration set includes more than one option, the findings from this paper is applicable.

Parallels to Other Empirical Studies. Apart from the literature that motivates the research question, the GPDS model also demonstrates its ability in explaining other prevailing 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, and found such measure increases employment. This can be represented in the GPDS model as lowering the unit information processing cost (or cognitive cost, λ) for a group of workers as essential information become easier to acquire and process. By introducing worker heterogeneity, the model predicts possibility of multiple equilibria, including cases where workers may be applying more to different firms. There could be higher chances of matching and improved social welfare. Nonetheless, Abebe et al. (2020) analyzed the impact of randomized job-fair interventions on labour market outcome. They found such events correct workers' beliefs about the labour market, thereby creating more matches at the end, particularly for the less educated. Deriving this from the GPDS model, it can be assumed that cognitive costs are high for the less educated, as a result, they will be fully reliant on their default strategies. Through adjusting the default strategies to match with the underlying productivity distribution, more efficient outcome can be achieved.

Policy Implications. This framework offers opportunities for policy interventions that targets information presentation to improve matching, rather than through reducing time or monetary search costs. Imposing stricter pay transparency laws could be a possible measure to mitigate information friction. For example, EU has passed the Pay Transparency Directive in 2023, which requires companies to disclose information on the starting salary or pay range for the advertized jobs, either in the vacancy details or ahead of interview. (Europa (2023)) While the main objective of this policy is to detect and reduce gender pay gap, the GPDS model postulates that revealing the wage range or the exact wages during the advertizing stage could have an additional benefit of lowering information friction due to inattention. However, since attention costs can also be interpreted as workers observing the exact wages but failed to process the information (Heng et al. (2020)) or have imprecise recall of wages when making decisions (Azeredo da Silveira et al. (2024)), pay transparency laws would not be helpful in this aspect. Furthermore, as workers might not enter labour market as a blank slate, based on the GPDS model, mismatch can be deconstructed to a consequence of innate bias, there can be policies to harmonize prior beliefs to match with market conditions to achieve more efficient outcome. Nonetheless, the possibility of equilibrium multiplicity and some equilibria being more efficient than others also advocates for equilibrium selection measures.

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. There are also no distinctions between cost of searching, processing, retaining and retrieval of information, which are all bundled together as attention cost. Even though this makes the model more generalizable, it might be worthwhile to disentangle the types of costs and isolate their effect on workers' application behaviour. Another potential limitation is in regards to the exceptional level of rationality demanded by this model. Even though it is possible for empirical data to exhibit patterns that resemble or trend towards the predicted equilibrium outcomes, achieving such outcomes in practice could be challenging. Similar to challenges faced by quantal response equilibrium, to achieve equilibrium outcome in the GPDS 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 one's rationality. Furthermore, workers could have wrong perception of their attention costs, or allocate disproportionate attention across the firms, creating another layer of imprecision that may be better modelled using behavioural inattention (Gabaix (2019)). However, despite this limitation, the GPDS model serves as a valuable baseline to understand workers' equilibrium behaviour under inattention, and upon which behavioural elements can be further incorporated.

Extensions. One of the main contributions of the paper is proving the existence of multiple equilibria, which brought forth an equilibrium selection problem. This opens up the possibility of exploring equilibrium selection mechanism, such as through workers learning from past experiences. Since job search occurs over one's lifetime, which constitutes repeated interactions and learning based past feedback, inattentive workers could enter the market uninformatively, or hold certain prior beliefs. They acquire and process information in each period and revise their beliefs about the underlying wage states, and learn to best respond to their opponents given the information they have received. The long-run equilibrium consists of workers optimally allocating their attention resources. Policies tackling attention allocation and easier information processing may stimulate faster learning process and influence which equilibrium the workers coordinate on when multiple equilibria exist. This provides potential for social planner to nudge workers towards socially preferred equilibrium outcome.

Another question that has not been examined extensively is the impact of competition on attention. The GPDS model illustrates a competitive landscape, where many workers can simultaneously compete for vacancies. Suppose there are more workers than firms, greater competition among workers could lead to a decrease in wages as firms, having higher probability of meeting at least a worker, can extract higher markdowns. Workers might respond by executing less effort, or paying less attention, which can in turn makes it harder for efficient matching outcome. This hypothesis could merit future exploration. Last but not least, experiments could be carried out as a next step to further validate the findings and investigate the plausible mechanisms to induce more efficient outcome.

6 Conclusion

This paper seeks to enhance the modelling of job search by providing a generalized framework that fully captures the role of worker's inattention on labour market outcome. It specifically accounts for imperfect search arising from both cognitive limitation and bias, where workers can have diverse priors. I show that their equilibrium behaviours are swayed by bias in their default search strategies. I also prove the existence of multiple equilibria, where workers can use the same or different application strategies that comprise of them applying narrowly, with high probability to different firms. Heterogeneity in attention costs could also lead to greater asymmetry in workers' strategies as well as larger variety of equilibrium application strategies. The model shows that workers' narrow search behaviour and diversity in application strategies, exhibited in empirical findings, could naturally arise from inattention, thereby closing the gap between theoretical predictions and empirical results in the literature. Most importantly, it posits that mismatch can happen even in a small set-up, but highlights a silver lining that some equilibria may be more efficient than others amidst the cognitive constraints. Equilibrium where workers apply asymmetrically generates higher social welfare, and potentially also lower monopsony power, than them applying symmetrically. The differences across the equilibria support exploration of equilibrium selection measures depending on policy objectives. Nonetheless, the model contributes to the grand vision of deconstructing macro phenomenon into individual decision-making problems, which postulates the plausibility of bottom-up policies to mitigate individual-level information frictions that could have a profound effect on market-level mismatch.

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., Pohlan, 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 Results

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:

$$\frac{d\mathcal{L}}{dq_1^i(\mathbf{s})} : \int_{\mathbf{w}} [(1 - \frac{q_1^{-i}(\mathbf{s})}{2})w_1 - (q_1^{-i}(\mathbf{s}) + \frac{1 - q_1^{-i}(\mathbf{s})}{2})w_2]F^i(d\mathbf{s}|\mathbf{w})G^{-i}(d\mathbf{w}) = 0 \tag{86}$$

$$\frac{d\mathcal{L}}{dq_1^{-i(\mathbf{s})}(\mathbf{s})} : \int_{\mathbf{w}} [(1 - \frac{q_1^i(\mathbf{s})}{2})w_1 - (q_1^i(\mathbf{s}) + \frac{1 - q_1^i(\mathbf{s})}{2})w_2]F^{-i}(d\mathbf{s}|\mathbf{w})G^i(d\mathbf{w}) = 0 \tag{87}$$

$$\frac{d\mathcal{L}}{F^i(\mathbf{s}|\mathbf{w})} : F^i(\mathbf{s}|\mathbf{w}) = \exp[-(1 + \frac{\int_{\mathbf{w}}[q_1^i(\mathbf{s})(1 - \frac{q_1^{-i}(\mathbf{s})}{2})w_1 + (1 - q_1^i(\mathbf{s}))(q_1^{-i} + \frac{1-q_1^{-i}(\mathbf{s})}{2})w_2]G^{-i}(d\mathbf{w})}{\lambda_i})] \tag{88}$$

$$\frac{d\mathcal{L}}{F^{-i}(\mathbf{s}|\mathbf{w})} : F^{-i}(\mathbf{s}|\mathbf{w}) = \exp[-(1 + \frac{\int_{\mathbf{w}}[q_1^{-i}(\mathbf{s})(1 - \frac{q_1^i(\mathbf{s})}{2})w_1 + (1 - q_1^{-i}(\mathbf{s}))(q_1^i(\mathbf{s}) + \frac{1-q_1^i(\mathbf{s})}{2})w_2]G^i(d\mathbf{w})}{\lambda_{-i}})] \tag{89}$$

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 $F^i(\mathbf{s}|\mathbf{w}) = F^{-i}(\mathbf{s}|\mathbf{w}) = F^e(\mathbf{s}|\mathbf{w})$, $q_1^i(\mathbf{s}) = q_1^{-i}(\mathbf{s}) = q_1^e(\mathbf{s})$,

$$F^e(\mathbf{s}|\mathbf{w}) = \exp[-(1 + \frac{\int_{\mathbf{w}}[q_1^e(\mathbf{s})(1 - \frac{q_1^e(\mathbf{s})}{2})w_1 + (1 - q_1^e(\mathbf{s}))(q_1^e(\mathbf{s}) + \frac{1-q_1^e(\mathbf{s})}{2})w_2]G^e(d\mathbf{w})}{\lambda})] \tag{90}$$

$$q_1^e(\mathbf{s}) = \frac{2\mathbb{E}[w_1] - \mathbb{E}[w_2]}{\mathbb{E}[w_1] + \mathbb{E}[w_2]}, \text{where } \mathbb{E}[w_1] = \int_{\mathbf{w}} w_1 F(d\mathbf{s}|\mathbf{w})G(d\mathbf{w}), \mathbb{E}[w_2] = \int_{\mathbf{w}} w_2 F(d\mathbf{s}|\mathbf{w})G(d\mathbf{w}) \tag{91}$$

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 a 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),

$$\frac{p}{1-p}\exp(\frac{2w_1 - w_2 - q(w_1 + w_2)}{2\lambda}) = \frac{q}{1-q} \tag{92}$$

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

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

$$\frac{p}{1-p}\exp(\frac{(2-q)w_1}{2\lambda}) = \frac{q}{1-q} \tag{93}$$

Express p in terms of q,

$$p = \frac{qA^{q-2}}{1-q+qA^{q-2}}, \text{ where } \exp(\frac{w_1}{2\lambda} = A), A > 0 \tag{94}$$

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

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

$f'(q) > 0$ 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 $w_1 = 0$, $w_2 > 0$,

$$\frac{p}{1-p}\exp(\frac{-(1+q)w_2}{2\lambda}) = \frac{q}{1-q} \tag{95}$$

Express p in terms of q,

$$p = \frac{qB^{1+q}}{1-q+qB^{1+q}}, \text{ where } \exp(\frac{w_2}{2\lambda} = B), B > 0 \tag{96}$$

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

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

$g'(q) > 0$ 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 (92) and express p in terms of q,

$$p = \frac{qC^{-1}D^{-q}}{1-q+qC^{-1}D^{-q}}, \text{ where } \exp(\frac{2w_1 - w_2}{2\lambda}) = C, \exp(\frac{-(w_1+w_2)}{2\lambda}) = D, C, D > 0 \tag{97}$$

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

$$\lim_{q\to 0} h(q) = -p < 0 \tag{98}$$

$$\lim_{q\to 1} h(q) = 1-p > 0 \tag{99}$$

$h'(q) > 0$ 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

$$h(\cdot) = \frac{1}{1+\frac{1-p}{p}\exp(-(\frac{(2-q_1^{-i^*})w_1-(1+q_1^{-i^*})w_2}{2\lambda}))} - q_1^* - \frac{1}{1+\frac{1-p}{p}\exp(-(\frac{(2-q_1{}^*)w_1-(1+q_1{}^*)w_2}{2\lambda}))} + q_1^{-i^*} \tag{100}$$

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

$$h(0, 1, w_1, w_2) = \frac{1}{1 + \frac{1-p}{p}\exp(-(\frac{w_1 - 2w_2}{2\lambda}))} - \frac{1}{1 + \frac{1-p}{p}\exp(-(\frac{2w_1 - w_2}{2\lambda}))} + 1 > 0 \tag{101}$$

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

$$h(1, 0, w_1, w_2) = \frac{1}{1 + \frac{1-p}{p}\exp(-(\frac{2w_1 - w_2}{2\lambda}))} - \frac{1}{1 + \frac{1-p}{p}\exp(-(\frac{w_1 - 2w_2}{2\lambda}))} - 1 < 0 \tag{102}$$

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

$$\frac{1}{\lambda}[(1 - q_1^{-i^*} + \frac{q_1^{-i^*}}{2})w_1 - (q_1^{-i^*} + \frac{1 - q_1^{-i^*}}{2})w_2] = \log(\frac{1-p}{p}\frac{q_1^*}{1 - q_1^*}) \tag{103}$$

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

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

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

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

And for (28),

$$\frac{1}{\lambda}[(1 - q_1^* + \frac{q_1^*}{2})w_1 - (q_1^* + \frac{1 - q_1^*}{2})w_2] = \log(\frac{1-p}{p}\frac{q_1^{-i^*}}{1 - q_1^{-i^*}}) \tag{104}$$

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

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

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

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

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 $\lambda \to 0^+$ with no other wage conditions,

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

With that, I specified the wage condition, which is small cognitive costs (λ) 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 (103) could be met in the limit but not (104), then equilibrium conditions do not hold, there will not be solution for $q_1^{-i^*} \to 0$ and $q_1^* \to 1$. 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 $q_1^{-i^*} \to 1$ and $q_1^* \to 0$.

Since (101) and (102) 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 $q_1^* \to q^{*-}$ and $q_1^{-i^*} \to q^{*+}$,

$$\begin{aligned} h(q^{*-}, q^{*+}, w_1, w_2) = & \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2w_1 - w_2 - q^{*+}(w_1+w_2))}{2\lambda}))} \\ & - \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2w_1 - w_2 - q^{*-}(w_1+w_2))}{2\lambda}))} - (q^{*-} - q^{*+}) < 0 \end{aligned} \tag{105}$$

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

$$\begin{aligned} h(q^{*+}, q^{*-}, w_1, w_2) = & \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2w_1 - w_2 - q^{*-}(w_1+w_2))}{2\lambda}))} \\ & - \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2w_1 - w_2 - q^{*+}(w_1+w_2))}{2\lambda}))} - (q^{*+} - q^{*-}) > 0 \end{aligned} \tag{106}$$

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

The sign of (105) and (106) would depend on the values of w_1, w_2 and λ. If λ 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 λ 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 λ relative to w. $h(\cdot)$ could potentially cuts 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 (105) and (106). If the inequalities break down, they combine to form the following inequality, which would result in <u>one</u> unique set of solution for $(q_1^*, q_1^{-i^*})$,

$$\frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2w_1 - w_2 - q^{*-}(w_1+w_2))}{2\lambda}))} - \frac{1}{1 + \frac{1-p}{p} \exp(-(\frac{(2w_1 - w_2 - q^{*+}(w_1+w_2))}{2\lambda}))} < 2\epsilon \tag{107}$$

$$\frac{1}{1+\frac{1-p}{p}ab} - \frac{1}{1+\frac{1-p}{p}\frac{a}{b}} < 2\epsilon, a = \exp(-(\frac{(2w_1 - w_2 - q^*(w_1+w_2)}{2\lambda})), b = \exp(-\frac{\epsilon(w_1+w_2)}{2\lambda}) \tag{108}$$

For any variation in λ, w_1 and w_2, a moves the first and second term of (108) in the same direction, making the extent of change in b essential. For large λ and small w_1 and w_2, it is more likely to fulfill (108), and therefore, a unique solution will be anticipated instead of multiple. The main purpose of (108) is for us to evaluate the potential threshold of λ, such that if λ 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 λ 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*, λ would have to be sufficiently small relative to expected payoff, which can directly mean that λ 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 λ relative to expected payoffs. Based on the wage determination equations (39) and (40), the wage condition can be reformulated as

$$2\max[z_1 - m_1(q_1^*, q_1^{-i^*}, \lambda), 0] > \max[z_2 - m_2(q_1^*, q_1^{-i^*}, \lambda), 0] > \frac{\max[z_1 - m_1(q_1^*, q_1^{-i^*}, \lambda), 0]}{2} \tag{109}$$

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*}, λ. Generally speaking, the extent at which the productivity could differ is approximately twice.

Return to the Section 2.3.1.

A.4 Wages given Feasible Asymmetric Subgame Strategies

Suppose $q_1^* \to 1$ and $q_1^{-i*} \to 0$, let RHS of the inequality (43) be $h(\cdot)$, $\lim_{q_1^* \to 1, q_1^{-i*} \to 0} h(\cdot) > 0$. Similarly, suppose $q_1^* \to 0$ and $q_1^{-i*} \to 1$, $\lim_{q_1^* \to 0, q_1^{-i*} \to 1} h(\cdot) > 0$. In both cases, $h(\cdot) \to \infty$ and (43) no longer holds. (for *i.*)

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

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

Same applies for $q_1^* \to q^{*-}$ and $q_1^{-i*} \to q^{*+}$. $h(\cdot)$ would depend on the exact value of q^*. Graphing for $0 < q^* < 1$ shows a minimum point at $q^* = \frac{1}{2}$. If q^* deviates away from $\frac{1}{2}$, the combined productivity relative to cognitive costs, $\frac{z_1+z_2}{\lambda}$, needs to be larger to satisfy the inequality condition. The uniformity of q^* can be induced by workers having less belief distortion and firms being more homogeneous. (for *ii.*)

Since wages are continuous functions of q_1^* and q_1^{-i*} when they exist, by intermediate value theorem, w_j^{ASPE} would increase continuously as q_1^* and q_1^{-i*} become less asymmetric, and it could exceed w_j^{SSPE}. Given the condition in (*ii.*), $\frac{z_1+z_2}{\lambda} \geq \frac{1-2q^*+2q^{*2}}{q^{*2}(1-q^*)^2}$, has been met, then in the limit, $w_j^{\text{ASPE}} \geq w_j^{\text{SSPE}}$. (for *iii.*)

Return to the Section 2.3.2.

A.5 Lower Monopsony Power in ASPE

Following SSPE system of equations (19), (17) and (18), I derive (1).

To derive (2), I follow ASPE system of equations (37), (38), (39), (40), workers' strategies are symmetrically asymmetrical, with each worker favouring different firm in similar balance of probability, $q_1^* = 1 - \epsilon$ and $q_1^{-i*} = \epsilon$:

$$q_1^* = 1 - q_1^{-i*} = \frac{1}{1 + \exp(-(\frac{(1-2\epsilon)w^{ASPE}}{2\lambda}))} \tag{110}$$

$$w^{ASPE} = \max\{z - \lambda \frac{(1-\epsilon+\epsilon^2)[1 - (\frac{1}{1-2\epsilon}\log(\frac{1-\epsilon}{\epsilon}\frac{1-\epsilon}{\epsilon}))^2(1-\epsilon)^2\epsilon^2]}{(1-\epsilon)\epsilon(1-\epsilon+\epsilon^2) - \frac{1}{2(1-2\epsilon)}\log(\frac{1-\epsilon}{\epsilon}\frac{1-\epsilon}{\epsilon}) \atop (1-\epsilon)^2\epsilon^2(1+2\epsilon-2\epsilon^2)}, 0\} \tag{111}$$

Since firms are homogeneous, they offer the same wage. Based on wage derivatives, (35), (36):

$$\frac{1-(1-q_1)(q_1)}{z-w} = \frac{(2w)(1-q_1)^2 q_1^2[(1+q_1)(1-q_1) + (2-q_1)(q_1)] \atop - 2\lambda(q_1)(1-q_1)[(1-q_1)(2-q_1) + q_1(1+q_1)]}{(2w)^2 q_1^2 (1-q_1)^2 - 4\lambda^2} \tag{112}$$

$$\frac{1-q_1(1-q_1)}{z-w} = -\{\frac{(2w)(q_1)^2(1-q_1)^2[(2-q_1)(1-q_1)+q_1(1+q_1)] - 2\lambda q_1(1-q_1)[(1-q_1)(1+q_1)+(q_1)(2-q_1)]}{(2w)^2 q_1^2(1-q_1)^2 - 4\lambda^2}\} \tag{113}$$

To ensure both equations are compatible and hold simultaneously, the following consistency condition between w, λ and ϵ needs to hold:

$$w = \frac{\lambda}{\epsilon(1-\epsilon)} \tag{114}$$

Under Lemma 2.2, ASPE cease to exist when $\lambda \to 0$ because expected payoff is driven to 0. Limit of $\lambda \to 0$ and $\epsilon \to 0$ or $\epsilon \to 1$ does not exist. However, given ASPE exists, for small λ, ϵ lay strictly between 0 and 1, under wage consistency condition (114), $w^{ASPE} > 0$ but can remain small and close to 0.

From (114), the threshold of λ for equilibrium wages to be higher in ASPE than SSPE:

$$w^{ASPE} \geq \frac{z}{2} - 2\lambda \Rightarrow \lambda \geq \frac{z}{2}\frac{\epsilon(1-\epsilon)}{1+2\epsilon(1-\epsilon)}$$

This gives rise to inequality (44), and ϵ is an implicit solution of equation (45) that comes from combining (110) and (111). This threshold indicates that λ has to be sufficiently large to balance the trade-off between productivity, z, and probabilistic behaviour of workers, $\epsilon(1-\epsilon)$.

As λ increases, but still satisfies Lemma 2.2, then by (110), $\epsilon \to 0.5$, and by (111), $w^{ASPE} \to z - 4\lambda$. $w^{ASPE} \geq w^{SSPE}$ if $z - 4\lambda \geq \frac{z}{2} - 2\lambda$, which simplifies to $\frac{z}{\lambda} \geq 4$. Since q_1^* and $q_1^{-i^*}$ are continuous, and wages are continuous functions of q_1^* and $q_1^{-i^*}$, by intermediate value theorem, w_j^{ASPE} could exceed w_j^{SSPE} as λ increases from 0 (Lemma 2.3).

Return to the Section 2.3.2.

A.6 Variations in Defaults Given Existence of Multiple Equilibria

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

$$\begin{aligned}\frac{dq_1^*}{dp} = &\frac{q_1^*(1-q_1^*)[(2-q_1^*)q_1^{-i^*}(1-q_1^{-i^*})(w_1+w_2) - 2\lambda(2-q_1^{-i^*})]}{(w_1+w_2)^2 q_1^{-i^*}(1-q_1^{-i^*})q_1^*(1-q_1^*) - 4\lambda^2}\frac{dw_1^*}{dp} \\ &- \frac{q_1^*(1-q_1^*)[(1+q_1^*)q_1^{-i^*}(1-q_1^{-i^*})(w_1^*+w_2^*) - 2\lambda(1+q_1^{-i^*})]}{(w_1^*+w_2^*)^2 q_1^{-i^*}(1-q_1^{-i^*})q_1^*(1-q_1^*) - 4\lambda^2}\frac{dw_2^*}{dp} \\ &+ \frac{[2\lambda(w_1^*+w_2^*)q_1^{-i^*}(1-q_1^{-i^*}) - 4\lambda^2]q_1^*(1-q_1^*)}{p(1-p)[(w_1^*+w_2^*)^2 q_1^{-i^*}(1-q_1^{-i^*})q_1^*(1-q_1^*) - 4\lambda^2]}\end{aligned} \tag{115}$$

$$\begin{aligned}\frac{dq_1^{-i^*}}{dp} = &\frac{q_1^{-i^*}(1-q_1^{-i^*})[(2-q_1^{-i^*})q_1^*(1-q_1^*)(w_1^*+w_2^*) - 2\lambda(2-q_1^*)]}{(w_1^*+w_2^*)^2 q_1^*(1-q_1^*)q_1^{-i^*}(1-q_1^{-i^*}) - 4\lambda^2}\frac{dw_1^*}{dp} \\ &- \frac{q_1^{-i^*}(1-q_1^{-i^*})[(1+q_1^{-i^*})q_1^*(1-q_1^*)(w_1^*+w_2^*) - 2\lambda(1+q_1^*)]}{(w_1^*+w_2^*)^2 q_1^*(1-q_1^*)q_1^{-i^*}(1-q_1^{-i^*}) - 4\lambda^2}\frac{dw_2^*}{dp} \\ &+ \frac{q_1^{-i^*}(1-q_1^{-i^*})[2\lambda(w_1^*+w_2^*)q_1^*(1-q_1^*) - 4\lambda^2]}{(1-p)p[(w_1^*+w_2^*)^2 q_1^*(1-q_1^*)q_1^{-i^*}(1-q_1^{-i^*}) - 4\lambda^2]}\end{aligned} \tag{116}$$

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.[5] Any deviation away from $p = \frac{1}{2}$ has a direct positive effect on $\frac{dq_1^*}{dp}$ and

[5]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.

$\frac{dq_1^{-i*}}{dp}$, 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.

A.7 Higher Social Welfare in ASPE than SSPE

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

Step 1: From (51), I first demonstrate that the expected productivity/ex-post welfare is higher under ASPE than SSPE:

$$((1-q^*)^2 - (1-q_1^*)(1-q_1^{-i*}))z_1 + (q^{*2} - q_1^* q_1^{-i*})z_2 > 0 \tag{117}$$

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

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

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

(117) is satisfied.

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

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

In the limit, the expected productivity difference between ASPE and SSPE tends to 0, but for any positive ϵ, 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 λ is negligible, then (51) always hold. For $\lambda > 0$, the magnitude of the RHS matters.

Relative cost of different equilibrium strategies. Obstructing away λ, define RHS of (51) to be a function $\mathbb{S}(q_j^*, q_j^{-i*}, q^*)$,

$$\mathbb{S}(\cdot) = \sum_{j=1}^{2} q_j^* \log \frac{q_j^*}{p_j} + \sum_{j=1}^{2} q_j^{-i*} \log \frac{q_j^{-i*}}{p_j} - 2q^* \log(\frac{q^*}{p}) - 2(1-q^*)\log(\frac{1-q^*}{1-p}) \tag{118}$$

Define $q_1^* \to 1$, $q_1^{-i*} \to 0$ or $q_1^* \to 0$, $q_1^{-i*} \to 1$ as $\lim_{pure}$. From (118),

$$\lim_{pure} \mathbb{S}(\cdot) = (1-2q^*)\log(\frac{1-p}{p}) - 2q^* \log(\frac{q^*}{1-q^*}) - 2\log(1-q^*) \tag{119}$$

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

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

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

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

$\lim_{pure} \mathbb{S}(\cdot)$ 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 (51) to break down.

Next, define $q_1^* \to q^{*-}$ and $q_1^{-i*} \to q^{*+}$, or $q_1^* \to q^{*+}$ and $q_1^{-i*} \to q^{*-}$ as $\lim_{mixed}$. From (118),

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

For small ϵ, the expression is positive but remains low, with a minimum at $q^* \to \frac{1}{2}$. 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 (51) is likely to hold.

Cognitive costs. To investigate the magnitude of λ for the inequality (51) to remain true. Reformulating (51):

$$\lambda < \frac{((1-q^*)^2 - (1-q_1^*)(1-q_1^{-i*}))z_1 + (q^{*2} - q_1^* q_1^{-i*})z_2}{\sum_{j=1}^2 q_j^* \log \frac{q_j^*}{p_j} + \sum_{j=1}^2 q_j^{-i*} \log \frac{q_j^{-i*}}{p_j} - 2q^* \log(\frac{q^*}{p}) - 2(1-q^*)\log(\frac{1-q^*}{1-p})} \tag{120}$$

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

$$0 < \lambda < \frac{(1-\frac{q^*}{2})w_1^* - (\frac{q^*+1}{2})w_2^*}{\log(\frac{1-p}{p}\frac{q^*}{1-q^*})} \tag{121}$$

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 λ for existence of ASPE is stricter than λ required for fulfilling the (120), thus welfare is in general higher given ASPE exist. For instance, if workers behave close to pure coordination, λ required for ASPE tp exist needs to be close to 0, but remain positive for (120). As we move from that to more symmetric strategies, λ for existence of ASPE is slightly relaxed, it would be necessary to verify

$$\frac{(1-\frac{q^*}{2})w_1^* - (\frac{q^*+1}{2})w_2^*}{\log(\frac{1-p}{p}\frac{q^*}{1-q^*})} < \frac{(1-q^*)^2 z_1 + q^{*2} z_2}{\log(\frac{1}{1-p}) + \log(\frac{1}{p}) - 2q^* \log(\frac{q^*}{p}) - 2(1-q^*)\log(\frac{1-q^*}{1-p})} \tag{122}$$

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 (122),

$$\frac{(\frac{1}{2} - q^*)w^*}{\log(\frac{1-p}{p}\frac{q^*}{1-q^*})} < \frac{((1-q^*)^2 + q^{*2})z}{\log(\frac{1}{1-p}) + \log(\frac{1}{p}) - 2q^* \log(\frac{q^*}{p}) - 2(1-q^*)\log(\frac{1-q^*}{1-p})} \tag{123}$$

For unbiased workers (i.e. $p = \frac{1}{2}$), $q^* = \frac{1}{2}$ when facing homogeneous firms, LHS of (123) 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 (122) 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:

$$\frac{(\frac{1}{2}-q^*)w_2^* + (1-\frac{q^*}{2})\delta x}{\log(\frac{1-p}{p}\frac{q^*}{1-q^*})} < \frac{((1-q^*)^2+q^{*2})z_2 + (1-q^*)^2 x}{\log(\frac{1}{1-p}) + \log(\frac{1}{p}) - 2q^*\log(\frac{q^*}{p}) - 2(1-q^*)\log(\frac{1-q^*}{1-p})} \tag{124}$$

As x becomes substantially large, unbiased workers would apply with higher probability to firm 1, q^* approaches 1, LHS of (124) would be

$$\lim_{x\to\infty, q^*\to 1} \frac{(\frac{1}{2}-q^*)w_2^* + (1-\frac{q^*}{2})\delta x}{\log(\frac{1-p}{p}\frac{q^*}{1-q^*})} = 0 \tag{125}$$

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

$$\lim_{x\to\infty, q^*\to 1} \frac{((1-q^*)^2+q^{*2})z_2 + (1-q^*)^2 x}{\log(\frac{1}{1-p}) + \log(\frac{1}{p}) - 2q^*\log(\frac{q^*}{p}) - 2(1-q^*)\log(\frac{1-q^*}{1-p})} = \frac{z_2}{\log(\frac{p}{1-p})} \tag{126}$$

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 δ is small, then (124) 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 δ, 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.8 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:

$$q_1 : (1 - q_1^{-i} + \frac{q_1^{-i}}{2})w_1 - (q_1^{-i} + \frac{1-q_1^{-i}}{2})w_2 = \lambda \log(\frac{1-p_1}{p_1}\frac{q_1}{1-q_1}) \tag{127}$$

$$q_1^{-i} : (1 - q_1 + \frac{q_1}{2})w_1 - (q_1 + \frac{1-q_1}{2})w_2 = \lambda \log(\frac{1-p_1^{-i}}{p_1^{-i}}\frac{q_1^{-i}}{1-q_1^{-i}}) \tag{128}$$

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

$$\frac{p_1}{1-p_1}\exp(\frac{(2-q_1^{-i*})w_1 - (1+q_1^{-i*})w_2}{2\lambda}) - \frac{q_1^*}{1-q_1^*} = 0 \tag{129}$$

$$\frac{p_1^{-i}}{1-p_1^{-i}}\exp(\frac{(2-q_1^*)w_1-(1+q_1^*)w_2}{2\lambda})-\frac{q_1^{-i*}}{1-q_1^{-i}}=0 \tag{130}$$

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

$$f(\cdot)=\frac{p_1}{1-p_1}\exp(\frac{(2-q_1^{-i*})w_1-(1+q_1^{-i*})w_2}{2\lambda})-\frac{q_1^*}{1-q_1^*} \tag{131}$$

$$g(\cdot)=\frac{p_1^{-i}}{1-p_1^{-i}}\exp(\frac{(2-q_1^*)w_1-(1+q_1^*)w_2}{2\lambda})-\frac{q_1^{-i*}}{1-q_1^{-i}} \tag{132}$$

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,

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

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

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 $\lambda \to 0$, by the system of equations for ASPE (53), (54), (55) and (56), $q_1^* = q_1^{-i*} = \frac{2w_1-w_2}{w_1+w_2}$, workers will behave symmetrically. This showcase the outcome of complete information. However, in this model, λ may be negligible but remain positive, $\lambda \to 0^+$, workers would not behave symmetrically, but can be fairly close.

Then, suppose $\lambda \to \infty$, rewriting workers' FOCs (127) and (128), taking a limit on both sides,

$$\text{LHS: } \lim_{\lambda\to\infty}\frac{(1-q_1^{-i*}+\frac{q_1^{-i*}}{2})w_1^*-(q_1^{-i*}+\frac{1-q_1^{-i*}}{2})w_2^*}{\log(\frac{1-p_1}{p_1}\frac{q_1^*}{1-q_1^*})}=\frac{(2-q_1^{-i*})w_1^*-(1+q_1^{-i*})w_2^*}{2\log(\frac{1-p_1}{p_1}\frac{q_1^*}{1-q_1^*})}$$

$$\text{RHS: } \lim_{\lambda\to\infty}\lambda=\infty$$

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

$$\lim_{q_1^*\to p_1^-}\frac{(2-q_1^{-i*})w_1-(1+q_1^{-i*})w_2}{2\log(\frac{1-p_1}{p_1}\frac{q_1^*}{1-q_1^*})}=\infty \text{ if } (2-q_1^{-i*})w_1-(1+q_1^{-i*})w_2<0$$

$$\lim_{q_1^*\to p_1^+}\frac{(2-q_1^{-i*})w_1-(1+q_1^{-i*})w_2}{2\log(\frac{1-p_1}{p_1}\frac{q_1^*}{1-q_1^*})}=\infty \text{ if } (2-q_1^{-i*})w_1-(1+q_1^{-i*})w_2>0$$

The same applies for (128), as $\lambda \to \infty$, limit on both sides equalize only if $q_1^{*-i} \to p_1^{-i}$.

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

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, (127) and (128), find the limit as $q_1^* \to 1, q_1^{-i*} \to 0$. For (127),

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

$$\text{LHS: } \frac{p_1}{1-p_1}\lim_{q_1^*\to 1, q_1^{-i*}\to 0}\exp(\frac{(2-q_1^{-i*})w_1-(1+q_1^{-i*})w_2}{2\lambda}) = \frac{p_1}{1-p_1}\exp(\frac{2w_1-w_2}{2\lambda})$$

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

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

Finding the limit for (128) ends up with similar conditions: $\lambda \to 0^+$, $2w_2 > w_1$, or if $p_1^{-i} \to 0$.

For $q_1^* \to 0, q_1^{-i*} \to 1$, (127) shows

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

$$\text{LHS: } \frac{p_1}{1-p_1}\lim_{q_1^*\to 0, q_1^{-i*}\to 1}\exp(\frac{(2-q_1^{-i*})w_1-(1+q_1^{-i*})w_2}{2\lambda}) = \frac{p_1}{1-p_1}\exp(\frac{w_1-2w_2}{2\lambda})$$

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

For LHS=RHS, $\lambda \to 0^+$ and $w_1 < 2w_2$ need to hold, such that $\lim_{\lambda\to 0^+}\exp(\frac{w_1-2w_2}{2\lambda}) = 0$. Otherwise, if $p_1 \to 0$, 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 (128), 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 λ 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. $p_1 \to 1$ and $p_1^{-i} \to 0$).

Step 4: I use the implicit function theorem to determine the presence of multiple solutions for the system of equations defined by (53), (54), (55) and (56). From (131) and (132), where $f(\cdot)$ and $g(\cdot)$ are continuously differentiable functions. Define the Jacobian matrix to be $J(f,g) = \begin{bmatrix} \frac{\partial f}{\partial q_1^*} & \frac{\partial f}{\partial q_1^{-i*}} \\ \frac{\partial g}{\partial q_1^*} & \frac{\partial g}{\partial q_1^{-i*}} \end{bmatrix}$.

$$\det(J(f,g)(a,b)) = \frac{\partial f}{\partial q_1^*}(a,b)\cdot\frac{\partial g}{\partial q_1^{-i*}}(a,b) - \frac{\partial f}{\partial q_1^{-i*}}(a,b)\cdot\frac{\partial g}{\partial q_1^*}(a,b) \tag{133}$$

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),

$$\frac{\partial f}{\partial q_1^*} = \frac{p_1}{1-p_1}\frac{1}{2\lambda}\exp(\frac{(2-q_1^{-i^*})w_1-(1+q_1^{-i^*})w_2}{2\lambda})[(2-q_1^{-i^*})\frac{\partial w_1}{\partial q_1^*}-(1+q_1^{-i^*})\frac{\partial w_2}{\partial q_1^*}]-\frac{1}{(1-q_1^*)^2}$$

$$\frac{\partial f}{\partial q_1^{-i^*}} = \frac{p_1}{1-p_1}\frac{1}{2\lambda}\exp(\frac{(2-q_1^{-i^*})w_1-(1+q_1^{-i^*})w_2}{2\lambda})[(2-q_1^{-i^*})\frac{\partial w_1}{\partial q_1^{-i^*}}-(1+q_1^{-i^*})\frac{\partial w_2}{\partial q_1^{-i^*}}-w_1-w_2]$$

$$\frac{\partial g}{\partial q_1^*} = \frac{p_1^{-i}}{1-p_1^{-i}}\frac{1}{2\lambda}\exp(\frac{(2-q_1^*)w_1-(1+q_1^*)w_2}{2\lambda})[(2-{q_1}^*)\frac{\partial w_1}{\partial q_1^*}-(1+{q_1}^*)\frac{\partial w_2}{\partial q_1^*}-w_1-w_2]$$

$$\frac{\partial g}{\partial q_1^{-i^*}} = \frac{p_1^{-i}}{1-p_1^{-i}}\frac{1}{2\lambda}\exp(\frac{(2-q_1^*)w_1-(1+q_1^*)w_2}{2\lambda})[(2-{q_1}^*)\frac{\partial w_1}{\partial q_1^{-i^*}}-(1+{q_1}^*)\frac{\partial w_2}{\partial q_1^{-i^*}}]-\frac{1}{(1-q_1^{-i^*})^2}$$

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

Suppose $\lambda \to \infty$, 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 $q_1^* \to 0$ and $q_1^{-i^*}$,

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

It was found that $\det(J(f,g)) > 0$ for non-zero and reasonably small λ relative to expected payoffs. The same applies for $q_1^* \to 1$ and $q_1^{-i^*} \to 0$. 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.9 Heterogeneous Defaults on Equilibrium Search Behaviour

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

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

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

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

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

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 λ 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 λ 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.10 Heterogeneous Cognitive Costs on Equilibrium Wages in SSPE.

To show the impact of heterogeneous cognitive costs on equilibrium wages in SSPE. Based on equation (61),

$$\frac{dw^*}{d\lambda_1} = \frac{-64\lambda_2 + 8(z-w^*)}{-16w^* + 8\lambda_1 + 8\lambda_2 + 4z} \tag{134}$$

$$\frac{dw^*}{d\lambda_2} = \frac{-64\lambda_1 + 8(z-w^*)}{-16w^* + 8\lambda_1 + 8\lambda_2 + 4z} \tag{135}$$

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

Return to the Section 3.2.

A.11 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 subgame equilibrium.

Rewriting (62) and (63),

$$\exp(\frac{(2-q_1^{-i*})w_1 - (q_1^{-i*}+1)w_2}{2\lambda_1}) - \frac{q_1^*}{1-q_1^*} = 0 \tag{136}$$

$$\exp(\frac{(2-q_1^*)w_1-(q_1^*+1)w_2}{2\lambda_2})-\frac{q_1^{-i^*}}{1-q_1^{-i^*}}=0 \tag{137}$$

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

$$f(\cdot)=\exp(\frac{(2-q_1^{-i^*})w_1-(q_1^{-i^*}+1)w_2}{2\lambda_1})-\frac{q_1^*}{1-q_1^*} \tag{138}$$

$$g(\cdot)=\exp(\frac{(2-q_1^*)w_1-(q_1^*+1)w_2}{2\lambda_2})-\frac{q_1^{-i^*}}{1-q_1^{-i^*}} \tag{139}$$

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 $q_1^* = 0$,

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

For $q_1^* = 1$,

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

Assuming $\lambda_1 > \lambda_2$, for low λs 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 $q_1^* = 0$,

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

For $q_1^* = 1$,

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

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

These shows that for small λs 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 (62) and (63) and finding the limit as $q_1^* \to 1$, and $q_1^{-i^*} \to 0$. For (62),

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

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

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

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

Similarly, finding the limit for (63) as $q_1^* \to 1$, and $q_1^{-i*} \to 0$,

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

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

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

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

Furthermore, finding the limit as $q_1^* \to 0$, and $q_1^{-i*} \to 1$. For (62),

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

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

To satisfy LHS=RHS, $\lambda_1 \to 0^+$, $2w_2 > w_1$. For (63),

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

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

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

The conditions for ASE, where workers almost purely coordinate, to exist, would be $\lambda_1 \to 0^+$, $\lambda_2 \to 0^+$ 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 (138) and (139), 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 λ_1 and λ_2 are independently and also jointly important in influencing the equilibrium multiplicity.

Return to the Section 3.2.

A.12 Heterogeneous Cognitive Costs on Equilibrium Strategies in ASPE

To demonstrate heterogeneous cognitive costs on equilibrium strategies, I first start with (62) and (63),

$$\exp(\frac{(2-q_1^{-i*})w_1^* - (q_1^{-i*}+1)w_2^*}{2\lambda_1}) - \frac{q_1^*}{1-q_1^*} = 0 \tag{140}$$

$$\exp(\frac{(2-q_1^*)w_1^* - (q_1^*+1)w_2^*}{2\lambda_2}) - \frac{q_1^{-i*}}{1-q_1^{-i*}} = 0 \tag{141}$$

Set LHS of (140) and (141) 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 $q_1^* = \frac{1}{2}$,

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

$f(\cdot) = 0$ 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 $q_1^* = \frac{1}{2}$,

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

For $q_1^{-i*} > \frac{1}{2}$, $g(\cdot)$ will decreases. 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.13 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 $\lambda \to 0^+$, from (73), the socially optimal $\hat{q}$ is

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

$$\lim_{\lambda \to 0^+} \hat{q} = \frac{z_1}{z_1 + z_2} \tag{142}$$

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

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

$$\lim_{\lambda \to 0^+} q^* = \frac{2w_1^* - w_2^*}{w_1^* + w_2^*}, \text{ where } q^* \in (0,1) \tag{143}$$

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 $\lambda \to 0^+$, $q^* = \frac{1}{2}$ and $\hat{q} = \frac{1}{2}$, the equilibrium strategy would be efficient.

If $\lambda \to \infty$,

$$\lim_{\lambda \to \infty} \hat{q} = p, \lim_{\lambda \to \infty} q^* = p \tag{144}$$

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 $\lambda \to 0^+$ and $\lambda \to \infty$. This suggests that given firms are homogeneous, $q^* = \hat{q}$ regardless of prior beliefs being correct or incorrect at extreme λ values.

Step 2: Relaxing the assumption on extreme values of λ, and evaluating for intermediary range of λs.

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 (73) 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 $\lambda \to 0^+$, $\hat{q} = \frac{1}{2}$ and $q^* = \frac{1}{2}$. As λ 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 λ, $\frac{d\hat{q}}{d\lambda}$ and $\frac{dq^*}{d\lambda}$,

$$\frac{d\hat{q}}{d\lambda} = -\frac{\log(\frac{1-p}{p}\frac{\hat{q}}{1-\hat{q}})}{2z + \lambda\frac{1}{\hat{q}(1-\hat{q})}} \tag{145}$$

$$\frac{dq^*}{d\lambda} = \frac{(2-q^*)\frac{2q^*}{1-q^*} + (1+q^*)\frac{2(1-q^*)}{q^*} - 2\log(\frac{1-p}{p}\frac{q^*}{1-q^*})}{4(1-2q^*)z + (w_1^* + w_2^*) + \frac{4q^{*3} - 8q^{*2} + 2}{(q^*-1)^2 q^{*2}}\lambda} \tag{146}$$

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

When λ 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 λ 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 λ, workers should apply less to firm 1 in equilibrium, and at a point, as λ 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,

$$q^* = \hat{q} = \frac{2(z - w_1^*) + w_2^*}{4z - (w_1^* + w_2^*)} = \frac{z + (z - w_1^*) - (w_1^* - w_2^*)}{2z + (z - w_1^*) + (z - w_2^*)} \tag{147}$$

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 λs 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 $\lambda \to 0^+$, $\hat{q} = \frac{1}{2}$ and $q^* = \frac{1}{2}$. As λ increases, $\hat{q}$ and q^* converge to p, $\hat{q}$ and q^* are upward sloping curves.

Based on (145) and (146), when λ 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 λ, the relationship reverse. Since the curves are upward sloping, this indicates that at low λ, $\hat{q} > q^*$, workers should have applied more to firm 1, and as λ 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 (147), 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 λs 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 $\lambda \to 0^+$, $\hat{q} \to \frac{z_1}{z_1+z_2}$, and $q^* \to \frac{2w_1^*-w_2^*}{w_1^*+w_2^*}$, with the wage determination equations,

$$\lim_{\lambda\to 0^+} q^* = \frac{2w_1^* - w_2^*}{w_1^* + w_2^*} = \frac{(3-2q^*-q^{*2})z_1 - (q^{*2}+2q^*)z_2}{(2q^*-2q^{*2})(z_1+z_2)} \tag{148}$$

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

$$\lim_{\lambda\to 0^+} q^* = \frac{2w_1^* - w_2^*}{w_1^* + w_2^*} = \frac{(3-2q^*-q^{*2})x - (q^{*2}+2q^*)}{(2q^*-2q^{*2})(x+1)} \tag{149}$$

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

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

For intermediary range of λs. 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 λ, but $z_1 \neq z_2$:

$$\frac{d\hat{q}}{d\lambda} = -\frac{\log(\frac{1-p}{p}\frac{\hat{q}}{1-\hat{q}})}{z_1 + z_2 + \lambda\frac{1}{\hat{q}(1-\hat{q})}} \tag{150}$$

$$\frac{dq^*}{d\lambda} = \frac{(2-q^*)\frac{2q^*}{1-q^*} + (1+q^*)\frac{2(1-q^*)}{q^*} - 2\log(\frac{1-p}{p}\frac{q^*}{1-q^*})}{2(1-2q^*)(z_1+z_2) + (w_1^*+w_2^*) + \frac{4q^{*3}-8q^{*2}+2}{(q^*-1)^2q^{*2}}\lambda} \tag{151}$$

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 $\lambda \to 0^+$, then for any intermediary λ, 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 λ, 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 $\lambda \to 0^+$, $\hat{q} > q^*$, and as λ 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.

$$\text{Cases for } p > \frac{z_1}{z_1+z_2} = \begin{cases} \hat{q} > q^*, \text{given small } \lambda \\ q^* = \hat{q}, \text{threshold crossing at } \frac{z_1+(z_1-w_1^*)-(w_1^*-w_2^*)}{z_1+z_2+(z_1-w_1^*)+(z_2-w_2^*)} \\ \hat{q} < q^*, \text{given large } \lambda \end{cases} \tag{152}$$

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

$$\bar{\lambda} = \frac{z_1 - \frac{z_1+(z_1-w_1^*)-(w_1^*-w_2^*)}{z_1+z_2+(z_1-w_1^*)+(z_2-w_2^*)}(z_1+z_2)}{\log(\frac{1-p}{p}\frac{z_2+(z_2-w_2^*)+(w_1^*-w_2^*)}{z_1+z_2+(z_1-w_1^*)+(z_2-w_2^*)})} \tag{153}$$

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 λ 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.14 Heterogeneous Defaults on Socially Optimal Strategy

To explore the impact of heterogeneous defaults on socially optimal solution, $\hat{q}$, based on (82), find $\frac{d\hat{q}}{dp_1}$ and $\frac{d\hat{q}}{dp_1^{-i}}$:

$$\frac{d\hat{q}}{dp_1} = \frac{\frac{\lambda}{(1-p_1)p_1}}{2(z_1 + z_2 + \frac{\lambda}{(1-\hat{q})\hat{q}})} > 0, \frac{d\hat{q}}{dp_1^{-i}} = \frac{\frac{\lambda}{(1-p_1^{-i})p_1^{-i}}}{2(z_1 + z_2 + \frac{\lambda}{(1-\hat{q})\hat{q}})} > 0 \tag{154}$$

Express (82) in terms of $\hat{q}$ on the RHS and p_1 and p_1^{-i} on the LHS,

$$2z_1 - \lambda \log(\frac{1-p_1}{p_1}\frac{1-p_1^{-i}}{p_1^{-i}}) = 2\lambda \log(\frac{\hat{q}}{1-\hat{q}}) + 2\hat{q}(z_1 + z_2) \tag{155}$$

Given exogeneously determined value for z_1, z_2 and λ, 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 (155) 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 (155) 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.15 Heterogeneous Cognitive Costs on Socially Optimal Strategy

To determine the impact of heterogeneous cognitive costs on socially optimal solution, $\hat{q}$, based on (82), find $\frac{d\hat{q}}{d\lambda_1}$ and $\frac{d\hat{q}}{d\lambda_2}$:

$$\frac{d\hat{q}}{d\lambda_1} = \frac{-\log(\frac{\hat{q}}{1-\hat{q}}) - \lambda_1(\frac{1}{\hat{q}} + \frac{1}{1-\hat{q}})}{2z_1 + 2z_2} \tag{156}$$

$$\frac{d\hat{q}}{d\lambda_2} = \frac{-\log(\frac{\hat{q}}{1-\hat{q}}) - \lambda_2(\frac{1}{\hat{q}} + \frac{1}{1-\hat{q}})}{2z_1 + 2z_2} \tag{157}$$

Given $\lambda_1 > \lambda_2$, change in λ_1 by the same amount as λ_2 will lead to a larger decrease in $\hat{q}$.

Return to Section 4.1.

A.16 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, $\mathbb{W}(q_1^*, q_1^{-i^*}) > \mathbb{W}(\hat{q})$,

$$((1-\hat{q})^2 - (1-q_1^*)(1-q_1^{-i^*}))z_1 + (\hat{q}^2 - q_1^* q_1^{-i^*})z_2$$
$$> \lambda[\sum_{j=1}^{2} q_j^* \log \frac{q_j^*}{p_j} + \sum_{j=1}^{2} q_j^{-i^*} \log \frac{q_j^{-i^*}}{p_j^{-i}} - \sum_{j=1}^{2} \hat{q} \log \frac{\hat{q}}{p_j} - \sum_{j=1}^{2} \hat{q} \log \frac{\hat{q}}{p_j^{-i}}] \quad (158)$$

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

$$((1-\hat{q})^2 - (1-q_1^*)(1-q_1^{-i^*}))\frac{z_1}{\lambda_1+\lambda_2} + (\hat{q}^2 - q_1^* q_1^{-i^*})\frac{z_2}{\lambda_1+\lambda_2} > \frac{\lambda_1}{\lambda_1+\lambda_2}(\sum_{j=1}^{2} q_j^* \log \frac{q_j^*}{\frac{1}{2}}$$
$$- \hat{q} \log \frac{\hat{q}}{\frac{1}{2}} - (1-\hat{q}) \log \frac{1-\hat{q}}{\frac{1}{2}}) + \frac{\lambda_2}{\lambda_1+\lambda_2}(\sum_{j=1}^{2} q_j^{-i^*} \log \frac{q_j^{-i^*}}{\frac{1}{2}} - \hat{q} \log \frac{\hat{q}}{\frac{1}{2}} - (1-\hat{q}) \log \frac{1-\hat{q}}{\frac{1}{2}}) \quad (159)$$

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 λ_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 λ_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:

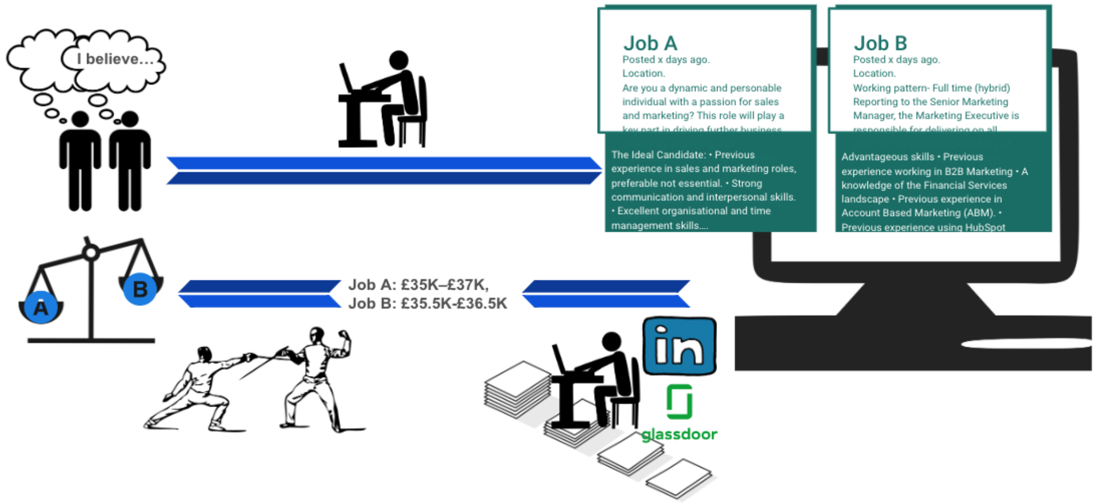


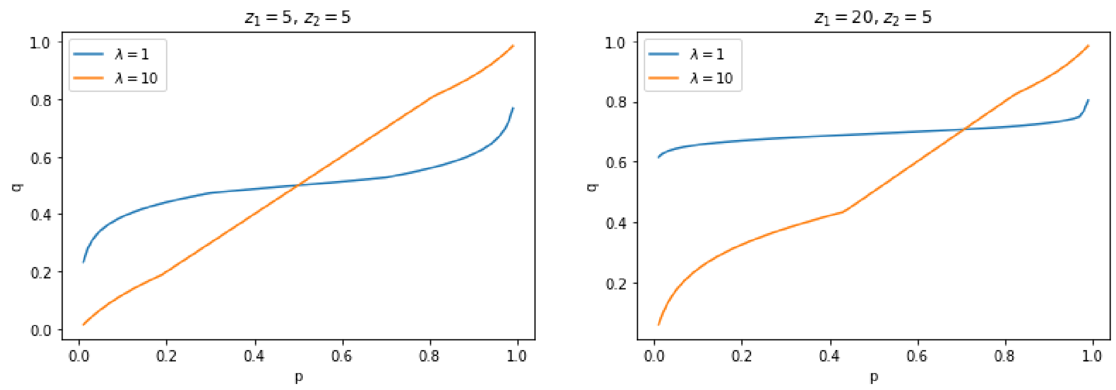
Figure 15: Workers' Decision Process

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 strategies and wages in SSPE, I plot the changes in p on q^* for a given set of exogenous λ, z_1 and z_2.



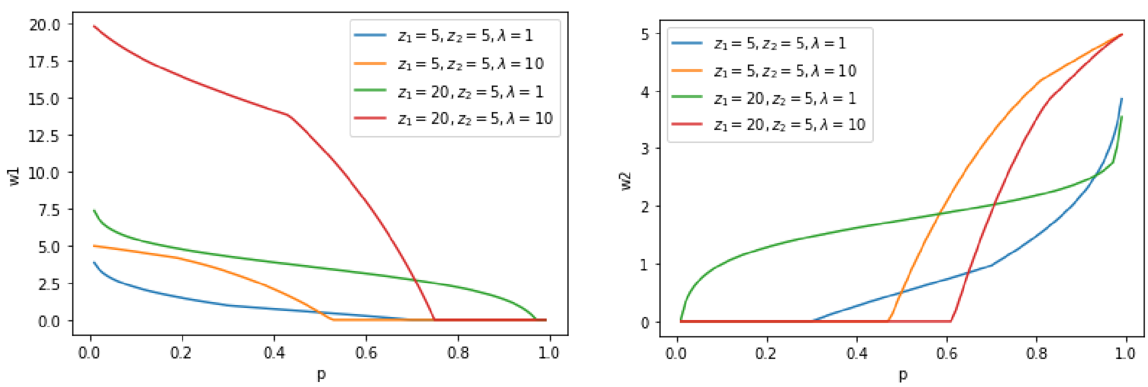
(a) Homogeneous Firms (b) Heterogeneous Firms

Figure 16: Effects of Default Strategies on Equilibrium Application Probabilities

Figure 16 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^*):



(a) Firm 1 (b) Firm 2

Figure 17: Effects of Default Strategies on Equilibrium Wages

Figure 17 shows a clear decreasing trend for w_1^* and increasing trend for w_2^* when p increases, complying with the proof for Proposition 1. The firm expecting lower 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

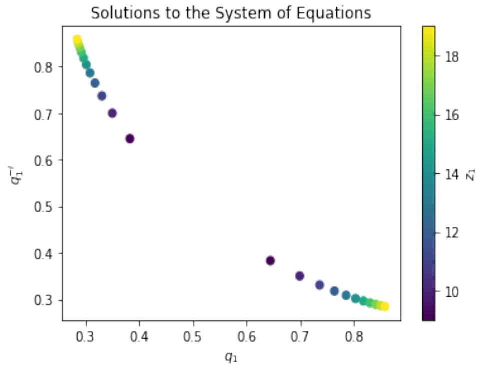
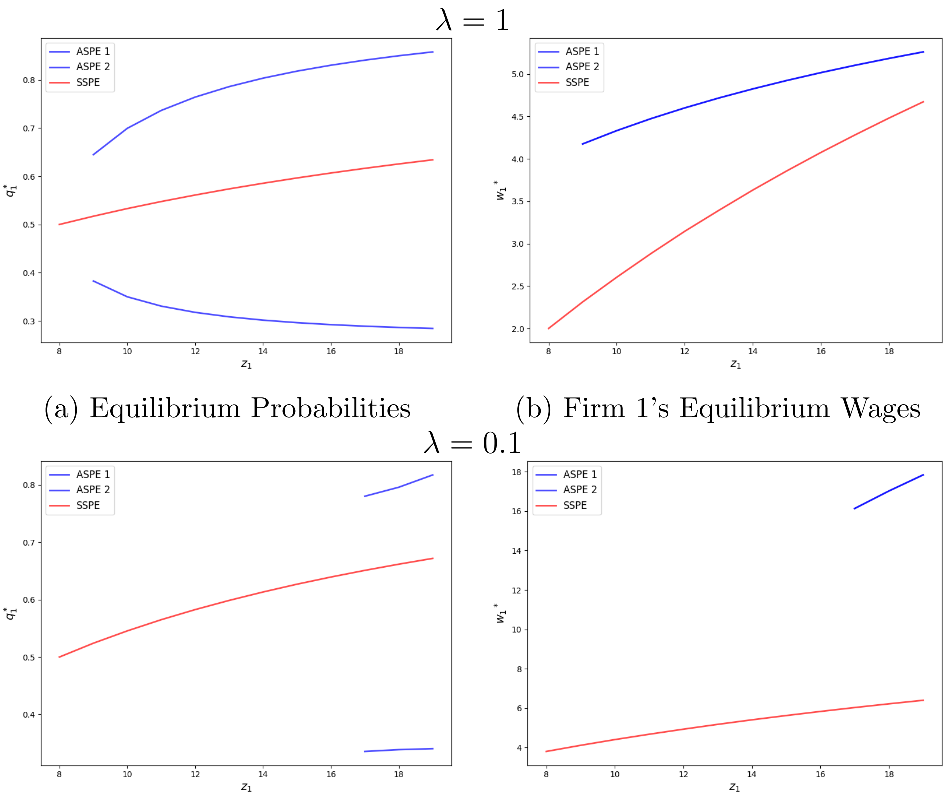


Figure shows ASPE solutions when $\lambda = 1$, $z_1 \in [8, 20]$, $z_2 = 8$, $p_1 = p_1^{-i} = \frac{1}{2}$. The colour spectrum displays a range of z_1 while fixing z_2 to show the impact of heterogeneous productivity on workers' asymmetric equilibrium behaviour.

Figure 18: 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 18 shows numerically that ASPE solutions, if exist, comprise of q_1^{-i*} and q_1^* going in opposite directions.



(a) Equilibrium Probabilities (b) Firm 1's Equilibrium Wages

(c) Equilibrium Probabilities (d) Firm 1's Equilibrium Wages

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

Figure 19: Effects of Varying Productivity on Equilibrium Probabilities and Wages

Equilibrium Solutions with different λ. Figure 19 demonstrates: (1) There could be three possible equilibria, SSPE always exist (red) and ASPE only exist when all necessary conditions are fulfilled (blue) – lowering λ does not necessarily imply existence of ASPE as wage condition may not be fulfilled.[6] (2) The equilibrium wages are lower when λ 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.

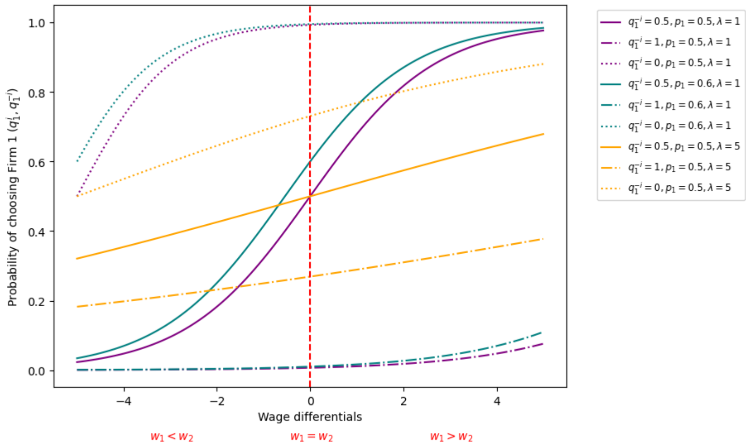


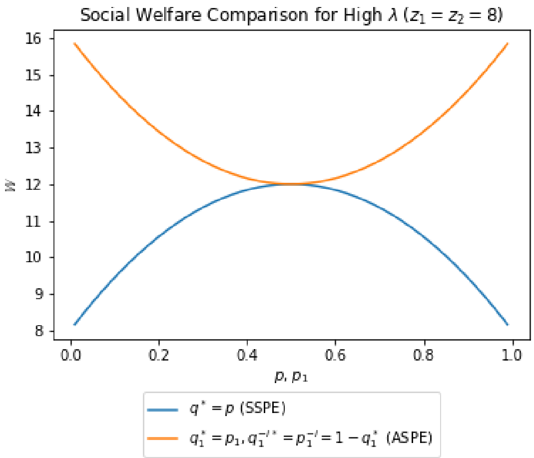
Figure 20: Changes in application probability to firm 1 given changes in wage differentials.

Sensitivity to wage differentials for changes in cognitive costs and defaults. Figure 20 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.

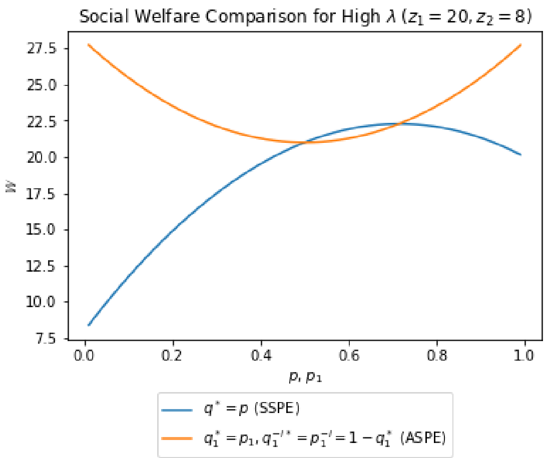
Return to the Section 2.3.2.

[6] To illustrate the example of $\lambda = 0.1$: As λ 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, λ 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).

B.4 Social Welfare Given Homogeneous and Heterogeneous Defaults with Infinitely Costly Search



(a) Homogeneous Productivity



(b) Heterogeneous Productivity

Figure 21: Social Welfare at High Cognitive Costs

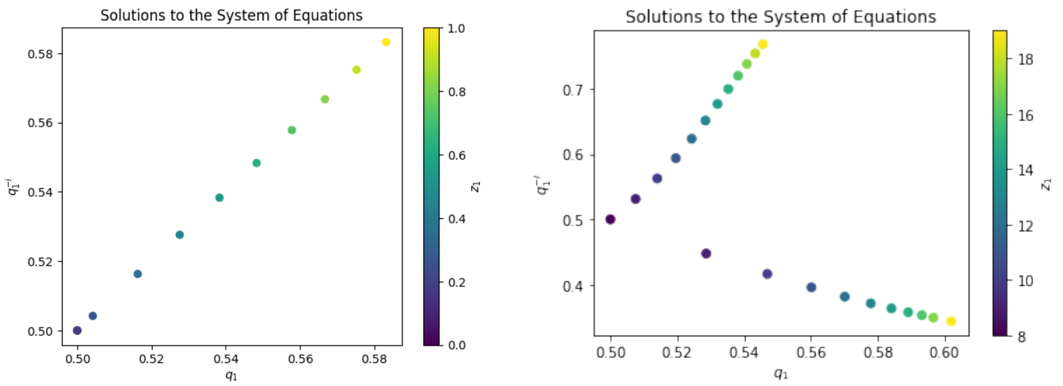
When information is infinitely expensive (i.e. $\lambda \to \infty$), workers resort to their default strategies, social welfare will be determined by the primitive default values. Figure 21 shows that when facing firms of the same productivity, if workers are homogeneous and apply symmetrically in the equilibrium, then welfare is maximized when $q^* = \frac{1}{2}$. If productivities differ, welfare will be maximized at $q^* = \frac{z_1}{z_1+z_2}$. This is where the equilibrium application probability matches the productivity distribution. However, if workers hold different defaults, assuming no relative belief distortion, 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 cost of lower application rate received by the more productive firm is compensated by the gain from higher probability of one-to-one matching.

Return to the Section 3.1.

B.5 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.

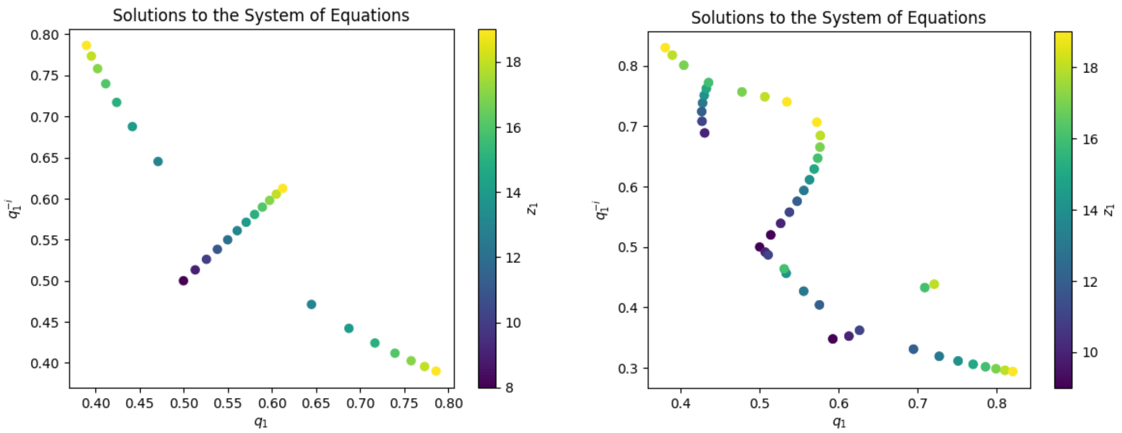


(a) $\lambda_1 = \lambda_2 = 3$ (b) $\lambda_1 = 3$, $\lambda_2 = 1$

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

Figure 22: 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.



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

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

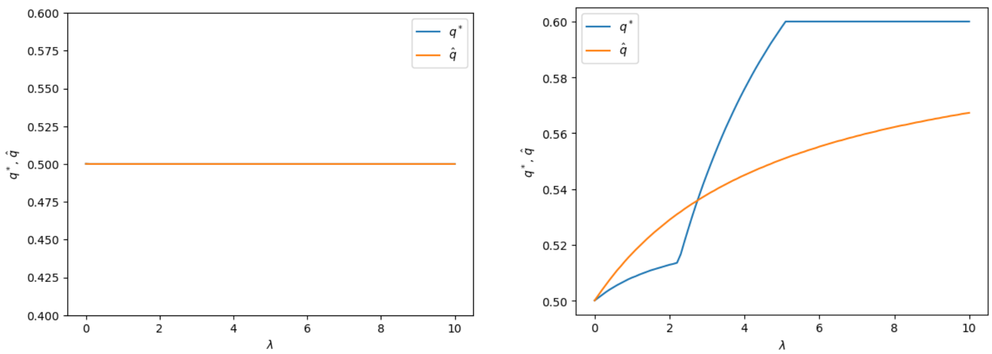
Figure 23: Equilibrium Solutions with Low Cognitive Costs

Return to the Section 3.2.

B.6 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 λs.

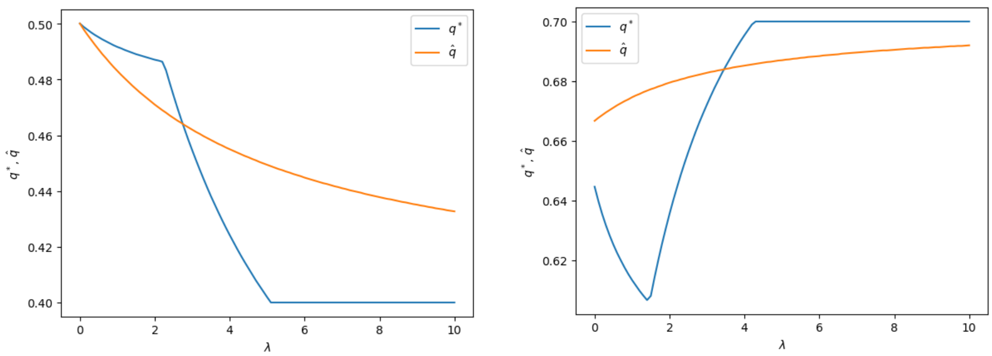
Figure 24a shows for homogeneous firms and correct prior beliefs ($p = \frac{1}{2}$). Figure 24b shows for homogeneous firms incorrect prior beliefs, biased towards firm 1 by default ($p > \frac{1}{2}$), and Figure 25a shows the other case of biased towards firm 2 by default ($p < \frac{1}{2}$). The 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) $p = 0.6$.

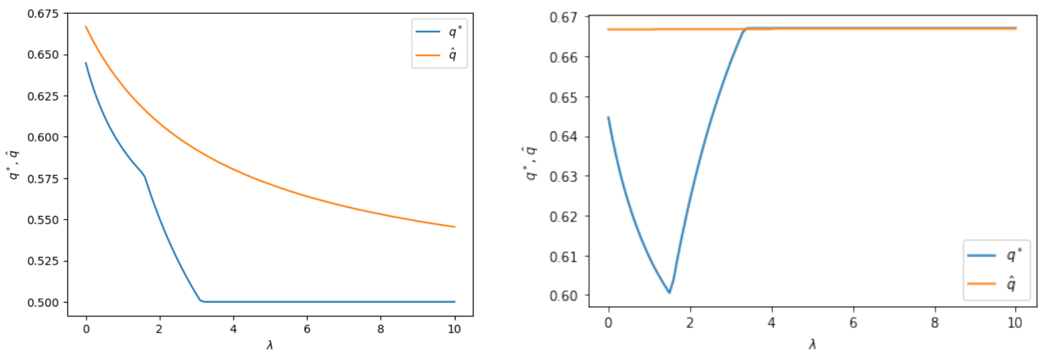
Figure 24: SSPE vs. Socially Optimal Application Probabilities (Homogeneous Firms)



(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 25: 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 26: 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 (72) would provide the socially optimal solution $(\hat{q}_1,\, \hat{q}_1^{-i})$, where $\hat{q}_1$ can differ from $\hat{q}_1^{-i}$:

$$(1-q_1^{-i})z_1 - q_1^{-i}z_2 = \lambda_1 \log(\frac{1-p_1}{p_1}\frac{q_1}{1-q_1}) \tag{160}$$

$$(1-q_1)z_1 - q_1 z_2 = \lambda_2 \log(\frac{1-p_1^{-i}}{p^{-i}}\frac{q_1^{-i}}{1-q_1^{-i}}) \tag{161}$$

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

$$\hat{q}_1 = \frac{1}{1+\frac{1-p_1}{p_1}\exp(-(\frac{(1-\hat{q}_1^{-i})z_1-\hat{q}_1^{-i}z_2}{\lambda_1}))} \tag{162}$$

$$\hat{q}_1^{-i} = \frac{1}{1+\frac{1-p_1^{-i}}{p_1^{-i}}\exp(-(\frac{(1-\hat{q}_1)z_1-\hat{q}_1 z_2}{\lambda_2}))} \tag{163}$$

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 (162), (163), 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,

$$q_1^* = \frac{1}{1+\exp(-\frac{(1-2q^{-i*})w^*}{2\lambda})} \tag{164}$$

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

$$\hat{q}_1 = \frac{1}{1+\exp(-\frac{(1-2\hat{q}^{-i})z}{\lambda})} \tag{165}$$

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 (164) and (165) combine to form,

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

This cannot hold since $z \geq w^*$. As a result, choice probabilities cannot be the same. Using (164) and (165),

$$\hat{q}_1 - q_1^* = \frac{\exp(-\frac{(1-2q^{-i*})w^*}{2\lambda}) - \exp(-\frac{(1-2\hat{q}^{-i})z}{\lambda})}{(1+\exp(-\frac{(1-2\hat{q}^{-i})z}{\lambda}))(1+\exp(-\frac{(1-2q^{-i*})w^*}{2\lambda}))} \tag{166}$$

$\hat{q}_1 - q_1^* > 0$ 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.