Unemployment Duration with Rationally Inattentive Firms

Jeong Ho (John) Kim Kyungmin Kim Marilyn Pease*

January 21, 2021

Abstract

We consider a worker's job search problem in which firms arrive sequentially, observe the worker's unemployment duration, and conduct an interview to learn about her unobservable productivity. Firms engage in fully flexible information acquisition subject to a cost proportional to the expected entropy reduction, as in the literature on rational inattention. We provide a closed-form characterization of equilibrium job search dynamics and demonstrate that endogenous information amplifies the "stigma" effect of long unemployment duration relative to exogenous information. We also show that lowering firms' information-acquisition costs has ambiguous implications for a worker's unemployment duration.

JEL Classification Numbers: D82, D83, J23

Keywords: duration dependence; job search; rational inattention.

*J. H. Kim: Emory University, jkim230@emory.edu; K. Kim: Emory University, kyungmin.kim@emory.edu; Pease: Indiana University, marpease@iu.edu. Kyungmin Kim is supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2020S1A5A2A03043516).

1 Introduction

It is well recognized that there is an inverse relationship between a worker's unemployment duration and her employment probability, namely, that it becomes harder for a worker to find a job, the longer she remains unemployed.[1] One prominent theory for such "negative duration dependence" is based on unobservable worker heterogeneity and firms' private information acquisition.[2] If firms are uncertain about workers' productivity but can interview (test) them, then a worker with lower productivity is less likely to pass the test. In this case, a longer unemployment duration indicates lower productivity, thereby resulting in a lower employment probability.

The goal of this paper is to incorporate flexible information acquisition into the information-based theory and evaluate its effects on unemployment duration. Most existing models assume that firms receive information according to an exogenous process and such information is either free or requires only fixed costs. In other words, firms have no (or little) control over the amount and type of information they receive about the worker, which is inconsistent with the significant resources firms spend doing due diligence to learn about potential employees in reality. In this paper, we endow firms with full flexibility in information acquisition and study how their information-acquisition incentives depend on a worker's unemployment duration, and in turn, influence the equilibrium job search dynamics.

We consider the simplest job search problem with unobservable worker heterogeneity: a worker is searching for a job and sequentially meets firms. The worker may or may not have high productivity, but firms do not directly observe her productivity. They observe her unemployment duration and choose how much information to acquire about her productivity. We accommodate flexible information acquisition by modeling each firm's interview as in the literature on rational inattention: a firm can choose any Blackwell experiment subject to a cost that is proportional to the expected entropy reduction.

We provide a complete characterization of the market equilibrium. In particular, we present closed-form formulae for each firm's optimal interview-and-hiring strategy and the evolution of market belief about the worker's type. This clean characterization owes much to the following two properties of our model: (i) each individual firm faces a binary choice problem (of whether to hire the worker or not), and (ii) information cost is entropy-based. Together, they imply that each firm optimally adopts a binary signal and its action takes a modified logit form, both of which help

[1] Real estate markets exhibit a similar phenomenon, namely, that a property becomes harder to sell as its days on market increase. See, e.g., Taylor (1999), Merlo and Ortalo-Magne (2004), and Tucker, Zhang and Zhu (2013). Our subsequent analysis and discussion can apply equally well to real estate markets.

[2] See Vishwanath (1989) and Lockwood (1991) for seminal theoretical contributions. Other prominent theories are based on depreciating human capital (e.g. Acemoglu, 1995), duration-based rankings (e.g. Blanchard and Diamond, 1994), and endogenous search intensity (e.g. Coles and Smith, 1998; Lentz and Tranaes, 2005).

simplify the law of motion for the market belief.[3]

We show that endogenous information (i.e., flexible information acquisition) amplifies the "stigma" effect, causing a worker's employment probability to drop faster than in the case of exogenous information. With exogenous information, a worker's employment probability falls over time simply because firms raise their hiring standard (i.e., cutoff signal) as the market belief (that the worker has high productivity) decreases. Endogenous information accelerates this process by making the amount of information firms acquire to vary over time. To see this, note that the worker is most likely to be the high type when she just enters the market. Relatively optimistic firms are likely to hire the worker without seeking precise information. As her unemployment duration increases, firms become more cautious, thereby acquiring more information and hiring with smaller probabilities.[4] Once the worker remains unemployed for a sufficiently long period of time, firms are fairly confident that she has low productivity and, therefore, again have little incentive to acquire information. In this case, a firm hires the worker only when it receives an unexpectedly good signal about her. Importantly, this change of optimal information acquisition magnifies the underlying falling-belief process: relative to the comparable exogenous-information case, a worker is more likely to be hired initially but less likely to be hired after some time.[5]

A related economic implication is that endogenous information makes the market belief about a worker's productivity fall more slowly at an early stage of her job search and also after she has stayed unemployed for a sufficiently long period of time. The market belief declines faster when the gap between the two worker types' hiring rates is larger and, therefore, a worker's unemployment duration induces more information acquisition about the worker. As explained above, when a worker's unemployment duration is either very short or very long, firms choose not to acquire much information, reducing informational content of their hiring decisions and, therefore, causing the market belief to move (decrease) slowly.

One novel lesson from our analysis is that reducing firms' information-acquisition costs has ambiguous implications for the worker's unemployment duration. Its effects crucially depend on the initial market belief p_0 ("prior reliance"). More information, driven by lower information-acquisition costs, induces a firm to take a non-default action (i.e., the action that is not optimal under the prior) more frequently. If p_0 is relatively large and so firms are likely to hire the worker, then more information induces firms to reject the worker more often, thereby raising the worker's unemployment duration. Conversely, if p_0 is relatively small, then more information increases the

[3]We caution that these two properties do not guarantee a closed-form characterization. As shown in Section 5, a firm's optimal interview strategy (not to mention the evolution of market belief) cannot be solved in closed form if it faces a fixed information capacity.

[4]Note that we are implicitly assuming that the initial market belief p_0 is large, but not too large. If p_0 is sufficiently large, then all firms hire the worker without acquiring information, in which case the market belief stays equal to p_0.

[5]This is conditional on firms hiring with positive probability. With exogenous information, firms may stop hiring all together once a worker remains unemployed for a sufficiently long time.

probability of hiring, thereby reducing the worker's unemployment duration. Clearly, the overall effect also depends on the worker's type: the high type wants her type to be revealed, while the low type wants the opposite. Therefore, the high type is more likely to benefit from a reduction in information-acquisition costs than the low type. We show that this "information effect" dominates when information is relatively cheap, while the aforementioned prior-reliance effect dominates otherwise.

As discussed above, our work belongs to the literature that explains negative duration dependence based on unobservable worker heterogeneity and firms' information acquisition. In particular, our model is closest to that of Lockwood (1991). The only substantial difference is that firms receive information according to an exogenously given technology in Lockwood (1991), while they have full control (flexibility) over information acquisition in our model.[6]

Several recent papers also incorporate rational inattention into labor market models.[7] For example, Cheremukhin et al. (2020) study to what extent rational inattention regarding whom to meet influences the market outcome in a two-sided matching environment. Wu (2020) applies a similar idea to a directed search environment. Briggs et al. (2017) study the implications of rational inattention on the part of workers in an environment where they can acquire information about non-wage characteristics prior to accepting a job. Bartoš et al. (2016) present a simple static model that illustrates how firms' attention choices can magnify the effects of prior beliefs and preferences. Finally, Acharya and Wee (2020) introduce firms' rational inattention regarding hiring into a labor market model and investigate how it interacts with aggregate productivity shocks.[8]

The remainder of this paper is organized as follows. Section 2 introduces our formal model. Section 3 provides a closed-form characterization of the market equilibrium. Section 4 analyzes how the market outcome (in particular, the worker's unemployment duration) depends on search frictions (measured by the arrival rate of firms) and information frictions (measured by the unit information cost of firms). Section 5 considers an alternative model in which firms' information constraints are given in the form of finite information capacity. Section 6 illustrates how the main lessons from our baseline model apply to two richer environments, each allowing for strategic actions on the part of the worker. Section 7 concludes.

[6]The seminal insight of Vishwanath (1989) and Lockwood (1991) has been utilized in various ways. See, for example, Taylor (1999), Zhu (2012), Palazzo (2017), Kaya and Kim (2018), and Zou (2019).

[7]More broadly, it is an active area of research to investigate the implications of flexible information acquisition in various economic problems. See the recent survey by Mackowiak et al. (2020).

[8]In Acharya and Wee (2020), as in our paper, a firm observes a worker's unemployment duration. Due to the complexity of their full model (featuring aggregate productivity and match-specific shocks and general type spaces), they provide a limited set of analytical results (Section II). Specifically, they provide a characterization for a firm's optimal interview strategy (similar to our Propositions 1 and 4) but do not study its dynamic implications, which are our main focus (see Proposition 2). In addition, they assume that firms face a fixed information processing constraint (i.e., the amount of information each firm can process is fixed). As we illustrate in Section 5, such a model has very different dynamic implications from our model (in which a firm can also choose how much information to acquire).

2 The Model

Physical environment. A worker is searching for a job. Time is continuous, and the worker's unemployment duration is denoted by $t \in \mathcal{R}_+$. The worker may or may not have high productivity. Formally, we use $\omega = L, H$ to denote the worker's type. The worker contributes y_ω to the hiring firm, where $0 \leq y_L < y_H$.[9] It is commonly known that at time 0, the worker has high productivity with probability $p_0 \in (0,1)$. There are infinitely many firms, which arrive sequentially according to a Poisson process with rate $\phi > 0$; that is, the worker faces an infinite sequence of firms that arrive at Poisson rate ϕ. Upon arrival, each firm observes the worker's unemployment duration t, *interviews* the worker, and then decides whether to hire the worker or not at wage w.[10] In order to avoid triviality, we maintain the assumption that $y_L < w < y_H$; that is, each firm wants to hire the worker if and only if she has high productivity (i.e., $\omega = H$). Job search continues until the worker is employed by a firm.

Interview. We model firms' information acquisition as in the literature on rational inattention. Let p denote a firm's prior belief about the worker's type and S denote an arbitrary ordered set. Each firm can choose any joint distribution $G : \{L, H\} \times S \to [0,1]$ that is consistent with prior p. Let $G(\omega, s)$ denote the probability that the realized signal is weakly less than s conditional on state ω. Then, a distribution G is feasible to the firm if and only if its marginal distribution over ω's coincides with p, that is,

$$p = \int_S G(H, s) ds.$$

As is typical in the literature, we assume that the cost of information is proportional to the expected reduction in entropy. Let λ denote the unit cost of information and p_s represent the firm's posterior belief after receiving $s \in S$. Then, the cost associated with G is given by

$$C(G; p) \equiv \lambda \left(H(p) - \mathbb{E}_G[H(p_s)] \right),$$

where $H(x)$ denotes entropy of the binary distribution $(1 - x, x)$, that is,

$$H(x) \equiv -(1 - x) \log(1 - x) - x \log x.$$

[9] The job can be either temporary or permanent. In the former case, y_ω represents a one-time contribution to the firm, while in the latter case, y_ω should be interpreted as the worker's life-time expected contribution. The same applies to wage w that will be introduced shortly.

[10] We assume that the market wage w is exogenously given. However, it can be micro-founded as follows: the worker's reservation utility is equal to w, irrespective of her type, and each firm makes a take-it-or-leave-it wage offer. Under this specification, by the usual Diamond-paradox logic, no firm offers strictly more than w and the worker gets hired only at wage w. In Section 6, we consider an extension of this micro-founded version in which the worker's reservation utility depends on her type.

Payoffs. Suppose that the worker is of type $\omega = L, H$ and hired at time t by a firm who chooses G with prior belief $p(t)$. Then, the worker's payoff is equal to $e^{-rt}w$, where $r > 0$ denotes her discount rate, and the firm's payoff (profit) is equal to $y_\omega - w - C(G; p(t))$. All other firms (that arrive before t) incur only information costs, as modeled above.

Equilibrium. We study weak perfect Bayesian equilibria of this dynamic game. Note that a firm's optimal hiring strategy is straightforward: it hires the worker if and only if her expected contribution $(1 - p_s)y_L + p_s y_H$ exceeds wage w. Since the posterior p_s is fully determined by the firm's prior p and interview strategy G, we focus on the latter two. Let G_t denote the interview strategy of the firm at t and $p(t)$ denote the prior belief that the worker is of type H. An equilibrium requires that for each t, (i) G_t is an optimal interview strategy given $p(t)$, and (ii) $p(t)$ is consistent with all G_s's such that $s < t$.

3 Equilibrium Characterization

In this section, we fully characterize equilibrium job search dynamics. We first analyze an individual firm's optimal interview-and-hiring problem. Then, we establish full equilibrium dynamics by intertemporally linking firms' optimal decisions.

3.1 Optimal Interview and Hiring

Consider a firm that is facing a worker and assigns probability $p \in (0, 1)$ to the event that the worker is of type H. Given p, the firm decides how much (and what) information to acquire about the worker and, depending on the interview outcome, whether or not to hire her. The latter (hiring) part of the firm's problem is straightforward. Let $\widehat{p}$ denote the probability of type H that makes the firm indifferent between hiring and nor hiring the worker; that is,

$$\widehat{p}(y_H - w) + (1 - \widehat{p})(y_L - w) = 0 \Leftrightarrow \widehat{p} = \frac{w - y_L}{y_H - y_L}.$$

The firm hires the worker if and only if its posterior after interview exceeds $\widehat{p}$.

The firm's optimal interview strategy is rather intricate to characterize, because the firm has full flexibility in its interview design and so can choose any distribution G that is consistent with p. We solve this problem by applying a recent technical insight in the literature on information design (acquisition), namely that the entropy-based cost function is *posterior separable*, and thus a solution to an information-acquisition problem can be found through *concavification* (see Gentzkow

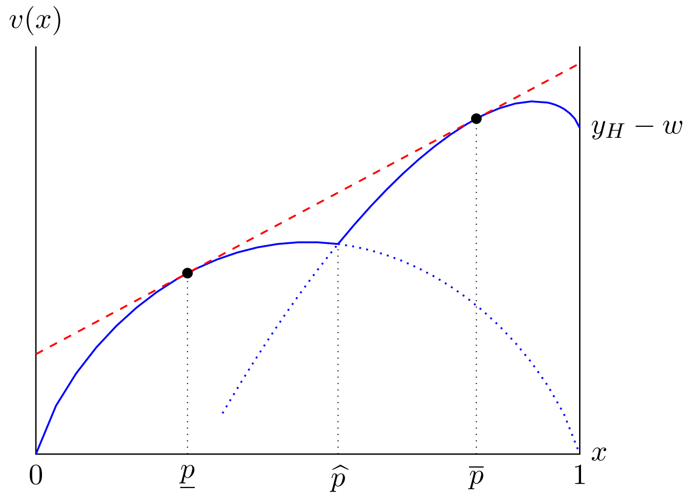

Figure 1: Solving the firm's optimal interview strategy through concavification. The parameter values used for this figure are $y_L = 1$, $w = 2$, $y_H = 3$, and $\lambda = 0.75$.

and Kamenica, 2014; Caplin et al., 2019).[11]

Suppose that the firm chooses a distribution $G(\omega, s)$. This induces a distribution $\tilde{G}$ over posteriors, where $\tilde{G}(x)$ denotes the probability that the firm's posterior is weakly below x. Let $\pi(x)$ denote the firm's expected profit when its posterior is equal to x; that is, $\pi(x) \equiv \max\{0, xy_H + (1-x)y_L - w\}$. Then, the firm's expected payoff from G can be written as

$$\begin{aligned} V(G;p) &= \int \pi(x) d\tilde{G}(q) - \lambda \left(H(p) - \int H(x) d\tilde{G}(x) \right) \\ &= \int \left(\pi(x) + \lambda H(x) \right) d\tilde{G}(x) - \lambda H(p). \end{aligned}$$

This implies that, excluding the constant term $-\lambda H(p)$, the firm's optimal interview problem can be written as[12]

$$\max_{\tilde{G} \in \Delta([0,1])} \int v(x) d\tilde{G}(x) \text{ subject to } \int x d\tilde{G}(x) = p,$$

where $v(x) \equiv \pi(x) + \lambda H(x)$. As is well-known, this problem can be solved through concavification (Aumann and Maschler, 1995; Kamenica and Gentzkow, 2011).

Figure 1 illustrates how concavification works in the current environment. The solid blue curve represents the value function $v(x)$. Its strict concavity below $\underline{p}$ or above $\underline{p}$ is due to the strict concavity of the entropy function $H(x) = -(1-x)\log(1-x) - x\log(x)$. The kink at $\widehat{p}$ arises

[11] Alternatively, one can utilize the fact that the firm faces a binary choice problem, and thus a general discrete choice characterization by Matějka and McKay (2015) applies. An important advantage of our concavification-based approach is that the same analysis applies unchanged to a broad class of information cost functions: as shown shortly, our analysis relies only on posterior separability of the entropy-based cost function and strict concavity of $H(\cdot)$.

[12] As is standard in the literature, $\Delta([0,1])$ denotes the set of all probability distributions over $[0,1]$.

because the firm hires the worker if and only if $x \geq \widehat{p}$. Given the particular shape of $v(p)$, it is straightforward that there are only two possibilities: If $p \leq \underline{p}$ or $p \geq \overline{p}$ then the firm has no incentive to split p into multiple posteriors and, therefore, acquires no information. If $p \in (\underline{p}, \overline{p})$ then it is optimal for the firm to split p into two posteriors, $\underline{p}$ and $\overline{p}$.

Certainly, the optimal solution when $p \in (\underline{p}, \overline{p})$ can be implemented through a binary signal: let $S = \{\ell, h\}$ denote the set of signal realizations, and q_ω denote the probability that $s = h$ conditional on $\omega = L, H$ (i.e., $q_\omega \equiv Pr\{s = h|\omega\}$ for $\omega = L, H$). Then, it suffices to set q_L and q_H so that

$$\frac{\underline{p}}{1-\underline{p}} = \frac{p}{1-p}\frac{1-q_H}{1-q_L} \text{ and } \frac{\overline{p}}{1-\overline{p}} = \frac{p}{1-p}\frac{q_H}{q_L}. \tag{1}$$

The values of $\underline{p}$ and $\overline{p}$ can be explicitly found as follows: as shown in Figure 1, they must have the same tangent. Therefore,

$$v'(\underline{p}) = \lambda H'(\underline{p}) = v'(\overline{p}) = \pi'(\overline{p}) + \lambda H'(\overline{p})$$

and

$$v'(\underline{p}) = \frac{v(\overline{p}) - v(\underline{p})}{\overline{p} - \underline{p}} = \frac{\pi(\overline{p}) + \lambda H(\overline{p}) - \lambda H(\underline{p})}{\overline{p} - \underline{p}}.$$

In other words, $v(x)$ must have the same slope at $\underline{p}$ and at $\overline{p}$ (the first equation), and $(\overline{p}, v(\overline{p}))$ must lie on the tangent to $v(x)$ at $\underline{p}$ (the second equation). Arranging the terms and combining the result with the firm's optimal hiring strategy, we obtain the following result.

Proposition 1 *Let*

$$\underline{p} \equiv \frac{1 - e^{(y_L - w)/\lambda}}{e^{(y_H - w)/\lambda} - e^{(y_L - w)/\lambda}} \in (0, 1) \text{ and } \overline{p} \equiv \frac{e^{(y_H - w)/\lambda}(1 - e^{(y_L - w)/\lambda})}{e^{(y_H - w)/\lambda} - e^{(y_L - w)/\lambda}} \in (\underline{p}, 1).$$

For each $p \in [0, 1]$, the firm's optimal interview-and-hiring strategy is as follows:

(i) If $p \leq \underline{p}$, then the firm does not acquire any information and does not hire the worker.

(ii) If $p \geq \overline{p}$, then the firm hires the worker without acquiring information.

(iii) If $p \in (\underline{p}, \overline{p})$, then the firm chooses a binary signal so that its posterior is equal to either $\underline{p}$ or $\overline{p}$. The firm hires the worker if and only if its posterior becomes $\overline{p}$.

It is intuitive that the firm acquires information if and only if p is neither too low nor too high: if p is rather extreme, then it is too costly for the firm to collect sufficiently strong evidence to deviate from its default action. It is also natural that, since information is costly, the firm acquires just enough information to make an efficient hiring decision. In particular, it is wasteful for the

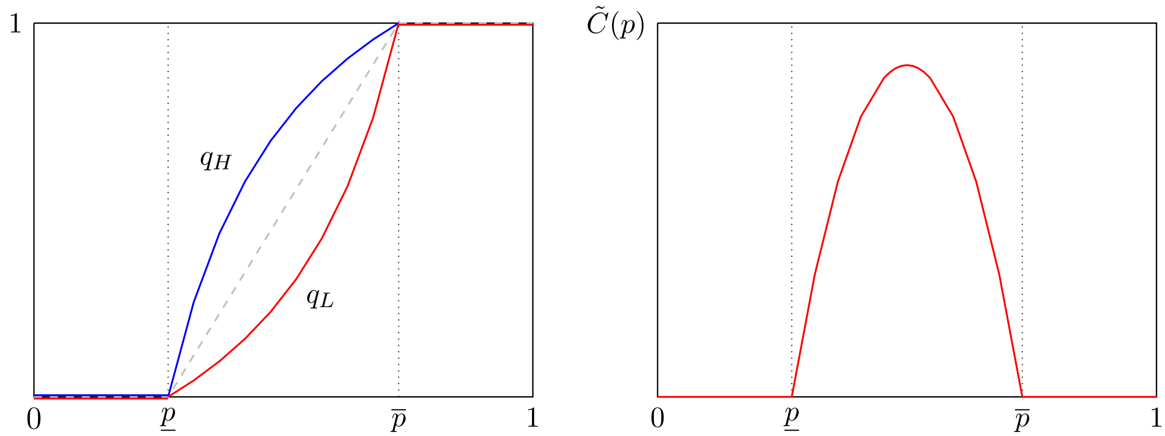

Figure 2: Optimal interview-and-hiring strategy as a function of prior p. The left panel depicts $q_H(p)$ (blue) and $q_L(p)$ (red), while the right panel shows the indirect total cost $\tilde{C}(p)$.

firm to produce multiple signals that lead to the same action, and thus the firm always adopts a binary signal.

Figure 2 illustrates how the firm's optimal interview strategy depends on its prior belief p. The left panel shows the optimal levels of q_H and q_L.[13] Clearly, q_H is uniformly above q_L, which simply means that h is always a good signal about the worker's type. Importantly, the two worker types are hired with similar probabilities near $\overline{p}$ and $\underline{p}$, while the probabilities are furthest apart when p is in the intermediate range. This is intuitive because, as usual, the firm has a stronger incentive to acquire information when there is more uncertainty. The right panel depicts the total attention cost that corresponds to the optimal interview strategy. It reaffirms that the firm indeed invests more into information acquisition, the more uncertain it is about the worker's type.

3.2 Equilibrium Dynamics

Proposition 1 presents an individual firm's optimal interview-and-hiring strategy as a function of its prior belief p. In our dynamic model where firms arrive sequentially, this belief evolves over time, depending on previous firms' (expected) behavior. We now derive the belief $p(t)$ at time t and illustrate the resulting equilibrium dynamics.

Notice that the job-finding rate of each worker type at time t is equal to $\phi q_\omega(t)$: firms arrive at rate ϕ and hire the type-ω worker with probability $q_\omega(t)$. This implies that the probability that the

[13]The gray dashed curve represents the unconditional probability of hiring, which can be easily shown to be linear in p: for any $p \in (\underline{p}, \overline{p})$,

$$q = pq_H + (1-p)q_L = \frac{p - \underline{p}}{\overline{p} - \underline{p}}.$$

type-ω worker stays unemployed until t is equal to $e^{-\int_0^t \phi q_\omega(x)dx}$, and thus

$$p(t) = \frac{p_0 e^{-\int_0^t \phi q_H(x)dx}}{p_0 e^{-\int_0^t \phi q_H(x)dx} + (1-p_0)e^{-\int_0^t \phi q_L(x)dx}}.$$

Equivalently, $p(t)$ evolves according to

$$\dot{p}(t) = -p(t)(1-p(t))\phi(q_H(t) - q_L(t)). \tag{2}$$

By equation (1), both q_H and q_L can be written as functions of p. Therefore, the above differential equation can be reduced to an ordinary differential equation of $p(t)$ and solved explicitly, as formally reported in the following result.

Proposition 2 *Suppose that* $p_0 \in (\underline{p}, \overline{p})$.[14] *In equilibrium,* $p(t)$ *evolves according to*

$$\dot{p}(t) = -\phi\frac{(p(t)-\underline{p})(\overline{p}-p(t))}{\overline{p}-\underline{p}} \Leftrightarrow p(t) = \frac{e^{-\phi t}(p_0-\underline{p})\overline{p} + (\overline{p}-p_0)\underline{p}}{e^{-\phi t}(p_0-\underline{p}) + (\overline{p}-p_0)}. \tag{3}$$

Proof. Solving the system of equations in (1), we get

$$q_L = \frac{\frac{p}{1-p} - \frac{\underline{p}}{1-\underline{p}}}{\frac{\overline{p}}{1-\overline{p}} - \frac{\underline{p}}{1-\underline{p}}} \text{ and } q_H = \frac{\frac{\overline{p}}{1-\overline{p}}\frac{1-p}{p}\left(\frac{p}{1-p} - \frac{\underline{p}}{1-\underline{p}}\right)}{\frac{\overline{p}}{1-\overline{p}} - \frac{\underline{p}}{1-\underline{p}}} = \frac{1 - \frac{\underline{p}}{1-\underline{p}}\frac{1-p}{p}}{1 - \frac{\underline{p}}{1-\underline{p}}\frac{1-\overline{p}}{\overline{p}}}.$$

Plugging this into the differential equation (2) and arranging the terms lead to equation (3):

$$\dot{p}(t) = -p(t)(1-p(t))\phi\frac{\left(\frac{\overline{p}}{1-\overline{p}}\frac{1-p(t)}{p(t)} - 1\right)\left(\frac{p(t)}{1-p(t)} - \frac{\underline{p}}{1-\underline{p}}\right)}{\frac{\overline{p}}{1-\overline{p}} - \frac{\underline{p}}{1-\underline{p}}} = -\phi\frac{(p(t)-\underline{p})(\overline{p}-p(t))}{\overline{p}-\underline{p}}.$$

■

The left panel of Figure 3 shows a typical path of $p(t)$ in our model (red solid). Provided that p_0 is rather close to $\overline{p}$, $p(t)$ initially decreases fairly slowly: this is when upon meeting a firm, both worker types are hired with probability close to 1. In this case, the worker's unemployment duration is more indicative of the possibility that the worker has not yet met a firm than that she is of the low type, and thus $p(t)$ does not move much. After some time, $p(t)$ decreases rather quickly: this is when firms acquire most information and, therefore, hire the high type with significantly higher probabilities than the low type. In this case, remaining on the market becomes a stronger signal that the worker is of the low type, causing $p(t)$ to drop quickly. If $p(t)$ gets near $\underline{p}$, the fall of

[14] By Proposition 1, $p(t)$ stays constant if $p_0 \leq \underline{p}$ or $p_0 \geq \overline{p}$.

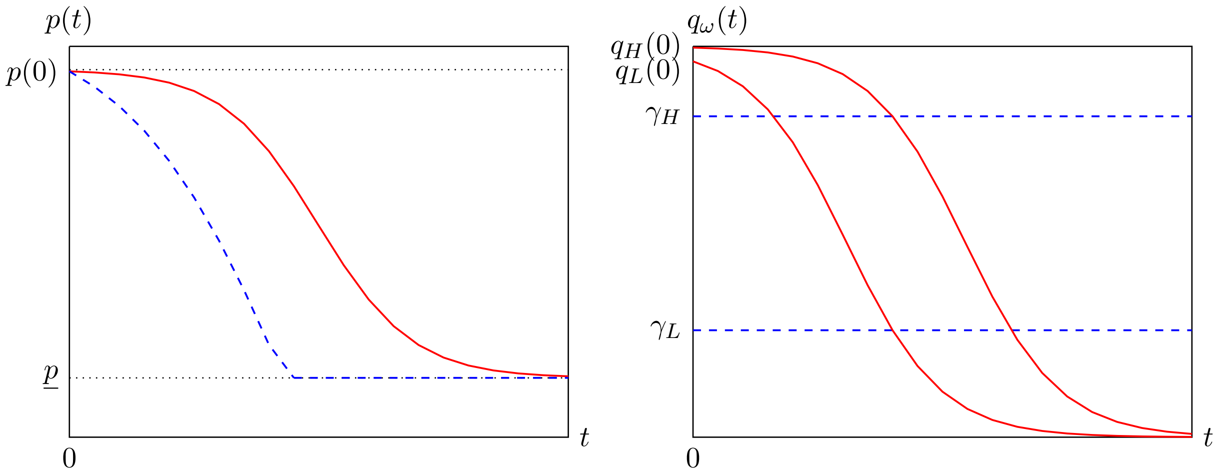

Figure 3: The red solid curves correspond to the equilibrium objects of our endogenous-information model, while the dashed blue curves represent those of the exogenous-information model. In the left panel, the upper dotted line indicates $\overline{p}$. The parameter values used for both panels are $y_L = 1$, $y_H = 2.25$, $w = 2$, $\phi = 0.3$, $\lambda = 0.5$, and $p_0 = 0.94$.

$p(t)$ slows down again: firms are so pessimistic about the worker's type that they rarely hire even the high-type worker. This last effect becomes stronger as $p(t)$ approaches $\underline{p}$. Therefore, although $p(t)$ constantly decreases, it never reaches $\underline{p}$ and converges to it asymptotically.

In order to highlight the effects of optimal (flexible) interview on equilibrium dynamics, consider the following exogenous-information case: each firm receives either ℓ or h, just as in its optimal interview strategy in Proposition 1. However, the probability that $s = h$ conditional on $\omega = L, H$ is exogenously fixed at γ_ω. Furthermore, γ_H and γ_L are such that

$$\overline{p}(1-\gamma_H)(y_H - w) + (1-\overline{p})(1-\gamma_L)(y_L - w) = 0 \text{ and } \underline{p}\gamma_H(y_H - w) + (1-\underline{p})\gamma_L(y_L - w) = 0.$$

In other words, the given binary signal makes an individual firm break even with hiring a worker, both when it has prior belief $\overline{p}$ but receives a bad signal ℓ (the first equation) and when it has prior belief $\underline{p}$ but receives a good signal h (the second equation). These imply that, just as in our main model, a firm always hires the worker if $p \geq \overline{p}$, hires only with a good signal if $p \in (\underline{p}, \overline{p})$, and never hires if $p \leq \underline{p}$. It is also straightforward to show that if $p_0 \in (\underline{p}, \overline{p})$ then, until $p(t)$ reaches $\underline{p}$, $p(t)$ continuously decreases according to

$$\dot{p}(t) = -p(t)(1-p(t))\phi(\gamma_H - \gamma_L) \Leftrightarrow p(t) = \frac{p_0 e^{-\phi(\gamma_H - \gamma_L)t}}{p_0 e^{-\phi(\gamma_H - \gamma_L)t} + (1-p_0)}. \tag{4}$$

The left panel of Figure 3 shows how this evolution of $p(t)$ (blue dashed) compares with that of our main model (red solid).

As is clear from equations (2) and (4), the main difference between endogenous information and exogenous information is that the difference between the two types' hiring rates varies over time in the former ($q_H(t)-q_L(t)$) but stays constant in the latter ($\gamma_H-\gamma_L$). As shown in Section 3.1, $q_H(t) - q_L(t)$ is small when $p(t)$ is near $\underline{p}$ or $\overline{p}$ (see the left panel of Figure 2) and $p(t)$ constantly decreases in both cases. Therefore, $q_H(t) - q_L(t) > \gamma_H - \gamma_L$ only when $p(t)$ belongs to an intermediate range (see the right panel of Figure 3). In other words, $p(t)$ decreases faster with exogenous information if $p(t)$ is close to $\overline{p}$ or $\underline{p}$, but the opposite holds if $p(t)$ is far away from both $\underline{p}$ and $\overline{p}$.

4 Comparative Statics

We now analyze how equilibrium dynamics and the worker's unemployment duration (and welfare) depend on market characteristics. In particular, we evaluate the effects of search frictions (ϕ) and information cost (λ).[15] We begin by illustrating their effects on overall equilibrium dynamics, which are summarized in the following lemma.

Lemma 1 *The two belief cutoffs, $\overline{p}$ and $\underline{p}$, are independent of ϕ. If λ increases, then $\overline{p}$ falls, while $\underline{p}$ increases. In addition,*

$$\lim_{\lambda\to\infty}\overline{p} = \lim_{\lambda\to\infty}\underline{p} = \widehat{p},\text{ while } \lim_{\lambda\to 0}\underline{p} = 0 \text{ and } \lim_{\lambda\to 0}\overline{p} = 1.$$

At each (fixed) $t \in \mathcal{R}_+$, the belief $p(t)$ decreases in ϕ, but increases in λ.

As formally shown in Section 3.1, the two belief cutoffs, $\overline{p}$ and $\underline{p}$, are determined by an individual firm's information-acquisition incentive. Therefore, they are independent of ϕ, which has no impact on a firm's *static* incentive. On the contrary, an increase in λ directly reduces a firm's information-acquisition incentive, thereby making the firm's hiring more prior-driven. In our model, this translates into the information-acquisition region $[\underline{p}, \overline{p}]$ shrinking as λ rises (see the right panel of Figure 4 and compare $\overline{p}$ and $\underline{p}$ to $\overline{p}'$ and $\underline{p}'$, respectively). The asymptotic results in Lemma 1 are also intuitive: if information is too costly to acquire, then the firm's hiring depends only on its prior. In the opposite case in which information is almost free, the firm always seeks to acquire further information and, therefore, never conditions its hiring decision solely on its prior.

Lemma 1 also argues that for each t, $p(t)$ decreases in ϕ and increases in λ. As depicted in

[15] Another natural question is how raising the market wage w affects the worker's overall welfare. It clearly helps the worker conditional on being hired, but reduces her unconditional employment probability, lengthening her unemployment duration. Using our closed-form formulae, one can show that the result is ambiguous in general and mainly depends on the prior belief p_0. Specifically, although the two worker types have different cutoffs, both types benefit from an increase in w if p_0 is sufficiently large, while the opposite holds if p_0 is relatively small (close to $\underline{p}$).

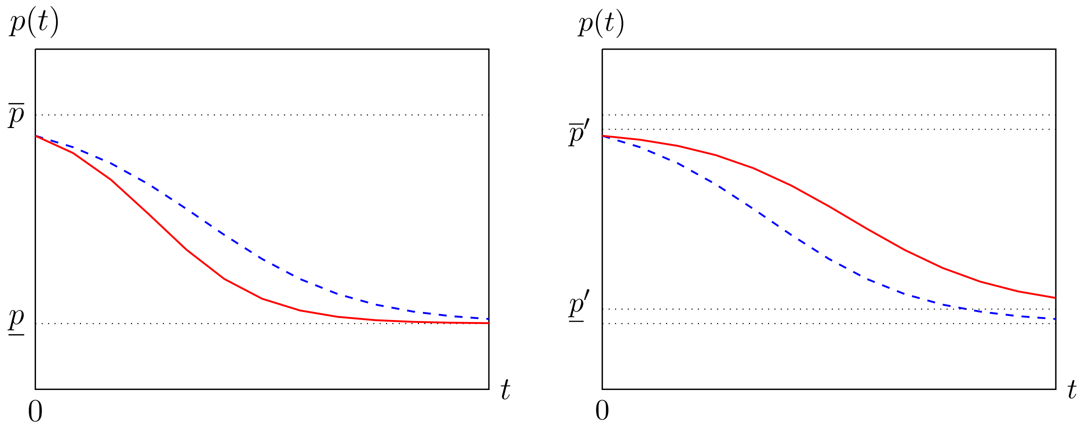

Figure 4: In both panels, the blue dashed curve depicts the original belief path (with $\phi = 0.5$ and $\lambda = 0.7$), while the red solid curve represents the belief path when either ϕ increases (to $\phi' = 0.7$ in the left panel) or λ increases (to $\lambda' = 0.85$ in the right panel). The other parameter values used for both panels are $w = 2$, $y_L = 1$, $y_H = 3$, and $p_0 = 0.7453$.

Figure 4, this means that the whole belief path $p(\cdot)$ shifts to the left if ϕ increases (left panel), while it shifts to the right if λ increases (right panel). To understand these results, recall that $p(t)$ declines over time because the low type is less likely to be hired by any given firm and so increasingly more likely to remain in the market than the high type. An increase in ϕ means that there are relatively more firms (that have interviewed but chosen not to hire the worker) by each time t, and thus $p(t)$ falls faster. On the contrary, an increase in λ induces each firm to acquire less information and so distinguish less between the two worker types, which makes $p(t)$ fall more slowly.

In order to systematically analyze the effects of ϕ and λ on unemployment duration, let $F_\omega(t)$ denote the probability that the type-ω worker becomes employed by time t. Since the conditional employment rate at time t is equal to $\phi q_\omega(t)$, $F_\omega(t)$ is given by

$$F_\omega(t) = 1 - e^{-\int_0^t \phi q_\omega(x)dx} \text{ for all } t \in \mathcal{R}_+ \text{ and } \omega = L, H.$$

This distribution is trivial if $p_0 \leq \underline{p}$ (in which case the worker never becomes employed) or $p_0 \geq \overline{p}$ (in which case the worker gets hired as soon as the first firm arrives). Even when $p_0 \in (\underline{p}, \overline{p})$, the specific structure of $q(t)$ enables us to derive this distribution in closed form, as formally reported in the following proposition.

Proposition 3 *Suppose that $p_0 \in (\underline{p}, \overline{p})$. Then, for each $\omega = L, H$, $F_\omega(t)$ is given by*

$$F_\omega(t) = \frac{1 - e^{-\phi t}}{1 + A_\omega},$$

where

$$A_L = \frac{(1-\underline{p})}{(1-\overline{p})}\frac{(\overline{p}-p_0)}{(p_0-\underline{p})} \textit{ and } A_H = \frac{\underline{p}}{\overline{p}}\frac{(\overline{p}-p_0)}{(p_0-\underline{p})}.$$

Proof. From equation (1), we get

$$q_L(t) = \frac{1-\overline{p}}{\overline{p}-\underline{p}}\frac{p(t)-\underline{p}}{1-p(t)} \text{ and } q_H(t) = \frac{\overline{p}}{\overline{p}-\underline{p}}\frac{p(t)-\underline{p}}{p(t)}.$$

Combining these with $p(t)$ in Proposition (2) leads to

$$q_\omega(t) = \frac{1}{1+e^{\phi t}A_\omega} \text{ for each } \omega = L, H.$$

Note that A_ω is independent of t. Then, by direct integration, we get

$$\begin{aligned} F_\omega(t) &= 1 - exp\left(-\int_0^t \phi q_\omega(x)dx\right) = 1 - exp\left(-\int_0^t \frac{\phi}{1+e^{\phi x}A_\omega}dx\right) \\ &= 1 - exp\left(-\phi t + \log(1+e^{\phi t}A_\omega) - \log(1+A_\omega)\right) \\ &= 1 - e^{-\phi t}\frac{1+e^{\phi t}A_\omega}{1+A_\omega} = \frac{(1+A_\omega)-(e^{-\phi t}+A_\omega)}{1+A_\omega} = \frac{1-e^{\phi t}}{1+A_\omega}. \end{aligned}$$

■

The strikingly clean solution of F_ω in Proposition 3 enables us to analyze the effects of ϕ and λ on unemployment duration in a simple fashion.

Corollary 1 *For both $\omega = L, H$, F_ω stochastically (in the sense of first-order stochastic dominance) decreases in ϕ.*

Proof. This result follows from the fact that for all t, $F_\omega(t)$ increases in ϕ. ■

Although the result is clear and simple to establish, there are two opposing forces. On the one hand, an increase in ϕ directly raises each worker type's employment rate $\phi q_\omega(t)$ at each time t. On the other hand, as shown in Lemma 1 (and the left panel of Figure 4), it makes $p(t)$ fall faster, which lowers $q_\omega(t)$ at each t. The existence of this latter effect makes $\phi q_\omega(t)$ rotate rather than shift, as shown in the left panel of Figure 5: if ϕ increases, then $\phi q_\omega(t)$ increases for lower values of t but decreases for higher values of t. Nevertheless, the former direct effect always outweighs the latter indirect effect, in the sense that $\int_0^t \phi q_\omega(x)dx$ increases in ϕ (see the right panel). Therefore, the distribution $F_\omega(t)$ stochastically decreases. We conclude that an increase in ϕ always shortens the worker's unemployment duration and improves her welfare.

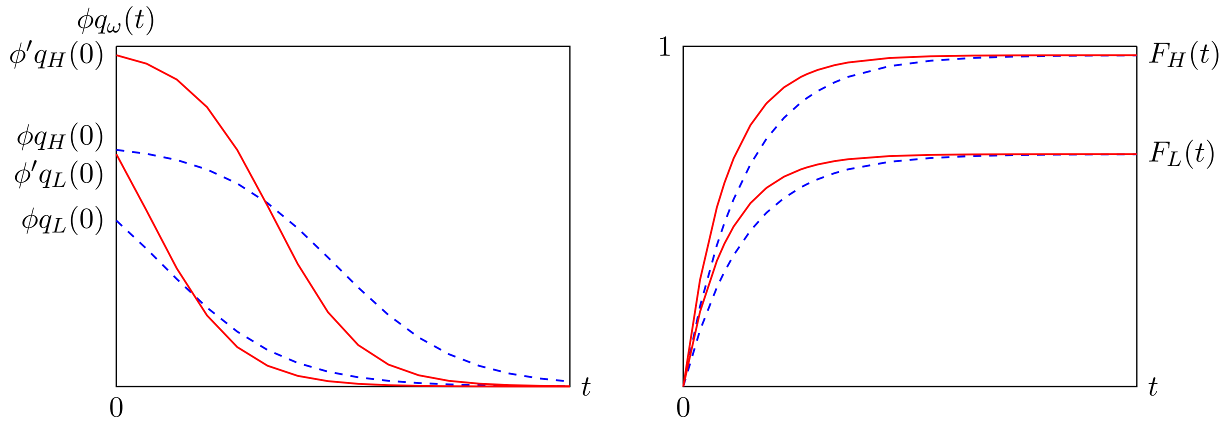

Figure 5: In both panels, the blue dashed curves represent $\phi = 0.5$, while the red solid curves correspond to $\phi' = 0.7$. The other parameter values used for this figure are identical to those of Figure 4.

Corollary 2 *If $p_0 \in [\widehat{p}, \overline{p})$, then F_L always (stochastically) decreases in λ, while F_H decreases in λ if and only if $\lambda \geq \lambda_H(p_0)$ for some $\lambda_H(p_0) \in (0, \infty)$. If $p_0 \in (\underline{p}, \widehat{p})$, then F_L decreases in λ if and only if $\lambda \leq \lambda_L(p_0)$ for some $\lambda_L(p_0) \in (0, \infty)$, while F_H always increases in λ.*

Proof. Proposition 2 implies that $F_\omega(t)$ moves in the opposite direction to A_ω for all t. Therefore, F_ω stochastically rises if and only if A_ω increases. In Appendix A, we show that A_ω moves in the way to support the pattern stated in the result. ■

Figure 6 visualizes Corollary 2, by showing the areas in which F_ω stochastically rises or falls as λ increases. Notice that there are effectively three cases: (i) both F_L and F_H stochastically decrease if λ is sufficiently large and $p_0 > \widehat{p}$; (ii) both stochastically increase if λ is sufficiently large and $p_0 < \widehat{p}$; (iii) F_L rises, while F_H falls, if λ is sufficiently small, regardless of p_0.

In order to understand this pattern, notice that an increase in λ has two distinct effects, prior reliance and the information effect. The former refers to the fact that an increase in λ (the price of information) induces firms to acquire less information and, therefore, rely more on their prior beliefs. If a firm's prior is relatively optimistic, then this is good news to the worker, regardless of her type, because putting more weight on the optimistic prior on the part of the firm implies a higher probability of hiring. Of course, it is bad news in the opposite scenario when a firm has a relatively pessimistic prior and thus its default action is not to hire. This explains why $\widehat{p}$ serves as an important cutoff in Corollary 2: with no information acquisition, a firm hires a worker if and only if $p(t) \geq \widehat{p}$.

An increase in λ has an additional effect, which depends on the worker's type. When information becomes more expensive, and the firm obtains less of it, the equilibrium signal becomes less informative about the worker's type. This benefits the low-type worker and harms the high-type

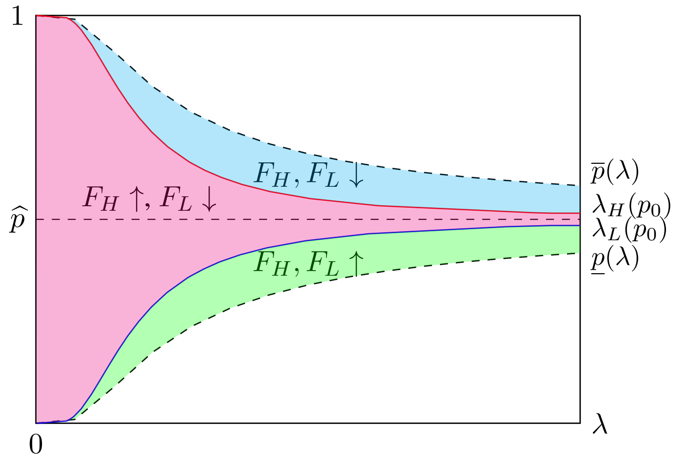

Figure 6: The blue curve represents $\lambda_L(p_0)$, while the red curve exemplifies $\lambda_H(p_0)$, in Corollary 2. In the figure, $\uparrow$ ($\downarrow$) means that F_ω rises (falls) in the sense of first-order stochastic dominance when λ increases. The parameter values used for this figure are $y_L = 1$, $y_H = 3$, $w = 2$, and $\phi = 0.5$.

worker. To see this clearly, suppose that λ is so close to 0 that each firm obtains almost perfect information. This is a good outcome for the high type (who wants to reveal her type) but not for the low type (who wants to hide her type). In this case, the low type enjoys an increase in λ, which reduces the precision of firms' hiring decisions, while the high type dislikes the change.

Corollary 2 indicates that the information effect dominates when λ is relatively small, while the prior-reliance effect dominates otherwise. This can be seen in the shaded regions in Figure 6. If λ is small (pink region), then enough information is already being revealed so that the information effect dominates, and F_H and F_L move in opposite directions. If λ is large enough (blue and green regions), however, prior reliance dominates and F_H and F_L move in the same direction as λ changes.

5 Finite Information-Processing Capacity

There are two prominent ways to introduce costly attention into economic problems, one in which, as in our main model, the agent incurs explicit costs proportional to the amount of information (e.g., Matějka and McKay, 2015; Jung et al., 2019) and the other in which the agent can freely process information up to some finite and fixed capacity (e.g., Sims, 2003; Acharya and Wee, 2020). In the current job search context, the former has an advantage in that it allows the amount of information firms choose to acquire to vary over time, which is arguably more realistic. Technically, however, the standard duality argument would seem to suggest that these two approaches produce qualitatively similar predictions. Specifically, if the unit cost λ of information in the former model is identical to the shadow price of the information-processing constraint in the latter model, then the two approaches prescribe an identical solution. Whereas this fundamental insight *does* apply to an individual firm's optimal interview problem, it *does not* apply to our job search dynamics; that is, the two modelling approaches yield very different dynamic outcomes. We demonstrate this by analyzing the alternative model with a finite information capacity and comparing its outcome to that of our main model in Section 3.

Setup. Recall that if a firm with prior p chooses a joint distribution $G : \{L, H\} \times S \to [0, 1]$, then the associated expected reduction in entropy is equal to

$$I(G; p) \equiv H(p) - \mathbb{E}_G[H(p_s)],$$

where p_s denotes the firm's posterior after observing $s \in S$ and $H(x) \equiv -(1-x)\log(1-x) - x\log x$ for all $x \in [0, 1]$. Instead of the cost-function specification in Section 2, we now assume that for some $\kappa > 0$, each firm can choose any distribution G such that $I(G; p) \leq \kappa$; that is, each firm can reduce expected entropy (without incurring explicit costs) up to κ.

Optimal interview. An individual firm's optimal (interview-and-hiring) problem can be analyzed just as in Section 3.1. The result closely resembles that of the main model but there are some important differences, as formally reported in the following proposition (which is a direct analogue to Proposition 1).

Proposition 4 *Let $\underline{p}^* \in (0, 1/2)$ and $\overline{p}^* \in (1/2, 1)$ be the values such that $H(\underline{p}^*) = H(\overline{p}^*) = \kappa$.*

- *If $p \leq \underline{p}^*$ or $p \geq \overline{p}^*$, then let $\underline{p}(p) \equiv 0$ and $\overline{p}(p) \equiv 1$.*
- *If $p \in (\underline{p}^*, \overline{p}^*)$, then let $\underline{p}(p)(< p)$ and $\overline{p}(p)(> p)$ be the values such that*

$$\underline{p}(p) \equiv \frac{1 - e^{(y_L - w)/\lambda(p)}}{e^{(y_H - w)/\lambda(p)} - e^{(y_L - w)/\lambda(p)}} \text{ and } \overline{p}(p) \equiv \frac{e^{(y_H - w)/\lambda(p)}(1 - e^{(y_L - w)/\lambda(p)})}{e^{(y_H - w)/\lambda(p)} - e^{(y_L - w)/\lambda(p)}}, \quad (5)$$

  where $\lambda(p)(> 0)$ is the value that satisfies

$$H(p) - \frac{\overline{p}(p) - p}{\overline{p}(p) - \underline{p}(p)} H(\underline{p}(p)) - \frac{p - \underline{p}(p)}{\overline{p}(p) - \underline{p}(p)} H(\overline{p}(p)) = \kappa. \quad (6)$$

Given $p \in (0, 1)$, the firm's optimal strategy is to choose a binary distribution that leads to posterior $\underline{p}(p)$ or $\overline{p}(p)$ and hire the worker if and only if its posterior becomes $\overline{p}(p)$.

Proof. See Appendix A. ∎

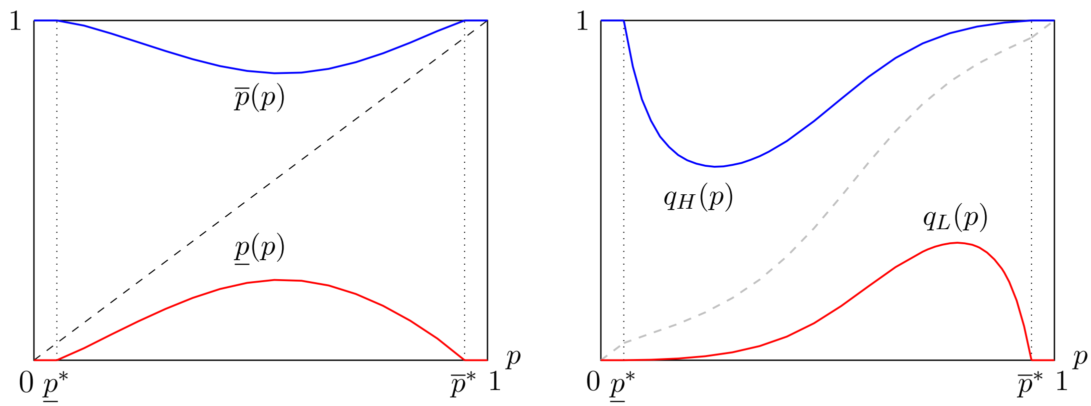

Figure 7: Optimal interview strategy as a function of prior p in the alternative model with finite information capacity. The left panel depicts two posteriors, $\overline{p}(p)$ and $\underline{p}(p)$, induced by the optimal signal, while the right panel presents the optimal binary signal, where $q_\omega = Pr\{s = h|\omega\}$ for $\omega = L, H$. The parameter values used for both panels are $y_L = 1$, $w = 2$, $y_H = 2.8$, and $\kappa = 0.2$.

The logic behind the optimality of a binary signal is identical to that in Section 3.1. Given that, the result follows from effectively the same analysis as in Section 3.1 (which yields the formulae in (5)) and the fact that either perfect learning is feasible (when $p \leq \underline{p}^*$ or $p \geq \overline{p}^*$), or the firm wishes to use up its information-processing capacity κ (which reduces to (6)).

A crucial difference from Proposition 1 is that a firm actively acquires information and conditions on acquired information, *regardless of its prior* p. Recall that in Proposition 1, $\underline{p}$ and $\overline{p}$ are independent of p, and a firm acquires information if and only if $p \in (\underline{p}, \overline{p})$. Obviously, this difference stems from the modelling choice: in Proposition 1, the firm faces the constant marginal cost of information, and thus it chooses not to acquire information if its marginal benefit is sufficiently small (which occurs when p is sufficiently close to 0 or 1). In contrast, according to Proposition 4, the firm faces no information costs until the capacity κ, and thus it chooses to acquire information no matter how small the corresponding marginal benefit is. As we delineate shortly, this implies very different job search dynamics than our main model.

Figure 7 shows how a firm's optimal interview strategy depends on its prior belief p. Both panels are in stark contrast to the corresponding ones in our baseline model, in which $\underline{p}$ and $\overline{p}$ are independent of p, and q_L and q_H are monotone increasing in p (see the left panel of Figure 2). To see why these different patterns arise, suppose that p is relatively close to 0. In this case, information is valuable only when it provides sufficiently strong evidence for high productivity and, therefore, induces the firm to hire the worker despite a low prior belief. Such (precise) information is so costly that the firm chooses simply not to acquire information in our baseline model. In the model with finite information capacity, since some amount of information is free, the firm still tries and chooses a signal that delivers accurate information about $\omega = H$. In fact, if p is below $\underline{p}^*$ then the

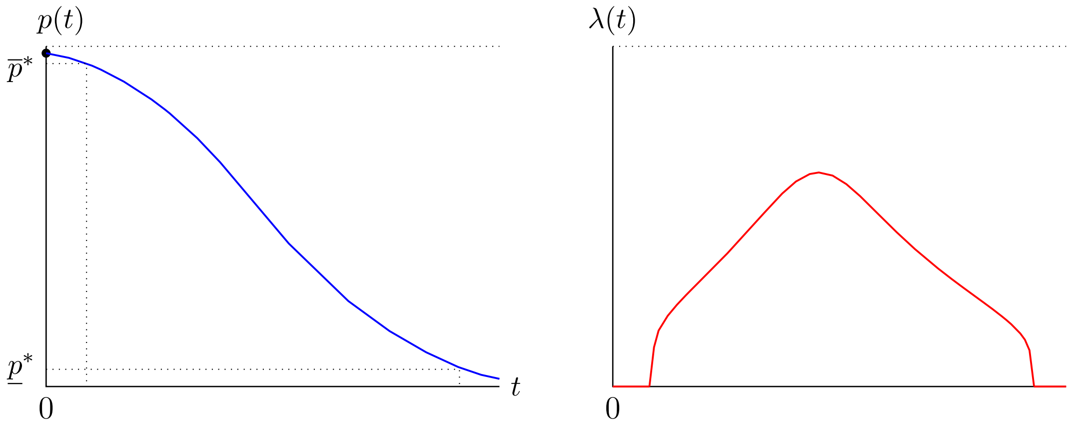

Figure 8: Equilibrium dynamics in the alternative model with limited information processing. The left panel depicts the evolution of $p(t)$ (starting from $p_0 = 0.98 > \overline{p}^* = 0.9495$), while the right panel shows the corresponding the shadow cost of information at each time. In this figure, $\phi = 0.6$, and the other parameter values are identical to those for Figure 7.

firm acquires perfect information, and thus $\overline{p} = q_H = 1$. The opposite reasoning applies when p is close to 1. If p is in the intermediate range, then both types of information, one revealing $\omega = H$ and the other revealing $\omega = L$, are valuable. In our baseline model, the firm responds to this by acquiring more information (see the right panel of Figure 2). With finite information capacity, the firm can only split the information capacity κ equally between the two states, and thus both $\overline{p}(p) - \underline{p}(p)$ and $q_H(p) - q_L(p)$ are relatively small.

Equilibrium dynamics. Given Proposition 4, the equilibrium dynamics can be characterized as in Section 3.2. The current alternative model does not permit a closed-form formula for the evolution of $p(t)$, primarily due to the fact that the shadow price of information, $\lambda(p)$, is defined only implicitly by equation (6). Nevertheless, $p(t)$ still evolves according to

$$\dot{p}(t) = -p(t)(1 - p(t))\phi(q_H(t) - q_L(t)),$$

where $q_\omega(t)$ denotes the probability that the type-ω worker is hired at time t (equivalently, the probability that the posterior becomes equal to $\overline{p}(p(t))$ conditional on $\omega = L, H$ and prior $p(t)$ according to Proposition 4). Therefore, we can still make inferences about the equilibrium dynamics.

Figure 8 depicts the resulting equilibrium dynamics. As shown in the left panel, the most crucial difference from our baseline model is that $p(t)$ strictly decreases over the entire region $(0, 1)$: recall that in our baseline model, $p(t)$ stays constant if $p(t) \leq \underline{p}$ or $p(t) \geq \overline{p}$ (see Figure 3). In fact, in the current model with limited information processing, $p(t)$ decreases even faster when it is close to 1 or 0, because it is when the firm acquires perfect information and, therefore, $q_H(t) = 1$

and $q_L(t) = 0$.

The right panel of Figure 8 shows how the shadow price of information changes over time. It is quasiconcave in general (see Lemma 2 in Appendix A). Initially, $p(t)$ is relatively high, in which case additional information is unlikely to change the firm's decision, so its price is low. In fact, as shown in Figure 8, $\lambda(t) = 0$ until $p(t)$ reaches $\overline{p}^*$. As $p(t)$ falls, $\lambda(t)$ increases, reflecting the fact that the firm relies increasingly more on its own information, raising its shadow price. Once $p(t)$ falls below a certain level, $\lambda(t)$ starts decreasing and eventually converges to 0. Now the default action (the optimal action conditional on prior $p(t)$) is not to hire, and it is less likely to change due to new information, the closer $p(t)$ is to 0.

In summary, the modeling choice between explicit information costs and finite information capacity is not innocuous in our environment. This is precisely because unemployment duration plays a key role in dynamically changing the information-acquisition problem faced by firms. With explicit information costs, firms change how much information they acquire over time, while with a finite information capacity, the perceived cost of information varies.

6 Extensions

In this section, we study two extensions of our baseline model and illustrate how our main lessons apply to richer environments. A formal analysis of each case can be found in Appendix B.

6.1 Adverse Selection

We first consider a model in which the two worker types have different reservation utilities, and each arriving firm makes a take-it-or-leave-it wage offer to the worker after an interview.[16] In this environment, the worker's reservation wage (the lowest wage she would accept) varies across types and over time, and firms must take the associated adverse selection into account.

Formally, let w_ω denote the type-ω worker's reservation utility and assume that $w_L < y_L < w_H < y_H$.[17] To focus on the more interesting case, we also assume that the worker is sufficiently patient or firms arrive fast enough such that

$$y_L - w_L < \int_0^\infty e^{-rt}(w_H - w_L)d\left(1 - e^{-\phi t}\right) = \frac{\phi}{r+\phi}(w_H - w_L) \Leftrightarrow \frac{r(y_L - w_L)}{\phi(w_H - y_L)} < 1. \quad (7)$$

[16] This model can be interpreted as the endogenous-information counterpart to Kaya and Kim (2018): they study the same physical environment but assume that firms (buyers) do not have control over information acquisition.

[17] This order embodies two (common) economic assumptions: (i) there are always gains from trade (i.e., $w_\omega < y_\omega$ for both $\omega = L, H$) and (ii) adverse selection is sufficiently severe (i.e., $y_L < w_H$). If either of these properties fails, then the analysis becomes simpler and the result becomes less interesting, both of which can be easily inferred from the subsequent discussion.

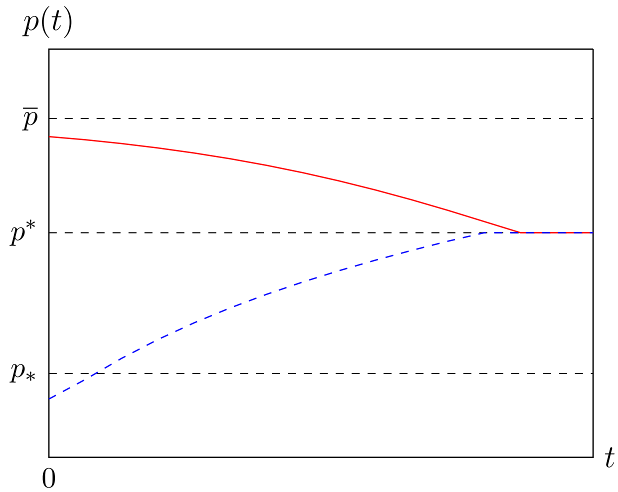

Figure 9: The blue (dashed) curve represents a belief path with a prior below p^*, while the red (solid) curve represents a belief path with a prior above p^*. The parameter values used in this figure are $y_L = 1$, $y_H = 3$, $c_L = 2/3$, $c_H = 2$, $\phi = 0.9$, $\lambda = 1$, and $r = 0.9$.

In other words, the low-type worker is willing to reject y_L (the highest wage a firm is willing to offer, knowing the worker's type) and wait if she expects to receive w_H from the next arriving firm. As is well-known in the literature, in equilibrium, no firm offers strictly more than w_H, the high type accepts w_H with probability 1, and the low type's reservation wage, denoted by $w(t)$, lies between w_L and w_H.

There are two immediate but important differences from the baseline model. First, even if a firm receives a negative signal ℓ about the worker, it may still try to hire her by offering a low wage $w(t)$, provided that $w(t)$ does not exceed y_L. The offer might not be accepted by the worker (if she is the high type) but would yield payoff $y_L - w(t)$ if she is the low type. This clearly affects firms' interview strategies as well as their hiring decisions. Second, it is no longer the case that the worker's employment probability becomes vanishingly small in the long run: if it were the case, then the low-type worker's reservation wage would become close to w_L. But then, firms would want to hire the low-type worker by offering a wage below y_L.

Figure 9 shows how these changes influence job search dynamics. If p_0 is sufficiently large (above $\overline{p}$), there is no change: all firms offer w_H without acquiring information, and thus $p(t)$ stays equal to p_0. If p_0 is slightly below $\overline{p}$, then firms acquire information, but they still offer w_H with a high probability. In this case, the low-type worker's reservation wage is too high (exceeding y_L), hiring occurs only at w_H (equivalently, only when the firm receives $s = h$), and so the dynamics for $p(t)$ are exactly the same as in the baseline model. One non-trivial difference is that once $p(t)$ reaches a certain level, p^*, it remains constant thereafter. It plays a similar role to $\underline{p}$ in the baseline model. However, $p(t)$ reaches p^* in finite time (whereas $p(t)$ converges to $\underline{p}$ asymptotically in our baseline model). In addition, at p^*, the worker's employment probability does not vanish away and firms continue to acquire information. Still, $p(t)$ remains equal to p^* because $w(t) = y_L$ (which renders each firm to be indifferent between hiring the low type and not), and thus it is possible that the firm offers $y_L$with just enough probability to compensate for the difference in the probability of w_H.[18]

If p_0 falls short of p^*, then the belief $p(t)$ *increases* over time. In this case, it is easy to check that $w(t) < y_L$ and the low-type worker always accepts $w(t)$. This implies that the low-type worker is hired as long as she meets a firm, whereas the high-type worker is hired only when the firm draws $s = h$. Taken together, the probability that the remaining worker is of the high type increases over time. The region below p^* can be further divided, depending on whether firms acquire information or not. If $p(t)$ is sufficiently small (below p_* in the figure) then, as in the baseline model, firms have no incentive to acquire information. In this case, they only offer $w(t)$. If $p(t)$ is rather close to p^* (above p_*), then firms acquire information and offer differential wages, depending on the signal realization. In this case, they do not simply offer $w(t)$, because $w(t)$ is close to y_L and so it is crucial (relatively more profitable) for firms to identify and hire the high-type worker.

6.2 Endogenous Search Intensity

In our baseline model, firms become more pessimistic about the worker's productivity (i.e., $p(t)$ falls over time) and, therefore, are less likely to hire the worker as her unemployment duration rises. This implies that the worker is particularly eager to meet firms at an early stage of her job search when a meeting will likely translate into employment. Our second extension allows the worker to control her search intensity and studies its effects on job search dynamics.

Specifically, we consider a model in which the low-type worker can privately increase her search intensity ϕ to $\overline{\phi}$ at flow cost $c(> 0)$ at any time, while such an option is not available to the high type (i.e., the high-type worker's search intensity is fixed at ϕ). The latter assumption not only simplifies the analysis but also makes firms' inference problems more interesting, as shown shortly.[19] To avoid triviality, we assume that c is positive but sufficiently small.[20] All other aspects

[18]To be precise, let q_ω denote the probability that the firm receives $s = h$ and so offers w_H, and σ denote the probability that the firm offers $w(t)$ conditional on $s = \ell$. When $q_H > q_L$, $p(t)$ does not change if

$$q_H = q_L + (1 - q_L)\sigma \Leftrightarrow \sigma = \frac{q_H - q_L}{1 - q_L}.$$

[19]The main intuition and comparison to the baseline model remain, for example, if the high-type worker meets firms at rate $\overline{\phi}$ (for free), while the low-type worker has to pay for $\overline{\phi}$.

[20]Precisely, all subsequent analysis and discussion hold if

$$c < \frac{e^{(y_L - w)/\lambda} w(\overline{\phi} - \phi)(\phi(1 - \underline{p})p_0 - \overline{\phi}\underline{p}(1 - p_0))}{e^{(y_L - w)/\lambda}\phi(\phi(1 - \underline{p})p_0 - \overline{\phi}\underline{p}(1 - p_0)) + r\overline{\phi}(1 - p_0)(1 - e^{(y_L - w)/\lambda})}.$$

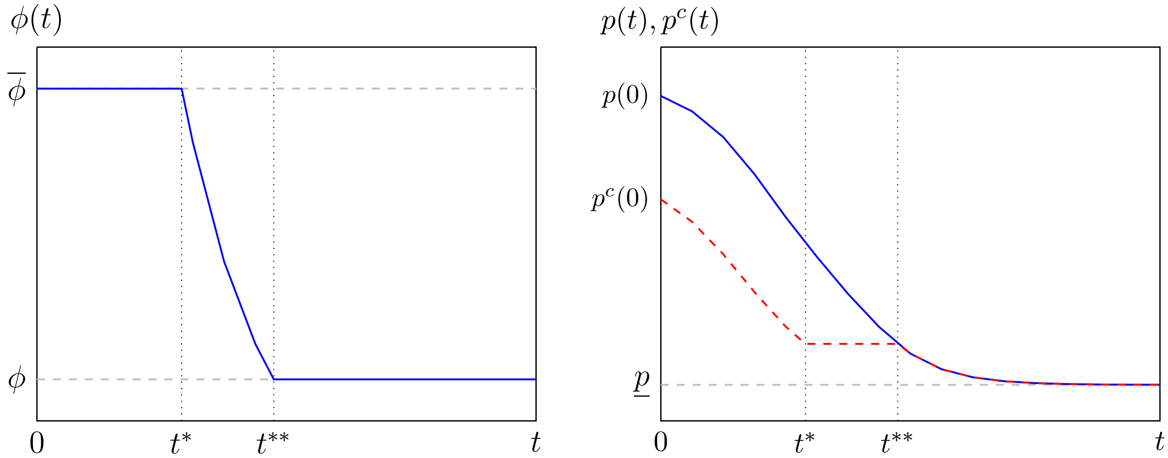

Figure 10: The left panel shows the low-type worker's equilibrium search intensity, while the right panel draws the resulting evolution of market beliefs. In the right panel, the solid blue curve depicts $p(t)$ (unconditional belief), while the dashed red curve shows $p^c(t)$ (conditional belief). The parameter values used for this figure are $y_L = 2$, $w = 5.5$, $y_H = 10$, $p_0 = 0.8$, $\phi = 0.3$, $\overline{\phi} = 1$, $\lambda = 2$, $r = 1$, and $c = 0.1$.

of the baseline model remain unchanged. Allowing for the low-type worker's mixing between ϕ and $\overline{\phi}$, we let $\phi(t)$ denote her expected search intensity at t.

To see the basic effects of endogenous search intensity, suppose that p_0 is slightly below $\overline{p}$ and firms behave just as in the baseline model; in particular, early firms do not acquire much information and hire with high probability. Then, the low-type worker would increase her search intensity at the beginning of job search. If firms anticipate that, however, they would become more reluctant to hire, because the worker is more likely to be the low type conditional on meeting. This directly lowers firms' hiring incentives and also affects their information-acquisition incentives, thereby altering the overall job search dynamics.

To be precise, let $p^c(t)$ denote the probability that the worker is the high type *conditional* on facing a firm. If the low type chooses ϕ (i.e., does not increase her search intensity), then $p^c(t)$ coincides with the *unconditional* belief $p(t)$. If the low type chooses $\overline{\phi}$, however, $p^c(t)$ becomes strictly smaller than $p(t)$: formally, by Bayes' rule,

$$p^c(t) = \frac{p(t)\phi}{p(t)\phi + (1 - p(t))\overline{\phi}} = \frac{p(t)}{p(t) + (1 - p(t))\overline{\phi}/\phi} < p(t).$$

Given $p^c(t)$, the firm's problem is just as in the baseline model: it acquires information if and only if $p^c(t) \in (\underline{p}, \overline{p})$ and hires the worker only when $s = h$. The difficulty lies in the fact that the low-type worker's optimal search intensity (and thus the relationship between $p(t)$ and $p^c(t)$) varies over time and must be determined jointly with firms' optimal strategies. However, it remains true that firms' beliefs, both $p(t)$ and $p^c(t)$, are decreasing in t, which in turn implies that the low-type worker's expected search intensity decreases over time.[21] These monotone properties render the analysis tractable.

Figure 10 illustrates the resulting equilibrium dynamics. As shown in the left panel, the low-type worker's expected search intensity decreases over time. Initially, she chooses $\overline{\phi}$ with probability 1 and then mixes between ϕ and $\overline{\phi}$. Once firms become sufficiently reluctant to hire, however, she has little incentive to increase her search intensity and, therefore, chooses ϕ. Such search behavior by the low-type worker implies that, as shown in the right panel, the difference between $p(t)$ and $p^c(t)$ is large initially, decreases over time, and eventually vanishes. It is noteworthy that $p^c(t)$ remains constant over the interval when the low-type worker mixes, even though both $\phi(t)$ and $p(t)$ strictly fall over the interval. This is necessary to make the low-type worker indifferent between ϕ and $\overline{\phi}$

7 Conclusion

We present a model of job search, in which firms acquire information about a worker's unobservable productivity in a fully flexible manner. Specifically, we model firms' interview (information-acquisition) strategies and costs as in the literature on rational inattention. Whereas it is typically perceived that flexible information acquisition renders a model analytically intractable, our model admits a closed-form characterization of equilibrium job search dynamics, which helps delineate the roles of information costs and prior beliefs for unemployment duration and worker welfare. Importantly, we find that (i) flexible information acquisition (endogenous information) amplifies the stigma effect, leading to a faster drop in the worker's employment probability than with exogenous information, and (ii) lowering firms' information-acquisition costs may or may not increase the worker's unemployment duration, depending not only on her unobservable productivity but also on market prior beliefs about it. We believe that these implications can be taken to data or field experiments, as have been done in some important recent studies (e.g., Bartoš et al., 2016; Jarosch and Pilossoph, 2019), and our model can be fruitfully exploited to better understand other related phenomena in the data.

References

Acemoglu, Daron, "Public policy in a model of long-term unemployment," *Economica*, 1995,

[21] This remains true *even* though the low type meets more firms and, therefore, has more employment opportunities.

pp. 161–178. 1

Acharya, Sushant and Shu Lin Wee, "Rational inattention in hiring decisions," *American Economic Journal: Macroeconomics*, 2020, *12* (1), 1–40. 3, 15

Aumann, Robert J. and Michael Maschler, *Repeated Games with Incomplete Information* 1995. 6

Bartoš, Vojtěch, Michal Bauer, Julie Chytilová, and Filip Matějka, "Attention discrimination: Theory and field experiments with monitoring information acquisition," *American Economic Review*, 2016, *106* (6), 1437–75. 3, 23

Blanchard, Olivier Jean and Peter Diamond, "Ranking, unemployment duration, and wages," *The Review of Economic Studies*, 1994, *61* (3), 417–434. 1

Briggs, Joseph, Andrew Caplin, Daniel Martin, Christopher Tonetti, and GSB Stanford, "Due Diligence: Job Search with Rationally Inattentive Workers," Technical Report, NYU working paper 2017. 3

Caplin, Andrew, Mark Dean, and John Leahy, "Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy," *mimeo*, 2019. 6

Cheremukhin, Anton, Paulina Restrepo-Echavarria, and Antonella Tutino, "Targeted search in matching markets," *Journal of Economic Theory*, 2020, *185*, 104956. 3

Coles, Melvyn G and Eric Smith, "Marketplaces and matching," *International Economic Review*, 1998, pp. 239–254. 1

Gentzkow, Matthew and Emir Kamenica, "Costly persuasion," *American Economic Review*, 2014, *104* (5), 457–62. 5

Jarosch, Gregor and Laura Pilossoph, "Statistical discrimination and duration dependence in the job finding rate," *The Review of Economic Studies*, 2019, *86* (4), 1631–1665. 23

Jung, Junehyuk, Jeong Ho (John) Kim, Filip Matejka, and Christopher A Sims, "Discrete Actions in Information-Constrained Decision Problems," *The Review of Economic Studies*, 2019, *86* (6), 2643–2667. 15

Kamenica, Emir and Matthew Gentzkow, "Bayesian persuasion," *The American Economic Review*, 2011, *101* (6), 2590–2615. 6

Kaya, Ayça and Kyungmin Kim, "Trading dynamics with private buyer signals in the market for lemons," *The Review of Economic Studies*, 2018, *85* (4), 2318–2352. 3, 19, 28

Lentz, Rasmus and Torben Tranaes, "Job search and savings: Wealth effects and duration dependence," *Journal of labor Economics*, 2005, *23* (3), 467–489. 1

Lockwood, Ben, "Information externalities in the labour market and the duration of unemployment," *The Review of Economic Studies*, 1991, *58* (4), 733–753. 1, 3

Mackowiak, Bartosz Adam, Filip Matejka, and Mirko Wiederholt, "Rational Inattention: A Review," *mimeo*, 2020. 3

Matějka, Filip and Alisdair McKay, "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *American Economic Review*, 2015, *105* (1), 272–98. 6, 15

Merlo, Antonio and Francois Ortalo-Magne, "Bargaining over residential real estate: evidence from England," *Journal of urban economics*, 2004, *56* (2), 192–216. 1

Palazzo, Francesco, "Search costs and the severity of adverse selection," *Research in Economics*, 2017, *71* (1), 171–197. 3

Sims, Christopher A, "Implications of rational inattention," *Journal of monetary Economics*, 2003, *50* (3), 665–690. 15

Taylor, Curtis R, "Time-on-the-market as a sign of quality," *The Review of Economic Studies*, 1999, *66* (3), 555–578. 1, 3

Tucker, Catherine, Juanjuan Zhang, and Ting Zhu, "Days on market and home sales," *The RAND Journal of Economics*, 2013, *44* (2), 337–360. 1

Vishwanath, Tara, "Job search, stigma effect, and escape rate from unemployment," *Journal of Labor Economics*, 1989, *7* (4), 487–502. 1, 3

Wu, Liangjie, "Partially directed search in the labor market," *mimeo*, 2020. 3

Zhu, Haoxiang, "Finding a good price in opaque over-the-counter markets," *The Review of Financial Studies*, 2012, *25* (4), 1255–1285. 3

Zou, Junyuan, "Information Traps in Over-the-Counter Markets," *mimeo*, 2019. 3

A Omitted Proofs

Proof of Corollary 2. It remains to show that A_ω moves in the direction stated in the corollary. For A_H, one can show that

$$\frac{\partial A_H}{\partial \lambda} = \frac{\frac{\partial \underline{p}}{\partial \lambda}\overline{p} - \frac{\partial \overline{p}}{\partial \lambda}\underline{p}}{\left(p_0 - \underline{p}\right)^2 \overline{p}^2} p_0 \left(\pi_H\left(\lambda\right) - p_0\right),$$

where

$$\pi_H\left(\lambda\right) = 1 - \frac{\left(e^{(y_H - w)/\lambda} - 1\right)^2}{\left(e^{(y_H - y_L)/\lambda} - 1\right)^2} e^{(w - y_L)/\lambda} \frac{y_H - y_L}{y_H - w}.$$

It is straightforward to show that

$$\frac{\partial \pi_H}{\partial \lambda} < 0 \text{ and } \lim_{\lambda \to \infty} \pi_H = 1, \lim_{\lambda \to \infty} \pi_H = \widehat{p}.$$

Now, if $p_0 \leq \widehat{p}$, $\pi_H - p_0 > 0$ for all λ, and in turn, $\partial F_H\left(t\right)/\partial \lambda < 0$ for all λ in this case. If $p_0 > \widehat{p}$, there exists some $\lambda_H\left(p_0\right)$ such that $\pi_H\left(\lambda_H\right) = p_0$, and in turn, $\partial A_H / \partial \lambda \geq (\leq) \, 0$ if and only if $\lambda \leq (\geq) \, \lambda_H$. Therefore, $\partial F_H\left(t\right)/\partial \lambda \leq (\geq) \, 0$ if and only if $\lambda \leq (\geq) \, \lambda_H$ in this other case.

Similarly, one can show that

$$\frac{\partial A_L}{\partial \lambda} = \frac{\frac{\partial \underline{p}}{\partial \lambda}\left(1 - \overline{p}\right) - \left(1 - \underline{p}\right)\frac{\partial \overline{p}}{\partial \lambda}}{\left(p_0 - \underline{p}\right)^2 \left(1 - \overline{p}\right)^2} \left(1 - p_0\right)\left(\pi_L\left(\lambda\right) - p_0\right)$$

where

$$\pi_L\left(\lambda\right) = \overline{p}\underline{p}\frac{y_H - y_L}{w - y_L}.$$

Again, it can be easily shown that

$$\frac{\partial \pi_L}{\partial \lambda} > 0 \text{ and } \lim_{\lambda \to 0} \pi_L = 0, \lim_{\lambda \to \infty} \pi_L = \widehat{p}.$$

Now, if $p_0 < \widehat{p}$, there exists some $\lambda_L\left(p_0\right)$ such that $\pi_L\left(\lambda_L\right) = p_0$, and in turn, $\partial A_L / \partial \lambda \leq (\geq) \, 0$ if and only if $\lambda \leq (\geq) \, \lambda_L$. Therefore, $\partial F_L\left(t\right)/\partial \lambda \geq (\leq) \, 0$ if and only if $\lambda \leq (\geq) \, \lambda_L$ in this case. If $p_0 \geq \widehat{p}$, $\pi_L - p_0 < 0$ for all λ, and in turn, $\partial F_L\left(t\right)/\partial \lambda > 0$ for all λ in this other case. ■

Proof of Proposition 4. Notice that $H(p)$ represents the cost of getting perfect information: if a firm acquires perfect information, then its posterior becomes equal to either 0 (with probability $1 - p$) or 1 (with probability p). In this case, the total expected reduction in entropy is equal to

$$H(p) - ((1 - p)H(0) + pH(1)) = H(p).$$

An immediate implication is that perfect information is feasible to the firm (which surely wants to choose it) if and only if

$$H(p) \leq \kappa \Leftrightarrow p \leq \underline{p}^* \text{ or } p \geq \overline{p}^*.$$

Suppose $p \in (\underline{p}^*, \overline{p}^*)$. In this case, given λ, the analysis is identical to that of the main model. Specifically, it is optimal for the firm to choose a binary signal that leads to the following posteriors:

$$\underline{p} \equiv \frac{1 - e^{(y_L - w)/\lambda}}{e^{(y_H - w)/\lambda} - e^{(y_L - w)/\lambda}} \text{ and } \overline{p} \equiv \frac{e^{(y_H - w)/\lambda}(1 - e^{(y_L - w)/\lambda})}{e^{(y_H - w)/\lambda} - e^{(y_L - w)/\lambda}}.$$

It remains to show that there exists a unique $\lambda (> 0)$ that satisfies the following equation:

$$I(\lambda) \equiv H(p) - \frac{\overline{p} - p}{\overline{p} - \underline{p}} H(\underline{p}) - \frac{p - \underline{p}}{\overline{p} - \underline{p}} H(\overline{p}) = \kappa.$$

By direct calculus,

$$\begin{aligned} I'(\lambda) &= \frac{1}{(\overline{p} - \underline{p})^2} \left(\frac{\partial \overline{p}}{\partial \lambda} (p - \underline{p}) \left(\underline{p} \log \left(\frac{\underline{p}}{\overline{p}} \right) + (1 - \underline{p}) \log \left(\frac{1 - \underline{p}}{1 - \overline{p}} \right) \right) \right) \\ &\quad + \frac{1}{(\overline{p} - \underline{p})^2} \left(\frac{\partial \underline{p}}{\partial \lambda} (\overline{p} - p) \left(\overline{p} \log \left(\frac{\underline{p}}{\overline{p}} \right) + (1 - \overline{p}) \log \left(\frac{1 - \underline{p}}{1 - \overline{p}} \right) \right) \right). \end{aligned}$$

By Lemma 1, $\partial \underline{p} / \partial \lambda > 0$ and $\partial \overline{p} / \partial \lambda < 0$. In addition, whenever $\underline{p} < \overline{p}$,

$$\underline{p} \log \left(\frac{\underline{p}}{\overline{p}} \right) + (1 - \underline{p}) \log \left(\frac{1 - \underline{p}}{1 - \overline{p}} \right) > 0 > \overline{p} \log \left(\frac{\underline{p}}{\overline{p}} \right) + (1 - \overline{p}) \log \left(\frac{1 - \underline{p}}{1 - \overline{p}} \right).$$

It then follows that $I'(\lambda) < 0$ for any λ, and thus there exists at most one value of λ such that $I(\lambda) = \kappa$. Furthermore, such a λ exists if and only if

$$\lim_{\lambda \to 0} I(\lambda) = H(p) > \kappa > \lim_{\lambda \to 0} I(\lambda) = 0,$$

which holds if and only if $p \in (\underline{p}^*, \overline{p}^*)$. ■

Lemma 2 *$\lambda(p)$ is quasiconcave in $p \in (\underline{p}^*, \overline{p}^*)$.*

Proof. Recall that $\lambda(p)$ is implicitly defined by the following function:

$$J(\lambda, p) = I(\lambda) - \kappa = 0.$$

As shown in the proof of Proposition 4, $\partial J / \partial \lambda = I'(\lambda) < 0$. Together with the implicit function theorem, this implies that $d\lambda/dp$ has the same sign as

$$\frac{\partial J}{\partial p} = -\log\left(\frac{p}{1-p}\right) + \frac{H(\underline{p}) - H(\overline{p})}{\overline{p} - \underline{p}}.$$

The desired result follows from the fact that

$$\frac{\partial}{\partial p}\left(\frac{\partial J}{\partial p}\right) = -\frac{1}{p(1-p)} < 0.$$

■

B Analysis for Section 6

5.1 Adverse Selection

In this appendix, we explicitly construct the equilibrium described in Section 6.1.[22]

Stationary Path

We begin by showing that there exists a unique value of p^* such that once $p(t)$ reaches p^*, it stays constant thereafter. Observe that for $p(t)$ to stay constant, either $w(t) = y_L$, or firms must hire the worker with probability 1, without acquiring information. In the latter case, obviously, $p(t)$ stays constant. Suppose that the firm does acquire information and does not hire the worker with probability 1. If $w(t) < y_L$, then the firm hires the low type, whether the realized signal is ℓ (in which case the firm offers $w(t)$) or h (in which case it offers w_H). Since the firm hires the high type with probability less than 1 (only when $s = \ell$ and so the firm offers $w(t)$), $p(t)$ must increase in this case. If $w(t) > y_L$, then the firm hires the worker if and only if $s = h$. Since the high type generates signal h with a higher probability than the low type, $p(t)$ decreases in this case.

The above result implies that at p^*, the firm's optimal interview strategy would be identical to that of the baseline model. The only difference is that when signal ℓ is realized, the firm simply does not hire the worker in the baseline model, whereas now it can offer $w(t) = y_L$ with some positive probability.[23] From Section 3.1, it follows that the optimal probabilities, q_L^* and q_H^*, to

[22] It can be shown that it is the unique equilibrium unless $p_0 = \overline{p}$. This result and proof are effectively identical to those of Kaya and Kim (2018) and thus omitted.

[23] We assume that the low-type worker always accepts her reservation wage $w(t)$ but each firm may make a losing offer (below $w(t)$). Alternatively, one may assume that each firm should offer either $w(t)$ or w_H (i.e., cannot make a losing offer), but the low-type seller may randomize between accepting and rejecting $w(t)$.

generate signal h and offer w_H conditional on each type $\omega = L, H$ are given by

$$\frac{\underline{p}}{1-\underline{p}} = \frac{p^*}{1-p^*}\frac{1-q_H^*}{1-q_L^*} \text{ and } \frac{\overline{p}}{1-\overline{p}} = \frac{p^*}{1-p^*}\frac{q_H^*}{q_L^*}, \tag{8}$$

where

$$\underline{p} \equiv \frac{1-e^{(y_L-w_H)/\lambda}}{e^{(y_H-w_H)/\lambda}-e^{(y_L-w_H)/\lambda}} \text{ and } \overline{p} \equiv \frac{e^{(y_H-w_H)/\lambda}(1-e^{(y_L-w_H)/\lambda})}{e^{(y_H-w_H)/\lambda}-e^{(y_L-w_H)/\lambda}}.$$

Given q_L^* (i.e., given the fact that all subsequent firms offer w_H with probability q_L^* to the low type), the low-type worker's reservation wage is equal to

$$w_L + \int_t^\infty e^{r(s-t)}(w_H - w_L)d\left(1-e^{-\phi q_L^*(s-t)}\right) = \frac{\phi q_L^*}{r+\phi q_L^*}(w_H - w_L).$$

For $w(t) = y_L$, it must be that

$$q_L^* = \frac{r(y_L - w_L)}{\phi(w_H - y_L)}. \tag{9}$$

Equations (8) and (9) fully characterize p^*, q_L^*, and q_H^*: given the unique value of q_L^* in (9), p^* and q_H^* are fully determined by (8). Note also that inequality (7) ensures that p^*, q_L^*, and q_H^* are all well-defined in $(0,1)$.

Now it suffices to specify the low-type worker's acceptance strategy on the stationary path. Let σ^* denote the probability that each firm offers $w(t) = y_L$ conditional on receiving $s = \ell$ when its prior belief is p^*. Since both types accept w_H with probability 1, $p(t)$ stays constant if and only if

$$q_H^* = q_L^* + (1-q_L^*)\sigma^* \Leftrightarrow \sigma^* = \frac{q_H^* - q_L^*}{1-q_L^*}. \tag{10}$$

When $p(t) > p^*$

If $p(t) > p^*$, then the low-type worker's reservation wage $w(t)$ exceeds y_L. Then, the firm has no incentive to offer $w(t)$ (as it will be accepted only by the low type but $w(t) > y_L$) and hiring occurs only at w_H. This implies that the equilibrium outcome will be just as in our baseline model. Specifically, firms acquire information as in Proposition 1 (where w is replaced by w_H), and the belief stays constant if $p(t) \geq \overline{p}$ and continuously falls if $p(t) \in (p^*, \overline{p})$. The only difference is that once $p(t)$ reaches p^*, it stays constant thereafter: recall that in our baseline model, $p(t)$ asymptotically converges to $\underline{p}(< p^*)$.

When $p(t) < p^*$

If $p(t) < p^*$, then the low-type worker's reservation wage falls short of y_L. This implies that the low type always gets hired (because the firm will offer $w(t)$ conditional on receiving $s = \ell$), and thus $p(t)$ increases over time. Formally, letting $q_\omega(t)$ denote the probability that the firm receives signal h and offers w_H at time t, $p(t)$ increases according to

$$p(t+dt) = \frac{p(t)e^{-\phi q_H(t)dt}}{p(t)e^{-\phi q_H(t)dt} + (1-p(t))e^{-\phi dt}} \Leftrightarrow \dot{p}(t) = \phi p(t)(1-p(t))(1-q_H(t)). \quad (11)$$

There are two cases to consider, depending on whether $q_H(t) > 0$ or $q_H(t) = 0$. Clearly, the former corresponds to the case where the firm acquires information and offers differential wages, depending on the signal realization, while the latter is when the firm chooses to offer $w(t)$ without acquiring information. The following lemma, which is a straightforward modification of Proposition 1, shows under what condition each case arises.

Lemma 3 *Suppose $p(t) < p^*$. Then, the firm acquires information if and only if*

$$p(t) > \frac{1 - e^{(w(t)-w_H)/\lambda}}{e^{(y_H-w_H)/\lambda} - e^{(w(t)-w_H)/\lambda}}. \quad (12)$$

Notice that the only difference from Proposition 1 is that $w(t) - w_H$ appears in lieu of $y_L - w$, with w_H playing the role of w in the current extension. This follows from the fact that, since the firm now obtains $y_L - w(t)$ from hiring the low type even when $s = \ell$, its total benefit of choosing (q_H, q_L) is equal to

$$\begin{aligned} B(q_H, q_L) &= pq_H(y_H - w_H) + (1-p)q_L(y_L - w_H) + (1-p)(1-q_L)(y_L - w(t)) \\ &= pq_H(y_H - w_H) + (1-p)q_L(w(t) - w_H) + (1-p)(y_L - w(t)), \end{aligned}$$

instead of $B(q_H, q_L) = pq_H(y_H - w_H) + (1-p)q_L(y_L - w_H)$, which was the case in the baseline model.

The right-hand side in (12) decreases in $w(t)$. Combining this with the fact that both $p(t)$ and $w(t)$ increase over time when $p(t) < p^*$, it follows that there exists $p_*(< p^*)$ such that $q(t) > 0$ if and only if $p(t) \in (p_*, p^*)$.

(i) $p(t) \in (p_*, p^*)$**:** In this case, as shown above, the firm acquires information and offers w_H if and only if $s = h$. Precisely, given $w(t)$, by the same argument as in Section 3.1, the firm employs a binary signal so that

$$\frac{1 - e^{(w(t)-w_H)/\lambda}}{e^{(y_H-w_H)/\lambda} - 1} = \frac{p(t)}{1-p(t)} \frac{1-q_H(t)}{1-q_L(t)} \text{ and } \frac{e^{(y_H-w_H)/\lambda}(1 - e^{(w(t)-w_H)/\lambda})}{e^{(w(t)-w_H)/\lambda}(e^{(y_H-w_H)/\lambda} - 1)} = \frac{p(t)}{1-p(t)} \frac{q_H(t)}{q_L(t)}.$$

Meanwhile, $p(t)$ increases according to (11), and $w(t)$ increases according to

$$r(w(t) - w_L) = \phi q_L(t)(w_H - w(t)) + \dot{w}(t),$$

until $w(t)$ reaches y_L (or, equivalently, $p(t)$ arrives at p^*).

(ii) $p(t) \leq p_*$**:** In this case, the firm always offers $w(t)$ without acquiring information. Then, $p(t)$ and $w(t)$ continuously increase according to

$$\dot{p}(t) = \phi p(t)(1 - p(t)) \text{ and } r(w(t) - w_L) = \dot{w}(t),$$

respectively, until $p(t)$ hits p_*.

5.2 Endogenous Search Intensity

In this appendix, we construct the equilibrium introduced and discussed in Section 6.2.

Let $\phi(t)$ denote the low-type worker's expected search intensity at t. Then, the relationship between $p(t)$ and $p^c(t)$ is given as follows:

$$p^c(t) = \frac{p(t)\phi}{\phi p(t)\phi + \phi(t)(1 - p(t))} \Leftrightarrow \frac{p^c(t)}{1 - p^c(t)} = \frac{p(t)}{1 - p(t)} \frac{\phi}{\phi(t)}.$$

Since $\phi(t) \in [\underline{\phi}, \overline{\phi}]$, it is necessarily the case that $p^c(t) \leq p(t)$. In addition, given $p^c(t)$, the implied $(q_L(t), q_H(t))$, and the fact that the high-type worker's search intensity is fixed at ϕ, the unconditional belief $p(t)$ evolves according to

$$\dot{p}(t) = -p(t)(1 - p(t))(\phi q_H(t) - \phi(t) q_L(t)). \tag{13}$$

As argued in the main text, given $p^c(t)$, firms' optimal strategies are just as in the baseline model. Therefore, by Proposition 1,

$$q_L(t) = \frac{e^{(y_L - w)/\lambda}(p^c(t) - \underline{p})}{(1 - e^{(y_L - w)/\lambda})(1 - p^c(t))} \text{ and } q_H(t) = \frac{e^{(y_H - w)/\lambda}(p^c(t) - \underline{p})}{(e^{(y_H - w)/\lambda} - 1)p^c(t)}.$$

Applying these to (13), we get

$$\dot{p}(t) = -\frac{e^{(y_L - w)/\lambda}}{1 - e^{(y_L - w)/\lambda}}(1 - p(t))(\phi(1 - \underline{p})p(t) - \phi(t)\underline{p}(1 - p(t))) \left(\frac{\overline{p}}{1 - \overline{p}} - \frac{p(t)}{1 - p(t)} \right). \tag{14}$$

The result is trivial if either $p_0 \leq \underline{p}$ (in which case no firm hires) or $p_0 \geq \overline{p}/(\overline{p} + (1 - \overline{p})\phi/\overline{\phi})$

(in which case all firms hire without acquiring information). Therefore, we restrict attention to the case where $p_0 \in (\underline{p}, \overline{p}/(\overline{p} + (1-\overline{p})\phi/\overline{\phi}))$.

Let t^* be the last time at which the low-type worker chooses $\overline{\phi}$ with probability 1 and t^{**} be the first time at which the low-type worker chooses ϕ with probability 1. In other words, $\phi(t)$ is given by

$$\phi(t) = \begin{cases} \overline{\phi} & \text{if } t \le t^* \\ \in (\phi, \overline{\phi}) & \text{if } t \in (t^*, t^{**}) \\ \phi & \text{if } t > t^{**}. \end{cases}$$

Note that both t^* and t^{**} can be equal to 0, depending on p_0.

First phase ($t \le t^*$): For ease of notation, let $L \equiv e^{(y_L - w)/\lambda}$. Then, with initial condition $p(0) = p_0$, (14) yields

$$p(t) = \frac{\overline{p} + \frac{\overline{\phi}\underline{p}(1-\overline{p})}{\phi(1-\underline{p})}\beta e^{kt}}{1 + \frac{(\overline{\phi}\underline{p}+\phi(1-\underline{p}))(1-\overline{p})}{\phi(1-\underline{p})}\beta e^{kt}} = \frac{\phi\overline{p}(1-\underline{p}) - \overline{\phi}\underline{p}(1-\overline{p}) + \overline{\phi}\underline{p}(1-\overline{p})(\beta e^{kt}+1)}{\phi\overline{p}(1-\underline{p}) - \overline{\phi}\underline{p}(1-\overline{p}) + (\overline{\phi}\underline{p} + \phi(1-\underline{p}))(1-\overline{p})(\beta e^{kt}+1)},$$

where

$$k = \frac{L}{1-L}\frac{\phi\overline{p}(1-\underline{p}) - \overline{\phi}\underline{p}(1-\overline{p})}{1-\overline{p}} \text{ and } \beta = \frac{\phi(1-\underline{p})(\overline{p} - p_0)}{(1-\overline{p})(\phi(1-\underline{p})p_0 - \overline{\phi}\underline{p}(1-p_0))}.$$

Note that $\beta \ge 0$ because $p^c(0) \ge \underline{p}$. This also implies that $k \ge 0$. Plugging this into $q_L(t)$, we get $q_L(t) = k/(\overline{\phi}\left(1 + \beta e^{kt}\right))$.

Let $\overline{V}(t)$ denote the low type's expected payoff at time t. Then, we have

$$r\overline{V}(t) = -c + \overline{\phi}q_L(t)\left(\frac{w}{r} - \overline{V}(t)\right) + \dot{\overline{V}}(t) = -c + \frac{k}{(1+\beta e^{kt})}\left(\frac{w}{r} - \overline{V}(t)\right) + \dot{\overline{V}}(t),$$

which leads to

$$\overline{V}(t) = -\frac{c}{r+k} + \frac{kw - dk\beta e^{kt} + c_1 e^{t(k+r)}}{r(k+r)(\beta e^{kt}+1)}$$

where

$$c_1 = e^{-t^*(k+r)}\left(\left(\overline{V}(t^*) + \frac{c}{k+r}\right) r(k+r)(\beta e^{kt^*}+1) - kw + dk\beta e^{kt^*}\right).$$

Second Phase: $t \in (t^*, t^{**})$**.** Let $\underline{V}(t)$ be the low-type worker's value function in the second phase. Since she should be indifferent between ϕ and $\overline{\phi}$,

$$\begin{aligned} r\underline{V}(t) &= -c + \overline{\phi} q_L(t) \left(\frac{w}{r} - \underline{V}(t) \right) + \dot{\underline{V}}(t) \\ &= \phi q_L(t) \left(\frac{w}{r} - \underline{V}(t) \right) + \dot{\underline{V}}(t). \end{aligned}$$

Combining the two equations, we find that

$$r\underline{V}(t) = \frac{c\phi}{\overline{\phi} - \phi} + \dot{\underline{V}}(t) \Rightarrow \underline{V}(t) = e^{-r(t-t^*)} \left(\underline{V}(t^*) - \frac{c\phi}{r(\overline{\phi} - \phi)} \right) + \frac{c\phi}{r(\overline{\phi} - \phi)}.$$

Plugging the value function back into the low-type worker's indifference condition,

$$q_L(t) = \frac{c}{\overline{\phi} - \phi} \frac{1}{w/r - \underline{V}(t)} = \frac{c}{\overline{\phi} - \phi} \frac{1}{\frac{w}{r} - \frac{c\phi}{r(\overline{\phi}-\phi)} - e^{-r(t-t^*)} \left(\underline{V}(t^*) - \frac{c\phi}{r(\overline{\phi}-\phi)} \right)}.$$

From the definition of $q_L(t)$,

$$q_L(t) = \frac{L(p^c(t) - \underline{p})}{(1-L)p^c(t)} = \frac{L(\phi(1-\underline{p})p(t) - \phi(t)\underline{p}(1-p(t)))}{(1-L)\phi(t)(1-p(t))}.$$

Use value-matching to denote $\underline{V}(t^*) = \overline{V}(t^*) = V(t^*)$. Combining these yields

$$\phi(t) = \frac{p(t)}{1-p(t)} \frac{\frac{L}{1-L}\phi(1-\underline{p}) \left(\frac{w}{r} - \frac{c\phi}{r(\overline{\phi}-\phi)} - e^{-r(t-t^*)} \left(V(t^*) - \frac{c\phi}{r(\overline{\phi}-\phi)} \right) \right)}{\frac{d}{\overline{\phi}-\phi} + \frac{L}{1-L}\underline{p} \left(\frac{w}{r} - \frac{c\phi}{r(\overline{\phi}-\phi)} - e^{-r(t-t^*)} \left(V(t^*) - \frac{c\phi}{r(\overline{\phi}-\phi)} \right) \right)}. \tag{15}$$

Defining t^* **and** $V(t^*)$ Smooth pasting of the value function and the variable change $x^* \equiv \beta e^{kt^*} + 1$ give us

$$\begin{aligned} -re^{-r(t-t^*)} \left(V(t^*) - \frac{c\phi}{r(\overline{\phi} - \phi)} \right) &= \frac{-(c+w)k^2\beta e^{kt^*} + c_1 e^{t^*(k+r)}(k + r + r\beta e^{kt^*})}{r(k+r)(\beta e^{kt^*} + 1)^2} \\ \Leftrightarrow V(t^*) &= \frac{x^* rc \left(\frac{\phi}{\overline{\phi}-\phi} - 1 \right) + kw}{r(k + 2rx^*)}. \end{aligned}$$

Because $\overline{\phi} = \phi(t^*)$, we have

$$\overline{\phi} = \frac{p(t^*)}{1-p(t^*)} \frac{\frac{L}{1-L}\phi(1-\underline{p}) \left(\frac{w}{r} - V(t^*) \right)}{\frac{c}{\overline{\phi}-\phi} + \frac{L}{1-L}\underline{p} \left(\frac{w}{r} - V(t^*) \right)} \text{ where } \frac{p(t^*)}{1-p(t^*)} = \frac{\phi\overline{p}(1-\underline{p}) - \overline{\phi}\underline{p}(1-\overline{p}) + \overline{\phi}\underline{p}(1-\overline{p})x^*}{\phi(1-\underline{p})(1-\overline{p})x^*}.$$

Rearranging,

$$V(t^*) = \frac{w}{r} - \frac{1-L}{L}\frac{\phi(1-\overline{p})c}{(\overline{\phi}-\phi)(\phi\overline{p}(1-\underline{p})-\overline{\phi}\underline{p}(1-\overline{p}))}x^* = \frac{w}{r} - \frac{\overline{\phi}c}{k(\overline{\phi}-\phi)}x^*.$$

Together, these yield

$$x^* = \frac{k(w(\overline{\phi}-\phi)-\phi c)}{rc\overline{\phi}} \text{ and } V(t^*) = \frac{c\phi}{r(\overline{\phi}-\phi)}.$$

Plugging this back into (15) yields

$$\phi(t) = \frac{p(t)}{1-p(t)}\phi D \text{ where } D = \frac{\frac{L}{1-L}(1-\underline{p})\left(\frac{w}{r}-\frac{c\phi}{r(\overline{\phi}-\phi)}\right)}{\frac{c}{\overline{\phi}-\phi}+\frac{L}{1-L}\underline{p}\left(\frac{w}{r}-\frac{c\phi}{r(\overline{\phi}-\phi)}\right)}.$$

Characterization of Remainder of Equilibrium Note that x^* defines t^*. Then, the evolution of $p(t)$ is given by

$$\dot{p}(t) = -p(1-p)\left(\frac{\overline{p}}{1-\overline{p}}-\frac{p}{1-p}\right)\frac{L}{1-L}\phi(1-\underline{p})\left(\frac{c}{\overline{\phi}-\phi}-(1-\underline{p})\frac{L}{1-L}\left(\frac{w}{r}-\frac{c\phi}{r(\overline{\phi}-\phi)}\right)\right),$$

which leads to

$$p(t) = \frac{\overline{p}}{1+\frac{\overline{p}-p(t^*)}{p(t^*)}e^{A\overline{p}/(1-\overline{p})(t-t^*)}} \quad \text{where} \quad A = \frac{L\phi(1-\underline{p})}{1-L}\left(\frac{c}{\overline{\phi}-\phi}-(1-\underline{p})\frac{L}{1-L}\left(\frac{w}{r}-\frac{c\phi}{r(\overline{\phi}-\phi)}\right)\right).$$

Third Phase: $t > t^{**}$. The condition that $\phi = \phi(t^{**})$ directly defines t^{**}. Beliefs can be found from plugging $\phi(t) = \phi$ into (14), which yields

$$p(t) = \overline{p} - (\overline{p} - p(t^{**}))e^{(t-t^{**})\frac{L}{1-L}\frac{\phi(1-\underline{p})}{1-\overline{p}}}.$$