Unemployment Insurance and the Quality of Job Seekers*

Simon Margolin

Princeton University

October, 2024

[CLICK HERE FOR THE LATEST VERSION]

Abstract

This paper investigates the optimal design of unemployment insurance (UI) and explores how UI affects the quality of the job-seeker pool. Using a directed search model with endogenous search effort and asymmetric information between workers and recruiters, I show that higher levels of UI can exacerbate adverse selection by incentivizing low-productivity workers to apply for high-wage jobs and crowd out productive applicants. However, the model reveals a mitigating effect: increased UI reduces the number of applications from low-productivity workers, alleviating this adverse selection inefficiency. The model is calibrated using data from the Survey of Consumer Expectations and the Current Population Survey to quantify these two forces and the welfare effects of unemployment insurance, showing that the current level of benefits implemented in the US is close to, but slightly below, the optimal one.

JEL codes: E24; E32; J64; J65

*Simon Margolin: simon.margolin@princeton.edu. I am grateful to John Grigsby, Gianluca Violante and Richard Rogerson for their guidance and support. I thank Jerome Adda, Philippe Aghion, Mark Aguiar, Narek Alexanian, Emma Bursztejn, Nobuhiro Kiyotaki, Lukas Mann, Eugenia Menaguale, George Nikolakoudis, Bernardo Ribeiro, Alejandro Sabal, Aysegul Sahin, Thomas Sargent, Karthik Sastry, Andreas Schaab, John Sturm Becko, Jeffrey Sun, Chansik Yoon, Alicia Zhang, Owen Zidar and Thomas Zuber for their comments.

1 Introduction

Unemployment insurance (UI) plays a critical role in modern labor market policies. A well-established literature has examined the optimal design of unemployment insurance, identifying a key trade-off: while UI helps workers smooth consumption during periods of unemployment, it also introduces inefficiencies by distorting their job search effort. Prior studies have primarily focused on the quantitative aspect of this distortion, showing that job seekers tend to search less when they are more generously insured. This paper introduces an equilibrium search model that extends this discussion by exploring a new dimension: how unemployment insurance distorts the quality of the pool of job seekers.

This new distortion arises from two distinctive features of the model. First, job seekers choose both the number and the type of jobs they apply to. Second, worker productivity is private information, making it difficult for employers to separate out high- and low-productivity applicants. These two elements interact, as applications from low-productivity workers may crowd out the productive ones if the former send too many of them or send them too high. Unemployment insurance distorts both margins, and therefore the extent to which productive workers are crowded out of the market.

This paper argues that higher levels of unemployment insurance can improve matching efficiency through a previously unexplored channel. By reducing the intensity of search effort among low-productivity workers and limiting their excessive applications, unemployment insurance can mitigate employers' informational constraints and alleviate the adverse selection force that drives high-productivity workers out of the market.

I develop a dynamic directed search model in which firms post wages, and unemployed workers choose both the number and the directions of their job applications. Employers targeting high-productivity workers face the challenge of distinguishing them from low-productivity applicants, as productivity is not directly observable. They attract the best applicants by leveraging the greater patience of high-productivity workers, who have better outside options. To do so, they post jobs in markets with low job-finding rates, creating conditions that deter more impatient, low-productivity workers from applying.

This difference in patience or, equivalently, in tolerance for unemployment risk, is affected by the level of unemployment insurance. As UI increases for low-productivity workers, they become more willing to tolerate risk in their applications and apply to markets with lower job-finding rates, thereby crowding out high-productivity workers and forcing them to accept even lower job-finding rates to signal their type.

Yet, this higher level of unemployment insurance has an additional, opposite effect on high-productivity workers. As low-productivity job seekers become more willing to tolerate unemployment risk, they reduce the number of applications they send, thereby decreasing the likelihood that their applications will reach high-productivity markets. This channel mitigates adverse selection and its crowding-out effect on high-productivity applicants.

I characterize these different forces and show how they transform the traditional trade-off of unemployment insurance design. To do so, I solve the problem of a Ramsey planner who chooses the level of unemployment benefits in the decentralized equilibrium, subject to the UI budget constraint. I examine the new mechanisms by extending the decomposition of welfare gains from UI introduced by Baily (1978) and Chetty (2006). New terms appear in this decomposition, highlighting the direct impact of insuring low-productivity workers on high-productivity applicants, as well as its indirect externality on the fiscal burden of the unemployment system.

The ultimate impact of these forces on the optimal design of unemployment insurance is a quantitative question. To address it, I extend the model to incorporate aggregate productivity shocks, a key driver of unemployment risk. I show that the block recursivity of the directed search framework (Menzio and Shi, 2010, 2011) generalizes to simultaneous search, i.e. to a directed search equilibrium in which workers are allowed to visit multiple markets. This property allows me to solve for the equilibrium with aggregate risk relatively quickly. Interestingly, the interaction of endogenous search effort and adverse selection produces large unemployment fluctuations in response to aggregate productivity shocks. This is an important aspect of this framework given the difficulty for search models to replicate the employment dynamics observed in aggregate data (Shimer, 2005).

I calibrate the model using data from the Survey of Consumer Expectations Job Search Supplement (SCE-JS) (Faberman et al., 2022) and the Current Population Survey (CPS), and assess the welfare effects of unemployment insurance on workers. The quantitative exercise demonstrates that the two forces introduced in the model—adverse selection and endogenous search intensity—significantly affect the design of optimal unemployment insurance. While adverse selection and endogenous search intensity each independently push for less generous unemployment insurance, their interaction counteracts these effects by discouraging low-productivity workers from over-applying. I find an optimal level of unemployment benefits equal to 47.4% of workers' average wage, which is close to the replacement income observed in most US states, although slightly above the average.

This paper contributes to the literature on optimal unemployment insurance. Baily (1978) and Chetty (2006) derive a key trade-off for this insurance at the microeconomic level: consumption-smoothing benefits vs. moral hazard costs. These forces have been described in equilibrium frameworks with random search (Gautier et al., 2016; Landais et al., 2018) and directed search Acemoglu and Shimer (1999, 2000); Schmieder and Von Wachter (2016); Chaumont and Shi (2022); Cai and Heathcote (2023), which capture the congestion effect of job seekers' behavior. I contribute to this literature by examining how unemployment insurance affects both the quality of the job seeker pool and the informational constraints faced by recruiters. This framework incorporates the traditional moral hazard and congestion effects of UI, allowing for a comparison with the new externalities. Additionally, I quantify these channels in an economy characterized by an endogenous number of applications, asymmetric information, and aggregate risk.

This paper also builds on the literature on adverse selection in the labor market. Blanchard and Diamond (1994) and Engbom (2021) demonstrate the importance of asymmetric information between job seekers and employers for unemployment dynamics. Guerrieri (2008), Guerrieri et al. (2010), Guerrieri and Shimer (2014) and Chang (2018) lay the foundations for directed search models with private information. Davoodalhosseini (2019) analyzes optimal tax policies in a similar environment in which applicants can only visit one market at a time. The framework developed in this paper builds on Auster et al. (2022) who introduce multiple applications, or simultaneous search (Burdett and Judd, 1983; Chade and Smith, 2006; Albrecht et al., 2006; Gautier and Wolthoff, 2009; Kircher, 2009; Galenianos and Kircher, 2009; Wolthoff, 2018; Albrecht et al., 2020), and extends it by incorporating an endogenous number of applications, dynamics, aggregate risk, and by exploring the distortions associated with the unemployment insurance in this type of economies.

The paper is organized as follows. Section 2 describes a steady-state version of the model. Section 3 characterizes the effects of unemployment insurance on workers' behavior and welfare, and the problem of a planner choosing the optimal level of unemployment benefits. Section 4 extends the model to a dynamic environment with aggregate productivity shocks and calibrates the model using data from the SCE-JS and the CPS. Finally, section 5 solves for the optimal level of insurance under multiple scenarios about the tools available to the government.

2 Model

This section develops an equilibrium directed search model in which unemployed workers choose the number and type of jobs to apply for. Firms do not observe directly whether applicants' productivity but post truth-revealing contracts to separate out good and bad candidates. The model is described in steady state, although section 4.1 extends it to an environment with aggregate risk.

2.1 Environment

Time is discrete and all agents discount future payoffs at rate β. The labor market consists of three types of agents: homogeneous firms owned by a representative entrepreneur, low-productivity workers (L) and high-productivity workers (H). There is mass of workers of size one, among which a share m^H is of type H and a share $1 - m^H$ of type L. A match between an employer and a low-productivity worker produces z units of output in every period of activity, while a match with a high-productivity worker produces $z(1 + \xi)$. The parameter $\xi > 0$ governs the difference in productivity between types. I assume that the goods produced by these matches are sold on a competitive market with a price normalized to one, such that each firm's output is also its stream of revenue.

Employers seek workers by creating vacancies, at cost κ. Each of these vacancies has a posted wage which is constant throughout the duration of a match. The set of posted wages is denoted by $\mathcal{W}$. Every posted wage w defines a market in which vacancies and applications meet through a matching function that will be described in the next section. Unemployed workers look for jobs by sending applications to these markets. They can send as many applications as they want, and each of their applications can be directed towards a different type of job.

2.2 Matching

Every period, employers post vacancies and their associated wages, and unemployed workers direct their applications towards these different jobs. A market w is defined by all the vacancies that offer this wage and all the applications directed towards such postings.

Unemployed workers can send multiple applications *simultaneously*, to different types of jobs. Consequently, when a worker gets an offer from a job to which she applied, she may simultaneously obtain a better offer from another job to which she applied. As is common in the simultaneous search literature[1], I assume that two employers that make an offer to the same worker cannot compete ex post and commit to the wage they posted. Hence, a worker with multiple offers in hand accepts the one with the best wage and turns down the others.

As in Kircher (2009), an employer whose offer is turned down can immediately contact its next applicant. This assumption is named *perfect recall*[2]. *Effective* applications in the

[1]Notable exceptions include Albrecht et al. (2006) and Gautier and Wolthoff (2009) who introduce ex-post competition between firms that match with the same worker.

[2]Galenianos and Kircher (2009) build a model in which employers make at most one offer every period, and are thus left without a valid candidate as soon as their offer is turned down. Wolthoff (2018) develops a framework that nests these two extreme cases, i.e. perfect recall of applicants and no recall at all.

market for jobs offering a wage w are the ones from workers who were unsuccessful in their applications to better wages. With a large number of workers, the effective queue in the market for jobs of type w is its ratio of effective applications to vacancies, or

$$q(w) = \underbrace{\frac{\#\text{ applicants}}{\text{vacancies}}}_{\text{gross queue length}} \times Pr[\text{no better offer}]$$

Within a market, workers and firms cannot coordinate their actions and are matched through the matching function $m(q(w))$. This function is continuous, increasing and concave in $q \in \mathbb{R}$. I assume that $m(0) = 0$. The probability that a vacancy offering wage w meets an applicant willing to accept the job is

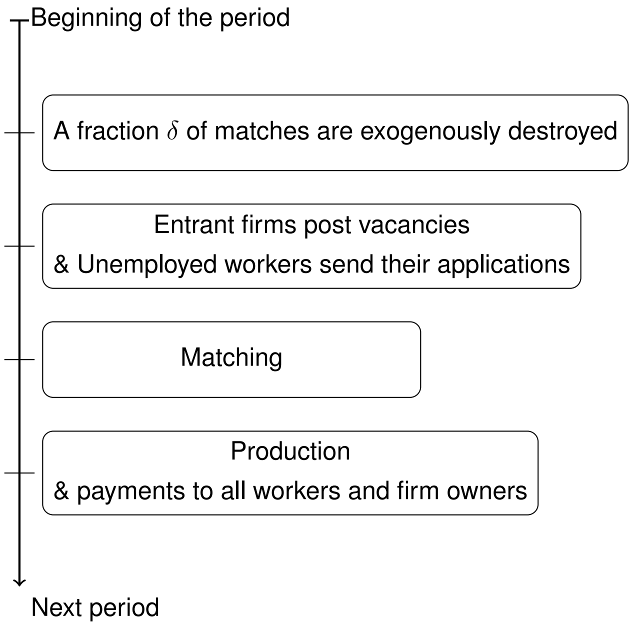
$$\mu\Big(q(w)\Big) = m\Big(q(w)\Big)$$

The probability that an unemployed applicant to market w gets an offer from this market is

$$\psi\Big(q(w)\Big) = \frac{m\Big(q(w)\Big)}{q(w)}$$

2.3 Timing

In every period, a fraction $\delta \in (0, 1)$ of existing matches is exogenously destroyed. Workers who lose their jobs must wait for the next period to be able to search for a new one. Firms losing their worker exit the market. Then, entrant firms post vacancies and their associated wages, and unemployed workers who lost their jobs in previous periods send their applications to the markets defined by these vacancies. The matching process described in the previous section occurs. The newly formed matches, as well as the incumbent matches that were not destroyed at the beginning of the period, produce, pay the promised wage to their worker, and give the remaining revenues to the representative firm owner. Finally, all unemployed workers receive their unemployment benefits.



2.4 Workers

Workers are hand-to-mouth and consume their entire disposable income every period. A worker of type $i \in \{L, H\}$, employed at wage w at the production stage (end of the period), has a lifetime utility

$$W^i(w) = v\Big((1-\tau)w\Big) + \beta\Big[\delta U^i + (1-\delta)W^i(w)\Big]$$

where $v(.)$ is a continuous and twice differentiable utility function such that $v'(.) \geq 0, v''(.) \leq 0$.

At the production stage, a worker is paid wage w, as prescribed by her contract, from which a fraction τ is taxed to fund the unemployment insurance system. Future payoffs are discounted at rate β. At the beginning of the next period, this job is exogenously destroyed with probability δ, in which case she obtains the value of unemployment corresponding to her type, U^i. If the match survives, she gets the value of this job next period.

An unemployed worker of type i has a lifetime utility, at the production stage, equal to

$$U^i = v\Big(b^i + b_0^i\Big) + \beta \max_a \Big\{S^i(a) - c(a)\Big\}$$

The government provides an unemployment benefit b^i, potentially specific to each type. The assumption of a type-specific unemployment benefit relies on the idea that the unemployment insurance system has more information about workers' employment history than recruiters and can therefore better identify their type. In the quantitative part of this paper, I compare results with and without this assumption, i.e. with type-specific and common benefits. In addition to this unemployment benefit, a worker receives an exogenous income b_0^i, $i \in \{L, H\}$, which is specific to each type[3].

In the next period, this unemployed worker will be able to search for a job. She will choose a discrete number a of applications to send simultaneously in order to maximize the net gains from searching. $c(a)$ is a convex cost function that satisfies $c(0) = 0$. $S^i(a)$ denotes the value of the optimal portfolio of a applications chosen by a worker of type i. It is the utility derived from the expected outcome of this strategy. Workers observe the posted wages $w \in \mathcal{W}$ and their associated queues $q(w)$, and choose the markets to which they send each of their a applications.

Without loss of generality, I order these applications from the one that offers the lowest wage to the one that offers the highest, that is $w_1 \leq w_2 \leq ... \leq w_a$. Once recruiters respond to these applications, a worker first looks at the result of her preferred application, w_a, which turns into a job offer with probability $\psi(q(w_a))$. If she does not get an offer from this preferred market, she goes to her second preferred application, w_{a-1}, which is successful with probability $\psi(q(w_{a-1}))$. Thus, the unconditional probability that this worker will end up in a job posted in market $a-1$ is $(1 - \psi(q(w_a)))\psi(q(w_{a-1}))$. More generally, the probability that a worker accepts an offer responding to its k-th, $k \in \{1, ..., a\}$, application is $\prod_{j=k+1}^{a} \Big[1 - \psi(q(w_j))\Big] \psi(q(w_k))$. If none of her a applications is successful, this worker remains unemployed. This outcome occurs with probability $\prod_{k=1}^{a} \Big[1 - \psi(q(w_k))\Big]$. Hence the value of the optimal portfolio of a applications sent by a worker of type i is

$$S^i(a) = \max_{(w_1, \ldots, w_a) \in \mathcal{W}^a} \left[\sum_{k=1}^{a} \underbrace{\left(\prod_{j=k+1}^{a} \Big[1 - \psi(q(w_j))\Big] \right) \psi(q(w_k))}_{\Pr\big[k\text{-th lowest app} = \text{first success}\big]} W^i(w_k) \; + \; \underbrace{\prod_{k=1}^{a} \Big[1 - \psi(q(w_k))\Big]}_{\Pr\big[\text{no job offer}\big]} U^i \right]$$

Following Kircher (2009) and Galenianos and Kircher (2009), one can formulate this portfolio choice recursively as

$$S^i(k) = \max_{w \in \mathcal{W}} \quad \psi(q(w)) W^i(w) + \Big[1 - \psi(q(w))\Big] S^i(k-1) \tag{1}$$

where $S^i(0) = U^i$. When choosing the direction of her k-th application, a worker maximizes the value of its expected outcome. With probability $\psi(q(w))$, this application provides this worker with a job in the target market, which yields a lifetime utility $W^i(w)$. With

[3]This exogenous income is meant to capture the fact that high-productivity workers may benefit from other sources of replacement income, on top of the government-provided benefits.

complementary probability $1 - \psi(q(w))$, this k-th application is unsuccessful and the worker gets her effective outside option, $S^i(k-1)$, which is the utility derived from her lower ranked applications. One can already notice that, in general, each application is directed to a different market because the fallback option $S^i(k-1)$ varies across k.

2.5 Firms

All firms are owned by a risk-neutral representative entrepreneur and post vacancies to maximize the discounted sum of expected profits. Firms make two choices: whether to post a vacancy and with what wage.

At the production stage, the discounted sum of expected profits generated by a job filled with a worker of type i is

$$J^i(w) = z(1 + \xi \cdot \mathbb{1}[i = H]) - w + \beta(1-\delta)J^i(w)$$

The flow profits from this match are equal to its output $z(1 + \xi \cdot \mathbb{1}[i = H])$ minus its labor cost w. At the beginning of the next period, the match is destroyed exogenously with probability δ. In this case, the job's value to the employer becomes null. If the match survives this exogenous shock, the firm obtains the value of next period's discounted profits, $J^i(w)$.

An entrant firm creates a vacancy at cost κ and chooses a wage w to solve

$$V = \max_{w} \quad \mu(q(w)) \cdot \big[\gamma(w)J^L(w) + (1-\gamma(w))J^H(w)\big] - \kappa \tag{2}$$

where V denote the value of an unfilled vacancy, $\mu(q(w))$ is the probability that this vacancy is filled, and $J^i(w)$ is the employer's discounted sum of profits once the job is filled with a worker of type $i \in \{L, H\}$. Vacancies that are not filled in the period of their creation exit the market.

$\gamma(w) \in [0,1]$ denotes the share of low-productivity workers in the pool of applicants for jobs offering a wage w. Recruiters know $\gamma(w)$ when they decide whether or not to post on market w. However, they cannot directly observe the productivity of the workers they meet. Consequently, they cannot condition their hiring and the wage they offer on the type of applicant they face.

2.6 Government

The government taxes all labor income at rate τ to pay for the benefits b^H and b^L paid to each H-type and L-type workers in every period of unemployment. I assume that the government budget is balanced, such that we can write the steady-state government budget constraint as

$$m^H \mathtt{u}^H b^H + (1 - m^H)\mathtt{u}^L b^L = \tau\Big(m^H(1-\mathtt{u}^H)\mathbb{E}\left[w^H\right] + (1-m^H)(1-\mathtt{u}^L)\mathbb{E}\left[w^L\right]\Big) \tag{3}$$

where m^H is the share of workers of type H in the overall workforce, $\mathtt{u}^i$ denotes the unemployment rate for workers of type $i \in \{H, L\}$, and $\mathbb{E}[w^i]$ denotes the average wage of workers of type $i \in \{H, L\}$. The left-hand side of equation (3) is the amount of benefits paid by the unemployment insurance every period. The right-hand side is the amount of revenues raised through the labor income tax.

2.7 Beliefs

For workers and firms to be able to make the choices described above, they must form beliefs about the size $q(w)$ and the composition of the pool $\gamma(w)$ of applicants they should expect for any potential wage, even if it is not posted in equilibrium, i.e. if $w \notin \mathcal{W}$.

First, these beliefs need to be consistent, as unemployed workers and entrant firms must expect the same application pool for any wage. Secondly, I follow the standard *market utility* approach (Wright et al., 2021) in the directed search literature, extended to frameworks with private information by Guerrieri et al. (2010) and Auster et al. (2022), for agents' out-of-equilibrium beliefs. When a firm considers the possibility of posting a wage $w \notin \mathcal{W}$, it expects the queue to be the one that requires the lowest job-finding probability capable of redirecting at least one application to this wage. More formally, the expected queue $q(w)$ is such that

$$S^i(k) \geq \psi(q(w))W^i(w) + \Big[1 - \psi(q(w))\Big]S^i_(k-1) \tag{4}$$

holds with equality for at least one type of worker i and one of her applications k, and with weak inequality for all $i \in \{H, L\}$ and $k \leq a^i$.

One way to understand this assumption is to consider it as a process of "tâtonnement", as described in Guerrieri et al. (2010). When a firm considers a deviation from equilibrium wages to $w \notin \mathcal{W}$, it first expects an infinitely long queue because it is the only firm in this hypothetical market and that many workers may prefer this wage to their equilibrium expected outcome. However, this long queue discourages job seekers, and they remove their applications from the queue up to the point where some of them are actually willing to continue applying to this market.

As explained by Guerrieri et al. (2010) and Auster et al. (2022), the *market utility* approach also determines the composition of the pool of applicants that firms expect for any out-of-equilibrium wage. The queue with the lowest job-finding probability needed to redirect an application can generally attract applications from a single type of worker. It only attracts workers of type L, $\gamma(w) = 1$, when condition (4) holds with equality for $i = L$ and for one $k \leq a^L$, and with strict inequality for $i = H$ and all $k \leq a^H$. Similarly, it only attracts workers of type H, $\gamma(w) = 0$, when it holds with equality for $i = H$ and for one $k \leq a^H$, and with strict inequality for $i = L$ and all $k \leq a^L$.

2.8 Separating equilibrium

2.8.1 Definition

A separating equilibrium is defined as a set of wages $\mathcal{W}$, a queue function $q(w)$, a market composition function $\gamma(w)$, numbers of applications a^i and values of unemployment U^i for $i \in \{H, L\}$, application portfolio values $S^i(k)$ for $i \in \{H, L\}$ and $k \in \{1, ..., a^i\}$ such that

1. *Worker optimization*: workers choose a number of applications a^i that maximizes their net value of searching, and a subset of the posted wages towards which they direct their a^i applications to solve (1).
2. *Firm optimization*: firms post wages that solve (2).
3. *Free entry*: firms enter in each market w until $V(w) = 0$.
4. *Incentive compatibility*: firms post wages that attract a single type of worker, i.e. such that $\gamma(w) \in \{0, 1\}$, by posting cream-skimming wages as defined below.
5. *Beliefs*: Workers and firms have beliefs consistent beliefs about the equilibrium functions $q(w)$ and $\gamma(w)$.
6. *Lemons condition*: unemployment delivers a higher lifetime utility to high-productivity workers than the highest wage that can be offered to low-productivity workers, or $U^H > W^H(z - \kappa(1 - \beta(1 - \delta))$.

2.8.2 Construction

In a separating equilibrium, each vacancy targets a specific type of worker. The *lemons condition* ensures that, in equilibrium, high-productivity workers are not willing to apply to vacancies that target low-productivity workers[4]. The wages posted for L-type workers and

[4] In practice, the estimated values of b_0^H and ξ ensure that the *lemons condition* holds with strict inequality. Auster et al. (2022) show that pooling equilibria exist when this condition is relaxed. It would be

their associated queues are therefore those that satisfy the first-order condition of L-type workers' application problem (1) and set the value of a vacancy (2) with $\gamma(w) = 1$ equal to zero (due to free entry). These conditions define the optimal portfolio $(w_1^L, ..., w_{a^L}^L)$, its associated queues $(q(w_1^L), ..., q(w_{a^L}^L))$, and the value of this portfolio $S_{a^L}^L$ for any $a^L \geq 0$. The number of applications sent by low-productivity workers is the one that maximizes the net gains from searching $S_{a^L}^L - c(a^L)$.

However, the markets that target H-type workers must propose wage-queue combinations that are unattractive to L-type applicants. Nothing prevents a low-productivity worker to send one of her applications to the H-type market. First, a firm that targets H-type workers chooses a wage that does not attract any of the existing applications sent by L-type workers, that is, a w such that the following incentive compatibility constraint is satisfied

$$\psi(q(w))W^L(w) + [1 - \psi(q(w))]S_{a^L-k}^L \leq S_{a^L-k+1}^L \quad \text{for any } k \in \{1, ..., a^L - 1\}$$

The left-hand side of this inequality represents the value for a worker of type L of redirecting her $a^L - k + 1$-th application to the considered posting w. This worker would receive an offer from this market with probability $\psi(q(w))$ and would fall back on her $a^L - k$ lower applications with complementary probability $1 - \psi(q(w))$. The latter event would provide a lifetime utility $S_{a^L-k}^L$. The right-hand side of this equation represents the value of the equilibrium strategy of a L-type worker when sending $a^L - k + 1$ applications.

As shown by Auster et al. (2022), checking this condition for the highest application sent by L-type workers is sufficient. The reason for this is that this condition exploits L-type workers' aversion to unemployment risk in their application targets. Due to the lemons condition, firms targeting H-type workers post wages that are higher than any wage posted on L-type markets. These jobs are thus more attractive than all the ones L-type workers apply for. However, these postings generate such high queues that L-type workers are not willing to take the risk on redirecting one of their applications to these markets with relatively low job-finding probabilities. H-type workers are still willing to apply because they have a higher value of unemployment and have more tolerance to risk in their application portfolio.

Through the same mechanism, L-type workers tolerate more risk in their highest application than in any other application. The fallback option for application $a^L - k$ is the value associated with the $a^L - k - 1$ lower applications. This fallback option increases with k, such that the highest application, indexed by a^L, is also the one that has the highest outside

interesting to understand the impact of the unemployment insurance in this type of equilibrium, although this is beyond the scope of the present paper.

option, and the highest risk tolerance. If the incentive compatibility constraint is satisfied for this highest application, it means that L-type workers find the posted wage w and its associated queue $q(w)$ too risky for it to be worth redirecting this application. Following this reasoning, the fact that they are even more risk averse for their lower applications implies that this wage would not be attractive for any of these lower applications neither. Hence, if the incentive compatibility constraint is satisfied for $k = 1$, it is satisfied for any $k \in \{1, ..., a^L - 1\}$.

This incentive compatibility constraint can be reformulated as

$$\psi(q(w))W^L(w) + [1 - \psi(q(w))]S^L_{a^L-1} \leq S^L_{a^L} \tag{IC1}$$

A second incentive compatibility constraint emerges from workers' endogenous number of applications. Even if they do not have any incentive to redirect their current applications to one of the H-type markets, L-type workers could still decide to send an additional application, indexed by $a^L + 1$, to one of these markets. To ensure that L-type workers do not make such a decision, firms targeting high-productivity workers post wages that satisfy

$$\psi(q(w))\left[W^L(w) - S^L_{a^L}\right] \leq c(a^L + 1) - c(a^L) \tag{IC2}$$

The left-hand side of this constraint is the additional utility derived by L-type workers from an additional application to market w, while the right-hand side is its cost.

3 Optimal unemployment insurance

3.1 Ramsey problem

I follow the tradition of the optimal unemployment insurance literature (Acemoglu and Shimer, 1999; Schmieder and Von Wachter, 2016; Cai and Heathcote, 2023; Le Barbanchon et al., 2024) in considering the problem of a constrained planner who maximizes the expected utility of a worker starting unmatched before the search and matching stage. The social welfare function is

$$SWF = m^H(S^H(a^H) - c(a^H)) + m^L(S^L(a^L) - c(a^L)) \tag{5}$$

I assume that the Ramsey planner is constrained by the decentralized equilibrium described in the previous section, but can choose type-specific and time-invariant unemployment benefits b^H and b^L satisfying the long-run government budget constraint defined in 2.6.

3.2 Exogenous number of applications: an extended Baily-Chetty formula

With an exogenous number of applications, i.e. fixed a^L and a^H, we can differentiate the social welfare function with respect to the policy instruments b^H and b^L. Search in still endogenous through the directions workers choose for their applications. However, the intensity margin is shut down.

The planner's first order condition with respect to the unemployment benefit for workers of type L, b^L, is

$$\underbrace{m^L\left[\frac{\partial S^L}{\partial b^L}+\frac{\partial S^L}{\partial \tau}\frac{\partial \tau}{\partial b^L}\right]}_{\text{Insurance}}+\underbrace{m^L\frac{\partial S^L}{\partial \tau}\left(\sum_{k=1}^{a^L}\frac{\partial \tau}{\partial w_k^L}\frac{dw_k^L}{db^L}\right)+m^H\frac{\partial S^H}{\partial \tau}\left[\frac{\partial \tau}{\partial b^L}+\sum_{k=1}^{a^L}\frac{\partial \tau}{\partial w_k^L}\frac{dw_k^L}{db^L}\right]}_{\text{Moral hazard fiscal externality}}$$
$$+\underbrace{m^H\frac{\partial S^H}{\partial \underline{w}^H}\frac{d\underline{w}^H}{db^L}}_{\text{Adverse selection direct effect}}+\underbrace{\left(m^L\frac{\partial S^L}{\partial \tau}+m^H\frac{\partial S^H}{\partial \tau}\right)\frac{\partial \tau}{\partial \underline{w}^H}\frac{d\underline{w}^H}{db^L}}_{\text{Adverse selection fiscal externality}}=0 \quad (6)$$

The planner balances consumption-smoothing benefits from the unemployment insurance with efficiency costs which arise from the distortions it introduces. Insurance benefits are the gains L-type workers get from transferring a marginal unit of income from employed workers to unemployed workers at constant aggregate revenues, and correspond to the first term in brackets in (6).

3.2.1 Moral hazard and the traditional cost of unemployment insurance

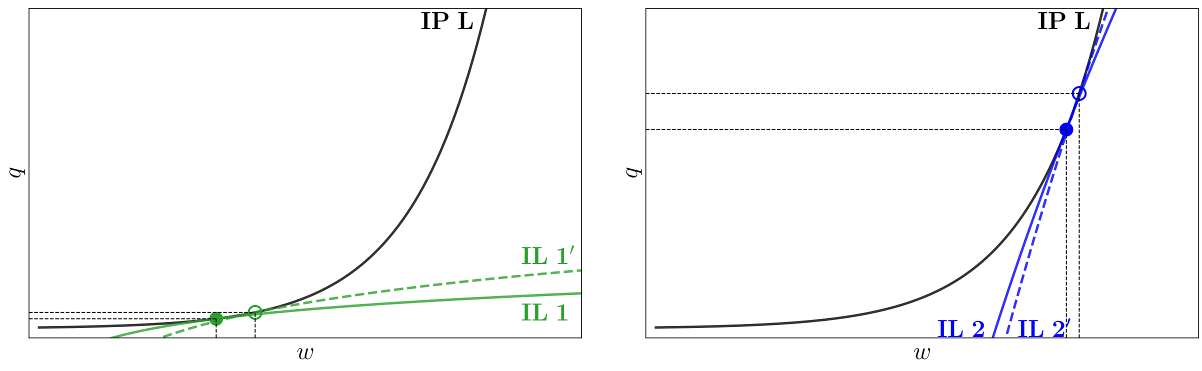
A first distortion results from the fact that unemployment benefits cannot be conditioned on the effort job seekers put into their search. Unemployed workers are effectively more tolerant to idiosyncratic risks associated with the search for a job when they are better insured. As a result, workers of type L target higher wages when they are better insured, as b^L increases. The other side of the market tailors its vacancies to respond to this shift by offering contracts that feature higher wages and a higher unemployment risk, in the form of longer queues. This moral hazard is what Acemoglu and Shimer (1999) coined "*market-generated moral hazard*".

The second and third terms of the first line of equation (6) generalize this efficiency cost to an economy with heterogeneous agents and multiple applications. The *moral hazard fiscal externality* can be decomposed into an externality within type L (second term), affecting the workers that are better insured, and an externality across types (third term), affecting workers who do not experience any change in their unemployment benefits but pay higher taxes as they contribute to the same system.

Increasing the unemployment benefit for L-type workers makes the overall unemployment insurance more costly first through a direct effect $\partial\tau/\partial b^L$, as every incumbent job seeker of type L receives higher benefits. Moreover, each of the a^L applications sent by unemployed workers of type L targets a market with higher wages and a lower employment probability. Thus, for each $k \in \{1, ..., a^L\}$, $dw_k^L/db^L > 0$. The fact that workers target higher wages has two effects on the cost of the unemployment insurance: it reduces the job finding rate, which increases the number of workers who claim benefits and decreases the number of workers who contribute through the labor income tax, and it increases the amount of taxes paid per job, as tax revenues are proportional to the level of wages. The first effect, on the mass of unemployed workers, makes the unemployment insurance more costly. The second effect makes it cheaper. The relative strength of these effects depends on which of the wages or queues is more sensitive to a marginal change in the level of unemployment benefits.

Figure 1 represents the equilibrium wage w and queue q for a given application sent by L-type unemployed workers. In this (w, q) space, the solid black line represents the isoprofit curve that satisfies the free entry condition for L-type markets. This isoprofit curve is defined by all the pairs (w, q) that give zero expected profits to a firm that posts a vacancy on a market that attracts L-type workers, i.e.

$$(w, q) \in \mathbb{R}^2 : \mu(q)J^L(w) = \kappa$$



(a) Low insurance (b) High insurance

Figure 1: Wage and queue elasticities depend on the level of insurance

The solid green curve (IL1) on the left panel represents the indifference curve that corresponds to L-type workers market utility for their first application. This indifference curve is defined as all the pairs of wages and queues that deliver a lifetime utility of S_1^L, when the outside option is unemployment, i.e.

$$(w, q) \in \mathbb{R}^2 : \psi(q)W^L(w) + [1 - \psi(q)]U^L = S_1^L$$

In this framework, the equilibrium wage-queue combination for L-type workers' first application is the tangent point between the isoprofit curve and this indifference curve. Note that this solid green indifference curve (IL 1) is relatively flat. This is due to the low level of insurance assumed in the left panel. A low unemployment benefit b^L yields a low value of unemployment U^L, which makes job seekers relatively averse to job finding risk for this lowest application. This aversion translates into a flat indifference curve, which means that they need a large increase in the offered wage to accept a slightly higher queue, i.e. a lower job-finding rate.

Suppose a reform marginally increases the level of unemployment benefits for low-productivity workers b^L. Unemployed workers tolerate more job-finding risk in their application and, in equilibrium, firms cater to this change by posting a higher wage, associated with a higher queue. Panel (a) in Figure 1 shows that, at a low level of insurance, this marginal increase in workers' risk tolerance has a large effect on the equilibrium wage and a relatively small effect on the equilibrium queue. Workers' indifference curve moves from (IL1) to (IL1'), the dashed green line, which drives the equilibrium wage-queue combination to the green circle. This equilibrium movement along the isoprofit curve generates a large change in the wage and a small change in the queue. At this low level of insurance, the queue is initially low and firms do not need a much longer queue to be compensated for an increase in the wage. The elasticity of wages with respect to unemployment benefits exceeds that of queue lengths.

Hence, at low levels of insurance, a marginal increase in unemployment benefits does not generate a large fiscal externality because it pushes workers towards higher wages, and therefore higher taxes, without a large increase in unemployment.

However, Panel (b) shows that a higher initial level of insurance is associated with steeper indifference curves. The resulting application strategy of job seekers places the equilibrium towards longer queues, as represented by the solid blue dot, because they tolerate more unemployment risk in order to reach higher wages. On this part of the isoprofit curve, firms need a larger increase in the queue and in their hiring probability to offset an increase in the wage they post. Therefore, a marginal increase in unemployment benefits has a small impact on the equilibrium wage and a large one on its associated queue. The fiscal externality is large because the employment effect, which reduces tax revenues, is large while the wage effect, which increases tax revenues, is small.

In practice, I calibrate unemployment benefits b^H and b^L to be 46.5% of their expected income when employed, the average replacement ratio in the US (Landais et al., 2018). In addition, both types of workers benefit from other source of replacement income, represented by b_0^H and b_0^L in the model. Therefore, the replacement income of unemployed workers is at least 46.5% of their expected income when employed. This value is high enough for job seekers to be relatively tolerant to long queues for every job they apply to. The endogenous distribution of posted contracts is located on the steeper part of the isoprofit curve represented in Figure 1, and queue lengths are much more sensitive than wages to changes in unemployment benefits.

These observations imply that $\sum_{k=1}^{a^L} \frac{\partial \tau}{\partial w_k^L} \frac{dw_k^L}{db^L}$, in equation (6), is positive. Higher benefits for unemployed workers of type L drive them to look for jobs offering higher, but less attainable, wages. This moral hazard effect makes unemployment insurance more expensive and requires a higher tax rate on labor income. This higher tax rate affects both types of workers negatively, i.e. $\partial S^H/\partial \tau < 0$ and $\partial S^H/\partial \tau < 0$.

The fiscal externality which emerges from moral hazard, represented by the second and third terms in equation (6), is negative, and reduces the planner's incentives to increase the unemployment benefits of workers of type L. This externality nests the fiscal cost depicted in Acemoglu and Shimer (1999). The latter corresponds to the case were $a^L = 1$ and $m^H = 0$.

3.2.2 Adverse selection: the new cost of unemployment insurance

With asymmetric information between applicants and recruiters, a new negative externality appears. The two terms of the second line of equation (6) represent this cost. A change in the level of insurance provided to workers of type L not only affects their search strategy, but also the set of incentive compatible contracts offered to workers of type H. The following proposition shows that an increase in unemployment benefits for workers of type L tightens the incentive compatibility constraint on markets of type H.

Proposition 1. *Suppose workers of type L and H send respectively a^L and a^H applications, and that $(\underline{w}_1^H, \underline{q}_1^H)$ denotes the lowest incentive compatible contract that can attract applications from H-type workers without receiving any application from a worker of type L, and that satisfies the zero-profit condition. Then, a budget-balanced increase in unemployment benefits provided to workers of type L, b^L, drives up the lowest incentive compatible wage and queue length. That is,*

$$\frac{d\underline{w}^H}{db^L} > 0 \quad \textit{and} \quad \frac{d\underline{q}^H}{db^L} > 0$$

With a higher level of insurance, L-type workers are willing to take more risk in their application portfolio. H-type markets, which offer high wages and small employment probabilities, become more attractive. Hence, firms that target H-type workers have to post contracts that provide even higher returns, with even higher risk, to avoid low-productivity workers. A complete proof of proposition 1 is provided in Appendix B.1.

Figure 2 shows the effect of an increase in unemployment benefits for low-productivity workers on the first incentive compatibility constraint that affects the wages posted for high-productivity workers. (IPL) and (IPH) are the isoprofit curves that satisfy the zero-profit condition for firms recruiting low-productivity workers and high-productivity workers. Both are increasing because firms accept to post a higher wage only if they are compensated by a longer queue, that is, a higher probability of filling their vacancy. (IPH) stands on the right of (IPL) because, for every queue q, a firm is willing to post a higher wage for H-type workers because they generate more revenue.

In this figure, I assume that both types of worker send two applications. In Panel (a), (IL1) and (IL2) represent the pairs of wage w and queue q that provide low-productivity workers with their "market utility", i.e. the equilibrium values of a portfolio of one application $S^L(1)$ and two applications $S^L(2)$. More formally, (IL1) is defined as

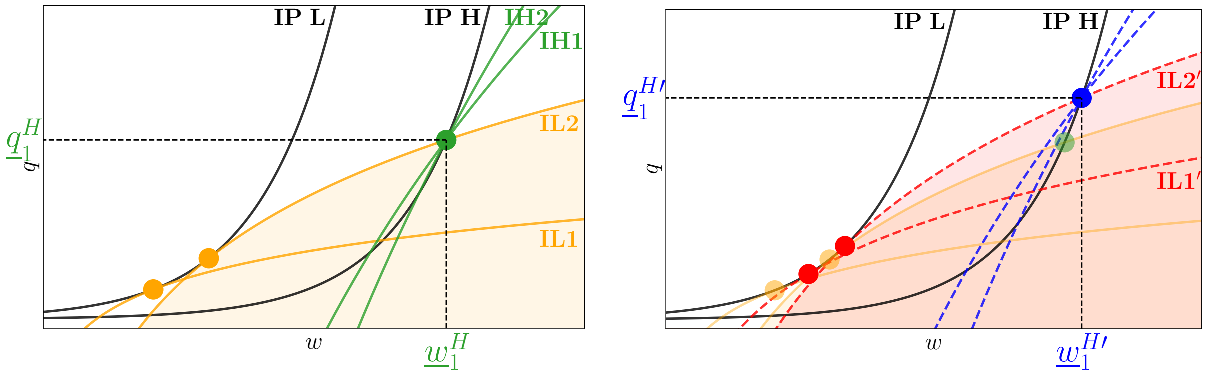
$$(w, q) \in \mathbb{R}^2 : \psi(q)W^L(w) + [1 - \psi(q)]U^L = S^L(1)$$

and (IL2) is defined as

$$(w, q) \in \mathbb{R}^2 : \psi(q)W^L(w) + [1 - \psi(q)]S^L(1) = S^L(2)$$

These two indifference curves have different slopes because each application has a different fallback option. The fallback of the lowest application, indexed by 1, is unemployment, while the fallback option of the second application also includes the expected outcome from the lowest application. Consequently, job seekers of type L are more willing to take some risk in their second-lowest application than in their lowest one. Indifference curve (IL2) is steeper than (IL1), and the corresponding applications are sent to different markets in equilibrium, represented by the two orange dots on Panel (a).

The shaded area under the curve (IL2) shows all the combinations of wage and queue that would be preferred by low-productivity workers compared to the one targeted by their highest application. As a result, any market that attracts H-type applications must be outside of this area. Otherwise, it would violate incentive compatibility. Moreover, because of free entry, we know that its wage-queue pair would be on (IPH), the isoprofit curve for high-productivity workers. Therefore, in equilibrium, any wage targeting H-type workers needs to be higher than w_1^H, the lowest incentive compatible wage. This wage is pinned down by the intersection of (IPH) and low-productivity workers' highest indifference curve



(a) Equilibrium with Asymmetric Information

(b) After an increase in UI benefits for low-productivity workers

Figure 2: Higher insurance for low-productivity workers tightens the constraint on high-productivity workers

(IL2).

I assume in this plot that this constraint is binding, i.e. that high-productivity workers would prefer to apply for jobs with lower wages and lower queues (down the isoprofit curve). Therefore, firms post the lowest incentive compatible wage $\underline{w}_1^H$, and high-productivity workers send both of their applications to this type of job. Hence, low-productivity workers' ability to redirect their applications towards H-type markets, due to the unobservability of their type, constrains H-type workers and forces them to apply to higher wages, at a higher risk.

Panel (b) shows how this constraint responds to an increase in unemployment benefits for low-productivity workers b^L. As explained in the previous subsection, job seekers tolerate more risk in their application portfolio when they are better insured. As a result, firms cater to these preferences by posting higher wages, which entail a higher risk of unemployment. When the incentive compatibility constraint is binding, this change in low-productivity workers' strategy drives up wages and queues on H-type markets. L-type workers' indifference curves become steeper and move to (IL1') and (IL2'). This upward movement constrains the set of incentive compatible wages and leads firms targeting high-productivity workers to post an even higher wage, $\underline{w}_1^{H\prime}$, with an even longer queue, $\underline{q}_1^{H\prime}$.

This upward movement in posted wages has a direct effect on the expected utility of high-productivity workers. This direct negative externality of unemployment insurance for low-productivity workers corresponds to the first term of the second line of equation (6). When the incentive compatibility constraint is binding, H-type workers would prefer to apply to jobs with lower wages and lower queues, but they are constrained to take more risk to reveal their type to recruiters.

In addition to this direct externality, the stronger adverse creates a negative fiscal externality on all workers. This new fiscal externality is characterized by the last term of the left-hand side of equation (6). By forcing high-productivity workers to accept lower job-finding rates, this adverse selection channel increases the mass of unemployed workers receiving benefits and reduces the mass of employed workers contributing to the system. Therefore, the balance of the unemployment insurance system requires a higher tax rate τ on all employed workers, which is detrimental to their lifetime utility S^L and S^H.

3.3 Endogenous search intensity: moral hazard costs vs. moral hazard benefits

The previous subsection showed that a new cost of unemployment insurance springs from the distortion of the directions of job search. In this subsection, I study the effect of unemployment insurance on another dimension of endogenous search: job seekers choose the number of applications they send, or their search intensity.

As for the directional aspect of job search, this intensity margin is distorted by the fact that an increase in unemployment insurance makes workers tolerate more risk in their application strategy. This higher tolerance for job-finding risk leads unemployed workers to send fewer applications.

However, this reduced intensity alleviates the adverse selection friction that the directional aspect of job search was reinforcing. When low-productivity workers send fewer applications, they are less likely to send one of them to a very high wage, associated with a very high queue. Figure 3 shows how the equilibrium responds to an increase in unemployment benefits for low-productivity workers now that they control the intensity of their search effort.

Before this increase in unemployment benefits, low-productivity workers were sending two applications to the markets represented by the orange dots. Now that they are better insured, they send each application to a higher target, but they send only one application. They are willing to take more risk in this application than in their previous lowest application, which explains the movement along the isoprofit curve (IPL) from the lowest orange dot to the red one. This is the new wage-queue combination characterizing the market to which low-productivity job seekers apply.

This change in low-productivity workers' search effort affects high-productivity markets in the following way. As they reach for higher targets, all their indifference curve shift towards riskier applications, crowding out high-productivity workers. On the other hand, they send one less application, which shifts their highest indifference curve to the one corresponding to their lowest, and now only, application. This movement relaxes the incentive compatibility constraint on high-productivity markets and reduces the crowding out of their workers. Figure 3 shows an example in which this second force dominates the first one. The wage posted for high-productivity workers goes down to $\underline{w}_1^{H\prime}$, together with its associated queue $\underline{q}_1^{H\prime}$. These unemployed workers prefer this wage-queue pair (blue dot) to the previous equilibrium (green dot), as shown by the fact that H-type workers' new indifference curves lie to the south-east of the previous equilibrium.

A higher level of insurance for low-productivity workers may therefore alleviate the adverse selection friction which crowds out high-productivity workers.

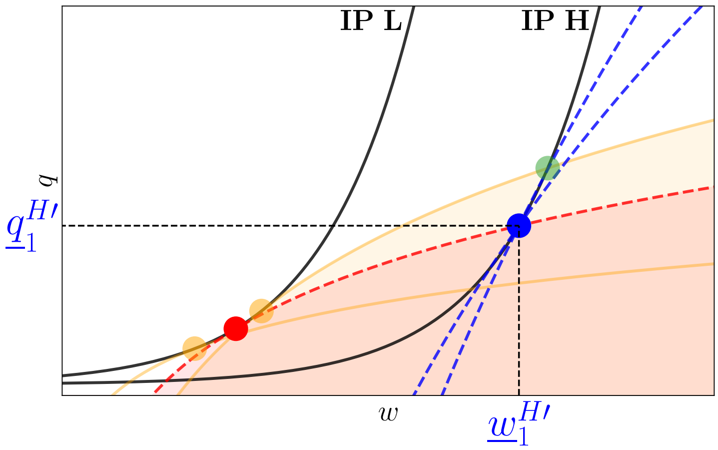


Figure 3: A higher level of insurance relaxes the constraint on high-productivity workers

Moreover, endogenous search intensity introduces an additional incentive compatibility constraint, described in section 2 as (IC2), which ensures that low-productivity workers do not have an incentive to send an extra application to the posted vacancies that try to hire high-productivity workers. The following proposition demonstrates that this constraint is relaxed when unemployment benefits increase for low-productivity workers.

Proposition 2. *Suppose workers of type L and H send respectively a^L and a^H applications, and that $(\underline{w}_2^H, \underline{q}_2^H)$ denotes the lowest contract that can attract applications from H-type workers without receiving any application from a worker of type L, and that satisfies the zero-profit condition. Then, a budget-balanced increase in unemployment benefits provided to workers of type L, b^L, drives down the lowest wage and queue length that satisfy the constraint (IC2). That is,*

$$\frac{d\underline{w}_2^H}{db^L} < 0 \quad \textit{and} \quad \frac{d\underline{q}_2^H}{db^L} < 0$$

See Appendix **??** for a proof. The reason for this relaxation of (IC2) is similar to the one described above for the first incentive compatibility constraint (IC1). When low-productivity workers are better insured, they are willing to take more risk in their application portfolio which, in terms of search intensity, implies that they are willing to send fewer applications. This effect reduces the desirability of an additional application reaching for high-productivity markets.

Endogenous search embeds a directional margin and an intensity margin. When workers are better insured, they aim for higher but fewer targets. Both margins yield the traditional cost of unemployment insurance: workers respond to a higher level of insurance by choosing a search strategy that features a higher risk of unemployment. However, the interaction of these two margins of endegenous search with asymmetric information work in opposite directions. Workers' directional response to a higher level of insurance reinforces adverse selection and crowds high-productivity workers out of the market. This channel increases the cost of unemployment insurance. On the other hand, the lower search intensity generated by an increase in unemployment insurance alleviates adverse selection by reducing low-productivity workers' incentive to send too many applications and crowding out high-productivity workers. This channel reduces the cost of unemployment insurance.

4 Calibration

I calibrate the model in two steps. First, I set some parameters exogenously. In a second step, I jointly estimate the remaining parameters to minimize the distance between moments generated by the model and moments taken from the Survey of Consumer Expectations Job Search Supplement (SCE-JS) and the Current Population Survey (CPS). The former contains a lot of information about workers' search behavior, and more particularly about their application strategies. I use the CPS to measure aggregate worker flows.

One of the moments used to estimate the model is the standard deviation of the unemployment rate over the cycle. In order to generate such fluctuations in the model, I introduce aggregate risk in the form of labor productivity shocks common to all matches, as in Shimer (2005) and Hagedorn and Manovskii (2008). This extension of the model contributes to the literature on unemployment fluctuations by being, to the best of my knowledge, the first model of simultaneous search with aggregate risk. Aggregate risk is important in particular for the design of an optimal unemployment insurance, as a large part of workers' unemployment risk results from aggregate shocks. The next subsection explains what these labor productivity shocks change to the model and how they translate into labor market fluctuations.

4.1 Aggregate risk

I introduce aggregate shocks by assuming that the common component of labor productivity, z_t, is indexed by t and follows an AR(1) process

$$z_t = \rho_z z_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_z)$$

where ρ_z is the persistence parameter of this process and σ_z is its volatility.

Firms post contracts stipulating a fixed wage over the duration of the match. Following Rudanko (2009), I show in Appendix B.3 that, under two-sided commitment, it is the type of contract that firms optimally choose to post even when they are allowed to choose any sequence of wages. Risk-neutral firms maximize profits by offering relatively low wages that are attractive for risk-averse workers due to their insurance properties. Fixed-wage contracts protect employed workers against aggregate shocks that occur throughout the duration of the match. However, these workers are still subject to aggregate risk once their match is destroyed and they have to deal with the aggregate state of the labor market to find a new job.

All equilibrium outcomes are now functions of the aggregate state and are therefore indexed by t. The aggregate state includes the level of productivity z_t and the distribution of workers across employment states and wages. Given that the distribution of posted wages is potentially different in each period, keeping track of the overall distribution of wages is computationally difficult. However, I show, with the following proposition, that the equilibrium is block recursive (Menzio and Shi, 2010, 2011). The only aggregate variable one needs to forecast to compute the value functions and policy functions of workers and firms is labor productivity z_t.

Proposition 3. *The equilibrium is block recursive. One can construct an equilibrium in which value functions, posted wages and queues do not depend on the distribution of workers across employment states and wages.*

See Appendix B.4 for a proof. This useful property is common in directed search frameworks (Menzio and Shi, 2011; Schaal, 2017; Wright et al., 2021). This proposition shows that it extends to simultaneous search environments, which allow workers to search in multiple directions.

Fixed wages imply that the flow of profits generated by a match varies along the cycle. Some firms can make losses if they commited to a high wage in the past and labor productivity undergoes a sufficiently negative shock. These losses are paid by the risk-neutral representative entrepreneur who owns all firms and can borrow from an external fund at a fixed rate r if its net aggregate revenue is negative. In other words, I assume that this is a small open economy.

Similarly, the unemployment insurance system may now suffer deficits during bad times. When a negative shock hits, unemployment rises, which imposes large expenses in benefits and smaller revenues from labor income taxation. The government finances its deficits by issuing one-period bonds at the interest rate r. Let B_t denote the amount of bonds issued in period t. Following the small open economy assumption, these bonds are held by agents outside of the economy. The government budget constraint in each period is

$$\begin{aligned} m^H \mathtt{u}_t^H b^H + (1-m^H)\mathtt{u}_t^L b^L + (1+r)B_{t-1} \\ = \tau\Big(m^H(1-\mathtt{u}_t^H)\mathbb{E}\left[w_t^H\right] + (1-m^H)(1-\mathtt{u}_t^L)\mathbb{E}\left[w_t^L\right]\Big) + B_t \end{aligned} \tag{7}$$

Every period, the government provides unemployment benefits and pays for the outstanding debt by collecting taxes on labor income and issuing new bonds. I assume that the economy starts from period 0 with no debt, i.e. $B_{-1} = 0$. The intertemporal budget constraint is

$$\sum_{t=0}^{\infty} \frac{m^H \mathtt{u}_t^H b^H + (1-m^H)\mathtt{u}_t^L b^L}{(1+r)^t} = \tau \sum_{t=0}^{\infty} \frac{m^H(1-\mathtt{u}_t^H)\mathbb{E}\left[w_t^H\right] + (1-m^H)(1-\mathtt{u}_t^L)\mathbb{E}\left[w_t^L\right]}{(1+r)^t} \tag{8}$$

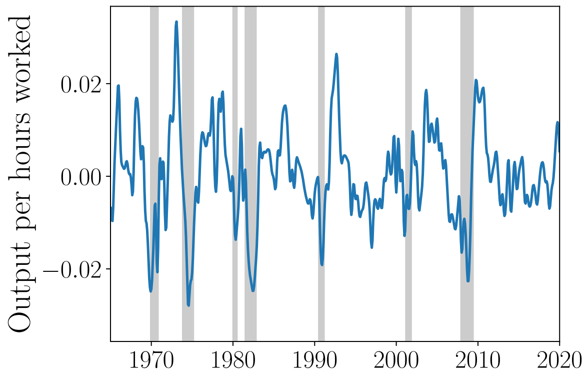
The unemployment insurance system can afford temporary deficits but needs to balance the discounted sum of benefit expenses with the discounted sum of tax revenues.

4.1.1 Amplification of productivity shocks

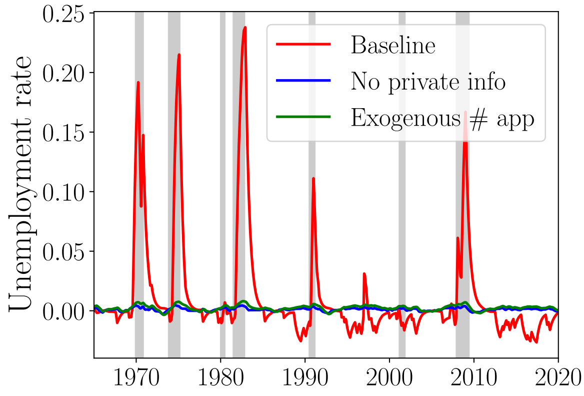
A long literature has documented the relative difficulty to generate the large unemployment fluctuations observed in the data from the small volatility of labour productivity using traditional search models (Shimer, 2005). The combination of endogenous search intensity with asymmetric information helps amplifying the effect of productivity shocks on unemployment fluctuations.

Figure 4 compares fluctuations in the unemployment rate generated by the business cycle movements in labor productivity in the US in the baseline model with endogenous applications and private information and in the same model without these elements. Panel (a) shows the log-deviations from trend of output per hours worked, as measured by the BLS. The grey bands highlight the recession periods defined by the NBER. Panel (b) shows the resulting deviations in the unemployment rate once these productivity shocks are fed into the model.

In the baseline model (red line on Panel (b)), unemployment peaks during recessions and slowly transitions back to its steady-state level afterward[5]. However, removing endogenous search intensity (green line) or asymmetric information between applicants and recruiters (blue line) reduces significantly the unemployment fluctuations generated by the model.



(a) Productivity



(b) Unemployment

Figure 4: Unemployment fluctuations generated by the US labor productivity series
Note: All variables are expressed in log-deviations from trend.

When a negative shock hits, firms offer lower wages and the marginal benefits from each application unemployed workers send decrease. Job seekers become more risk averse and are ready to accept lower wages to limit the shock on their job-finding rate. Because of incentive compatibility constraints, high-productivity workers are hampered in their ability to do so. Firms that target these workers do not want to cut their posted wages too much, otherwise they would not be able to identify their type. Consequently, high-productivity markets offer bundles of wages and queues that are still relatively high and which become less attractive in times of crisis. The marginal benefits from each application decreases more for high-productivity job seekers than for low-productivity ones. This effect leads them to send fewer applications and increases their unemployment rate significantly.

Hence, high-productivity workers are the ones who experience these large fluctuations in unemployment. When the number of applications they send is set exogenously the value of their portfolio of applications falls, but they cannot cut on their search intensity as they do in the baseline model. On the other hand, when there is no private information, they have more margins of adjustment, in particular the wages they target, which limit the fall in the value of their portfolio of application. Therefore, they have lower incentives to reduce their search intensity. It is the interaction of asymmetric information and endogenous search intensity that amplifies the fluctuations of unemployment for high-productivity workers.

[5]This effect is not observed during the 2001 recession because the output per hours worked, which is the source of aggregate fluctuations in the model, fell only slightly in this period, as shown on Panel (a).

4.2 Exogenous parameterization and functional forms

I set the length of a period to one month because it is the standard time frame for data frequency in the CPS and the SCE-JS. It is short enough for workers to send multiple applications simultaneously, i.e. without being able to wait for the result of each application before sending the next one.

I calibrate the persistence ρ_z and volatility σ_z of labor productivity shocks using the series of aggregate output per hour worked provided by the Bureau of Labor Statistics (BLS) from 1948 to 2020[6]. I measure an autocorrelation of 0.9625 and an unconditional standard deviation of 0.0104 for the log-deviations from trend of the HP filtered series with a smoothing parameter of 129,600.

Workers value the consumption of their disposable income through a CRRA utility function $v(w(1-\tau)) = [w(1-\tau)]^{1-\eta}/(1-\eta)$. I set the risk aversion parameter to a conventional value in the literature (Hornstein et al., 2011; Chaumont and Shi, 2022), that is $\eta = 2$. Workers and the representative entrepreneur who owns all the firms discount future payoffs at rate $\beta = 0.96^{1/12}$. The matching function is an urn-ball matching function (Burdett et al., 2001), $m(q) = 1 - e^{-\alpha q}$, with a parameter $\alpha \in (0, 1]$ that represents the share of suitable candidates in each pool of applicants (Petrongolo and Pissarides, 2001). I use a quadratic application cost function $c(a) = k_a \cdot a^2$ with a cost parameter $k_a > 0$.

For the exogenous separation rate δ, I use the average transition rate from employment to unemployment (EU rate) from 1990 to 2020, which I measure using data from the BLS. The interest rate of this small open economy is set to $r = 0.04$.

4.3 Estimated parameters

I jointly estimate the remaining six parameters $\{\kappa, k_a, m^H, \xi, b_0^H, b_0^L, \alpha\}$ using the method of moments. I minimize the squared percentage deviations between moments simulated from the model and their empirical counterparts,

$$\boldsymbol{p}^* = \arg\min_{\boldsymbol{p}\in\mathcal{P}} \sum_{m\in\mathcal{M}} \left(\frac{m^{model} - m^{data}}{m^{data}}\right)^2, \qquad \boldsymbol{p} = \{\kappa, k_a, m^H, \xi, b_0^H, b_0^L, \alpha\}$$

Although these parameters are jointly estimated, each of them is particularly informed by one of the targeted moments. Cross-sectional moments are simulated from an initial state corresponding to the median of the aggregate productivity grid and the steady-state

[6]The BLS measures output per hour worked at a quarterly frequency. I use a linear interpolation to extend it to a monthly frequency.

distribution of workers across employment states and wages. However, all agents are subject to aggregate risk looking forward. Time series moments are computed over a simulated path of 300 periods.

Table 1: Calibrated parameters

External calibration					
Parameter	Description	Value	Source		
β	Discount factor	0.99			
σ_z	Std. dev. of z	0.01	Output/hours (BLS)		
ρ_z	Autocorrelation of z	0.96	Output/hours (BLS)		
δ	Separation rate	0.02	EU rate (BLS)		
r	Interest rate	0.04			
Internal calibration					
Parameter	Description	Value	Target	Model	Data
κ	Vacancy creation cost	0.14	Unemployment rate	0.06	0.06
k_a	Application cost	0.04	Avg. nb. of applications	4.0	6.3
m^H	Share of H-type workers	0.31	St.dev(unemployment)	0.07	0.16
ξ	Productivity of H-type workers	0.44	Mean-min ratio of wages	1.20	1.5
b_0^H	Private replacement inc. (H)	0.18	Avg. nb. of offers	0.39	0.34
b_0^L	Private replacement inc. (L)	0.01	Acceptance rate	0.56	0.49
α	Matching parameter	0.45	Mismatch unemployment rate	0.01	0.01

The cost of creating a vacancy, κ, targets the average unemployment rate. The target is 0.057, the average unemployment rate over the 1948-2020 time frame in the CPS. The application cost parameter, k_a, targets the average number of applications that unemployed workers send, measured in the SCE-JS. The standard deviation of unemployment[7] informs the share of high-productivity workers. As explained in section 4.1.1, the job-finding rate of unemployed workers of type H fluctuates much more with the cycle than for workers of type L. Therefore, a larger share of high-productivity workers amplifies the response of unemployment to productivity shocks.

The difference in productivity between the two types of workers targets the mean-min ratio of residualized wages from Hornstein et al. (2011). This index is a relevant measure of wage dispersion in this context since the model focuses on the wage differences due to unobserved characteristics, i.e. the differences in residualized wages. Workers' exogenous replacement income, b_0^H and b_0^L, are informed by the average number of offers that workers

[7]Where unemployment is measured as the log-deviations from a trend estimated with a HP filter with smoothing parameter 129,600.

receive, and their average acceptance rate. These two parameters determine workers' outside option, and therefore the level of wages and queues that they target in their search, as well as their probability to decline an offer because they have a better one in hand. These two targets are taken from Faberman et al. (2022). Finally, parameter α can be interpreted as the share of job seekers who are mismatched (Petrongolo and Pissarides, 2001), i.e. who are looking for jobs in the wrong labor market. The target is the mismatch unemployment measured by Şahin et al. (2014), that is the difference between the observed unemployment rate and the unemployment rate without any mismatch of workers and vacancies across labor markets. The counterpart of this moment in the model is the difference in unemployment rate between the simulated economy and its equivalent with a value of $\alpha = 1$, which corresponds to an absence of mismatch. Table 1 presents the estimated parameters and the corresponding moments.

5 Quantitative results

I use the calibrated model to solve for the optimal level of unemployment insurance. I start from a benchmark model with no private information and an exogenous number of applications, which is similar to Acemoglu and Shimer (1999, 2000). Then, I add the key ingredients of this model one by one: an endogenous number of applications and asymmetric information between workers and recruiters.

I solve for the optimal unemployment benefits b^L and b^H, but I express them as a percentage of type-specific average wages in order to make them more easily comparable to the current US unemployment system, which is based on a replacement ratio. In a second exercise, I assume that government has less flexibility in its policies and has to pick a common replacement ratio for all workers.

5.1 Type-specific replacement ratio

Table 2 shows the optimal type-specific replacement ratios in the different benchmark models. The bottom right corner shows the results from a model without private information and with an exogenous number of applications[8]. The optimal replacement ratios are quite high, because unemployment insurance does not distort much workers' incentives in this benchmark. Insurance affects the wages posted in equilibrium and their associated queues, but it does not reduce workers' search intensity nor drives high-productivity workers out

[8]This exogenous number of applications is set to the number of applications they send in steady state in the baseline model.

of the market through adverse selection. The optimal level of insurance is higher for low-productivity workers because they have a lower private replacement income, b_0^L, than their high-productivity counterparts.

Table 2: Optimal replacement ratios

	Private information	No private information
Endogenous number of applications	$R^H = 44.3\%$ $R^L = 46.2\%$	$R^H = 44.2\%$ $R^L = 65.3\%$
Exogenous number of applications	$R^H = 50.0\%$ $R^L = 26.2\%$	$R^H = 53.8\%$ $R^L = 80.1\%$

Introducing an endogenous number of applications in this benchmark without private information reduces both levels of insurance. This is what can be seen in the top right corner of Table 2. This reduction in the optimal level of insurance springs from the traditional moral hazard cost: job seekers look for jobs with a lower intensity when they are better insured, which creates a negative fiscal externality forcing the government to raise the labor income tax rate. Both types of workers are subject to moral hazard and generating this fiscal externality, which explains why the planner decides to reduce both replacement ratios.

The bottom left corner of Table 2 introduces private information without endogenous search intensity. Compared to the first benchmark, in the bottom-right corner, this new element generates an adverse selection externality on high-productivity workers, which worsens with the risks low-productivity workers are wiling to take in their applications. Unemployment insurance pushes them towards higher targets, compounding this crowding-out of productive job seekers. This new externality comes from low-productivity workers' incentives to redirect their applications. As such the planner reduces their insurance significantly (from 80.1% in the first benchmark to 26.2% in this economy), while the optimal level of insurance for high-productivity workers remains approximately unchanged. The optimal unemployment insurance in this bottom-left economy with private information and an exogenous number of applications is regressive. It provides a higher replacement ratio to high-productivity workers than to low-productivity ones even though the former earn a higher income than the latter.

Finally, the upper left corner of Table 2 combines private information with an endogenous number of applications. It describes the main economy studied throughout this paper. Compared to the previous benchmark (bottom-left), it rebalances the optimal levels of insurance and provides similar replacement ratios to both types of workers. The reason for this is the new force, described in section 3, created by the interaction of endogenous search intensity with adverse selection. As low-productivity workers benefit from a higher level of insurance, they send fewer applications, which makes them less likely to reach high-productivity markets and alleviates the adverse selection constraints on these markets. This new channel leads the planner to insure low-productivity workers at a higher level, 46.2%.

This exercise highlights an important finding: the optimal level of insurance for low-productivity workers increases to reduce their job search effort, which is typically seen as a drawback of unemployment insurance.

5.2 Common replacement ratio

In practice, most states in the US use a common replacement ratio for all workers. Therefore, restricting the planner's policies to a unique replacement ratio for all workers is a relevant exercise to make the results more comparable with the current US system.

Table 3: Optimal common replacement ratio

	Private information	No private information
Endogenous number of applications	$R = 47.4\%$	$R = 64.6\%$
Exogenous number of applications	$R = 38.1\%$	$R = 79.7\%$

Table 3 shows the result of the same exercise as the one conducted in the previous subsection, but with a common replacement ratio for all workers. Similar patterns appear. A model with no private information and an exogenous number of applications (bottom-right corner) yields a high level of insurance. Adverse selection externalities are shut down, and the traditional moral hazard cost of unemployment insurance is relatively small given that workers can only adapt the directions of their applications and not their number. In this scenario, consumption-smoothing benefits are large while the fiscal cost from moral hazard is relatively low. Thus, the planner picks a high replacement ratio, namely 79.7%.

The introduction of a new margin of endogenous search (top right corner) generates an additional moral hazard cost, which makes optimal unemployment insurance less generous. The introduction of private information (bottom left corner) creates the adverse selection channel through which better insured low-productivity workers crowd out high-productive applicants. This new externality makes optimal unemployment insurance less generous. Although these two ingredients separately create incentives for the planner to make the unemployment insurance less generous, their interaction has a more nuanced result.

The model that includes both asymmetric information and endogenous search intensity (top left corner) yields an optimal replacement ratio of 47.4%. Both elements generate additional insurance costs, but their interaction mitigates the crowding-out of high-productivity workers by discouraging low-productivity job seekers to over-apply. The fact that workers search less intensely for a job when they are better insured may therefore be efficiency-enhancing.

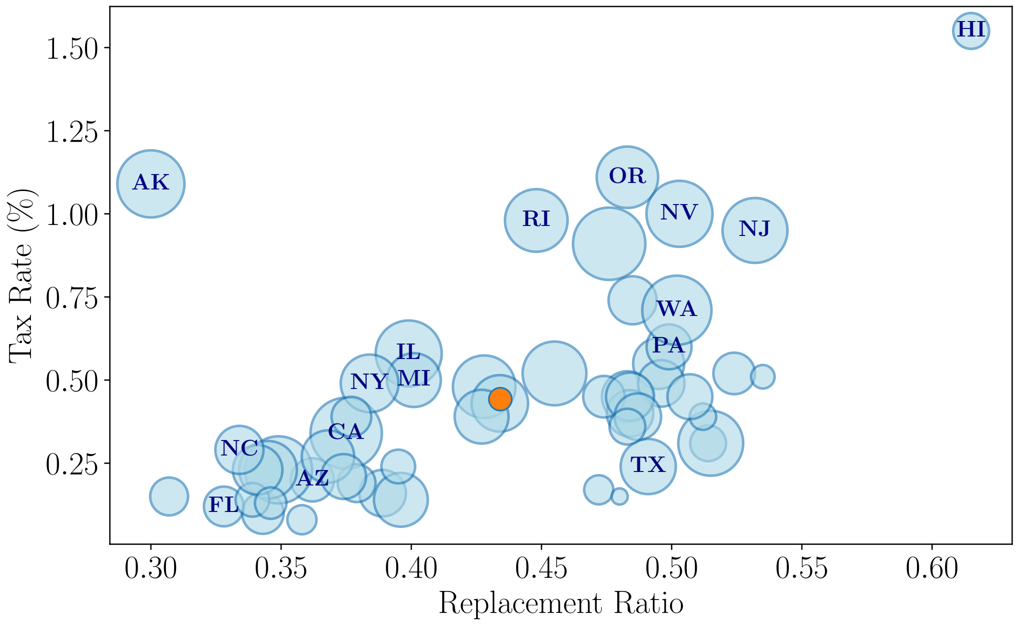


Figure 5: Replacement ratio and UI tax rate by state

Note: The replacement ratio for each state is measured as the average ratio of weekly benefits over wages. The tax rate is computed as the amount of taxes collected by the state's insurance system over the sum of its wages (not restricted to the taxable portion). The size of each circle represents the states' unemployment rates, and the orange circle corresponds to the average of both axes. All variables come from the Department of Labor's report on the UI system and are measured in the first quarter of 2024.

The combination of the traditional forces that affect the design of unemployment insurance, consumption-smoothing and moral hazard, with new externalities resulting from adverse selection and an endogenous number of applications, produce an optimal replacement ratio of 47.4%, which is slightly higher than the current average ratio applied in the US.

Figure A.9 shows the replacement income provided by states' unemployment insurance, as a percentage of the average wage, and the tax rate on labor income used to finance it. The relationship between these two variables is positive, as expected from the UI budget constraint. The orange circle represents the average across states along both dimensions. The average replacement ratio is 43.4%, which is slightly below the optimal level found through the quantitative exercise. This figure also shows that a large group of states provide a replacement income that is very close to the optimal one. A second important group provides replacement ratios between 0.3 and 0.4 which, according to the present findings, may be suboptimal.

6 Conclusion

This paper develops a framework in which workers' search behavior affects recruiters' cost of acquiring information about applicants. The resulting adverse selection alters the composition of the pool of job seekers by crowding out high-productivity workers. I show that this channel matters for the design of unemployment insurance, as it distorts workers' incentives to search and, therefore, the strength of adverse selection in the labor market.

I analyze these forces theoretically and quantify their impact in an extended model incorporating aggregate risk. The results reveal that the costs and benefits of these new externalities are significant, leading to an optimal unemployment benefit level of 47.4% of the average wage. This figure is slightly higher than the current average replacement ratio in the US, as it accounts for the fact that providing higher insurance discourages low-productivity workers from over-applying and crowding out high-productivity workers.

These findings are particularly relevant given new evidence that job application costs are decreasing (Martellini and Menzio, 2020), and job seekers are submitting more applications than in previous decades (Birinci et al., 2024). This shift in search technologies may alter governments' incentives to provide unemployment insurance and increase the importance of alternative labor market policies, such as those aimed at redirecting workers' search efforts (Behaghel et al., 2024).

References

Acemoglu, Daron and Robert Shimer, "Efficient unemployment insurance," *Journal of political Economy*, 1999, *107* (5), 893–928.

_ **and** _ , "Productivity gains from unemployment insurance," *European economic review*, 2000, *44* (7), 1195–1224.

Albrecht, James, Pieter A Gautier, and Susan Vroman, "Equilibrium directed search with multiple applications," *The Review of Economic Studies*, 2006, *73* (4), 869–891.

_ **, Xiaoming Cai, Pieter Gautier, and Susan Vroman**, "Multiple applications, competing mechanisms, and market power," *Journal of Economic Theory*, 2020, *190*, 105121.

Auster, Sarah, Piero Gottardi, and Ronald P Wolthoff, "Simultaneous search and adverse selection," 2022.

Baily, Martin N., "Some aspects of optimal unemployment insurance," *Journal of Public Economics*, 1978, *10* (3), 379–402.

Barbanchon, Thomas Le, Johannes F Schmieder, and Andrea Weber, "Job Search, Unemployment Insurance, and Active Labor Market Policies," *National Bureau of Economic Research*, 2024.

Behaghel, Luc, Sofia Dromundo, Marc Gurgand, Yagan Hazard, and Thomas Zuber, "The Potential of Recommender Systems for Directing Job Search: A Large-Scale Experiment," *IZA Discussion Paper*, 2024.

Birinci, Serdar, Kurt See, and Shu Lin Wee, "Job applications and labor market flows," *Review of Economic Studies*, 2024, p. rdae064.

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

Burdett, Kenneth and Kenneth L Judd, "Equilibrium price dispersion," *Econometrica: Journal of the Econometric Society*, 1983, pp. 955–969.

_ **, Shouyong Shi, and Randall Wright**, "Pricing and matching with frictions," *Journal of Political Economy*, 2001, *109* (5), 1060–1085.

Cai, Zhifeng and Jonathan Heathcote, "The great resignation and optimal unemployment insurance," *Federal Reserve Bank of Minneapolis, Research Division*, 2023.

Chade, Hector and Lones Smith, "Simultaneous search," *Econometrica*, 2006, *74* (5), 1293–1307.

Chang, Briana, "Adverse selection and liquidity distortion," *The Review of Economic Studies*, 2018, *85* (1), 275–306.

Chaumont, Gaston and Shouyong Shi, "Wealth accumulation, on-the-job search and inequality," *Journal of Monetary Economics*, 2022, *128*, 51–71.

Chetty, Raj, "A general formula for the optimal level of social insurance," *Journal of Public Economics*, 2006, *90* (10-11), 1879–1901.

Davoodalhosseini, Seyed Mohammadreza, "Constrained efficiency with adverse selection and directed search," *Journal of Economic Theory*, 2019, *183*, 568–593.

Engbom, Niklas, "Contagious unemployment," Technical Report, National Bureau of Economic Research 2021.

Faberman, R Jason, Andreas I Mueller, Ayşegül Şahin, and Giorgio Topa, "Job search behavior among the employed and non-employed," *Econometrica*, 2022, *90* (4), 1743–1779.

Galenianos, Manolis and Philipp Kircher, "Directed search with multiple job applications," *Journal of economic theory*, 2009, *144* (2), 445–471.

Gautier, Pieter A and Ronald P Wolthoff, "Simultaneous search with heterogeneous firms and ex post competition," *Labour Economics*, 2009, *16* (3), 311–319.

_ , José L Moraga-González, and Ronald P Wolthoff, "Search costs and efficiency: Do unemployed workers search enough?," *European Economic Review*, 2016, *84*, 123–139.

Guerrieri, Veronica, "Inefficient unemployment dynamics under asymmetric information," *Journal of Political Economy*, 2008, *116* (4), 667–708.

_ and Robert Shimer, "Dynamic adverse selection: A theory of illiquidity, fire sales, and flight to quality," *American Economic Review*, 2014, *104* (7), 1875–1908.

_ , _ , and Randall Wright, "Adverse selection in competitive search equilibrium," *Econometrica*, 2010, *78* (6), 1823–1862.

Hagedorn, Marcus and Iourii Manovskii, "The cyclical behavior of equilibrium unemployment and vacancies revisited," *American Economic Review*, 2008, *98* (4), 1692–1706.

Hornstein, Andreas, Per Krusell, and Giovanni L Violante, "Frictional wage dispersion in search models: A quantitative assessment," *American Economic Review*, 2011, *101* (7), 2873–2898.

Kircher, Philipp, "Efficiency of simultaneous search," *Journal of Political Economy*, 2009, *117* (5), 861–913.

Landais, Camille, Pascal Michaillat, and Emmanuel Saez, "A macroeconomic approach to optimal unemployment insurance: Theory," *American Economic Journal: Economic Policy*, 2018, *10* (2), 152–181.

Martellini, Paolo and Guido Menzio, "Declining search frictions, unemployment, and growth," *Journal of Political Economy*, 2020, *128* (12), 4387–4437.

Menzio, Guido and Shouyong Shi, "Block recursive equilibria for stochastic models of search on the job," *Journal of Economic Theory*, 2010, *145* (4), 1453–1494.

_ **and** _ , "Efficient search on the job and the business cycle," *Journal of Political Economy*, 2011, *119* (3), 468–510.

Petrongolo, Barbara and Christopher A Pissarides, "Looking into the black box: A survey of the matching function," *Journal of Economic literature*, 2001, *39* (2), 390–431.

Rudanko, Leena, "Labor market dynamics under long-term wage contracting," *Journal of monetary Economics*, 2009, *56* (2), 170–183.

Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L Violante, "Mismatch unemployment," *American Economic Review*, 2014, *104* (11), 3529–3564.

Schaal, Edouard, "Uncertainty and unemployment," *Econometrica*, 2017, *85* (6), 1675–1721.

Schmieder, Johannes F and Till Von Wachter, "The effects of unemployment insurance benefits: New evidence and interpretation," *Annual Review of Economics*, 2016, *8* (1), 547–581.

Shimer, Robert, "The cyclical behavior of equilibrium unemployment and vacancies," *American economic review*, 2005, *95* (1), 25–49.

Wolthoff, Ronald, "Applications and interviews: Firms' recruiting decisions in a frictional labour market," *The Review of Economic Studies*, 2018, *85* (2), 1314–1351.

Wright, Randall, Philipp Kircher, Benoît Julien, and Veronica Guerrieri, "Directed search and competitive search equilibrium: A guided tour," *Journal of Economic Literature*, 2021, *59* (1), 90–148.

APPENDIX

A Additional figures

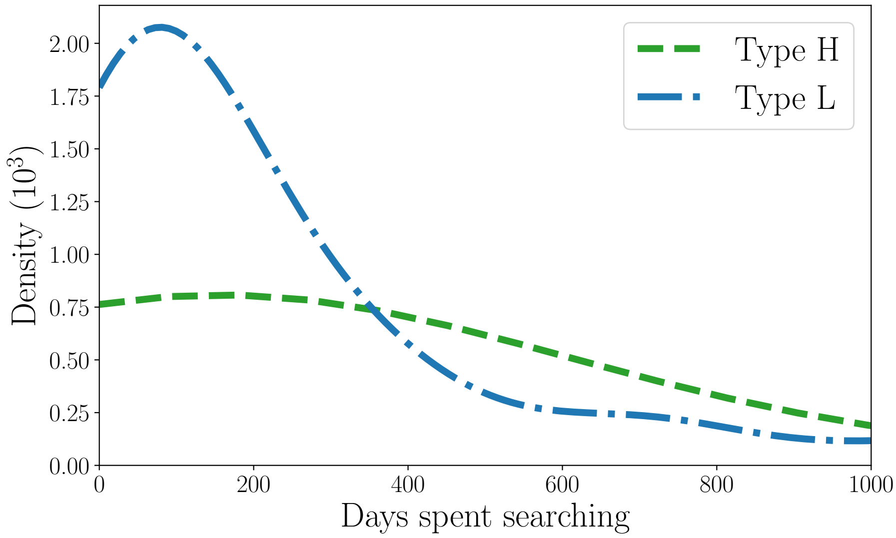


Figure A.6: Density of unemployment duration by type

Note: The two groups of unemployed workers are created by splitting the sample of unemployed workers in the SCE-JS based on their residual wage in their last job. This wage is residualized on education, age, age squared and gender. H-types and L-types are respectively the unemployed workers with a residualized wage above and below the median. The outcome variable is the density of answers to the question: "How many days have you spent looking for a job?".

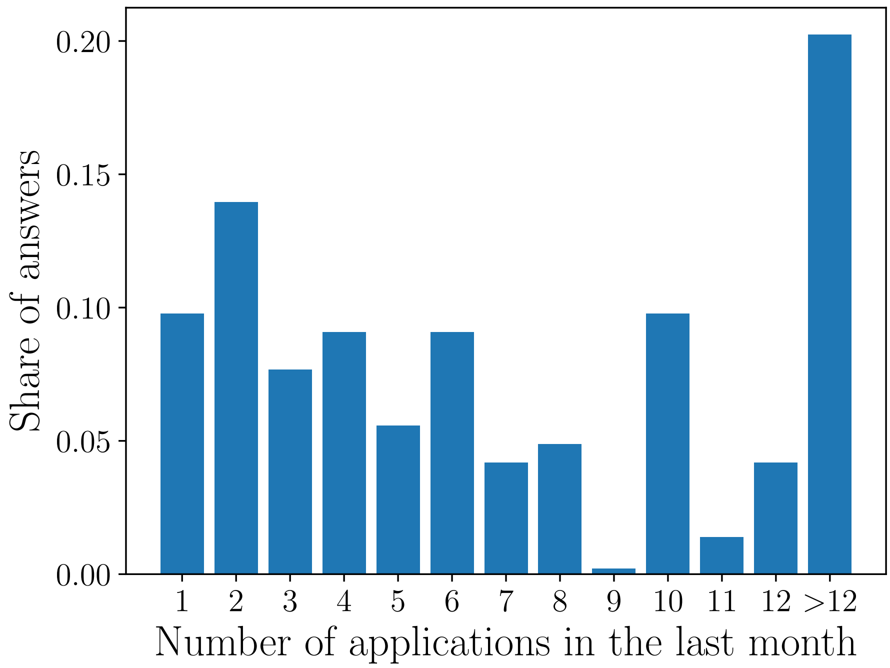


Figure A.7: Number of applications sent by unemployed workers over the past month

Note: This question is asked in the SCE-JS data.

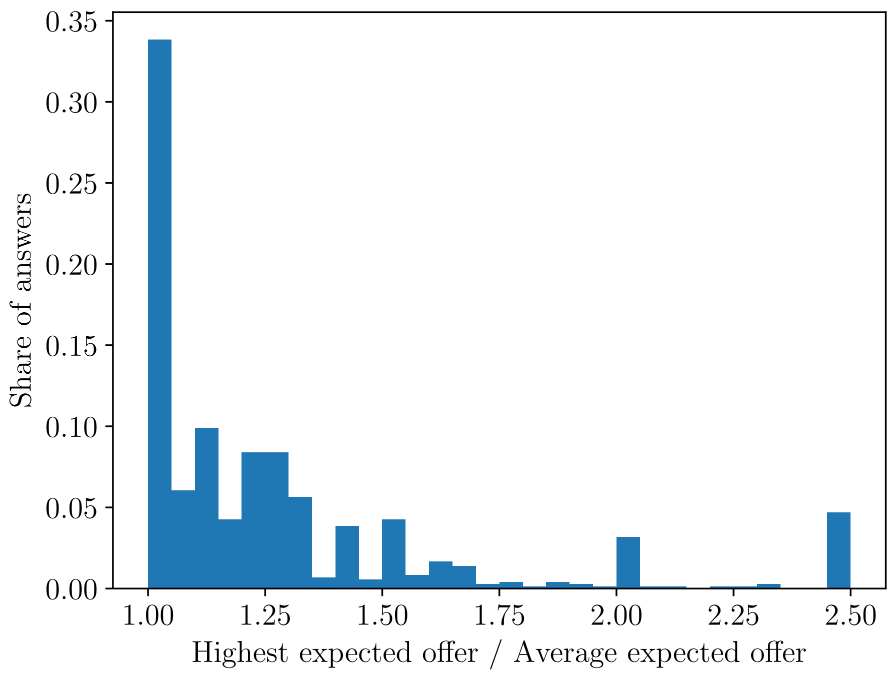


Figure A.8: Dispersion of expected offers

Note: This measure is constructed from two questions asked in the SCE-JS data, which correspond to the numerator and denominator of this index.

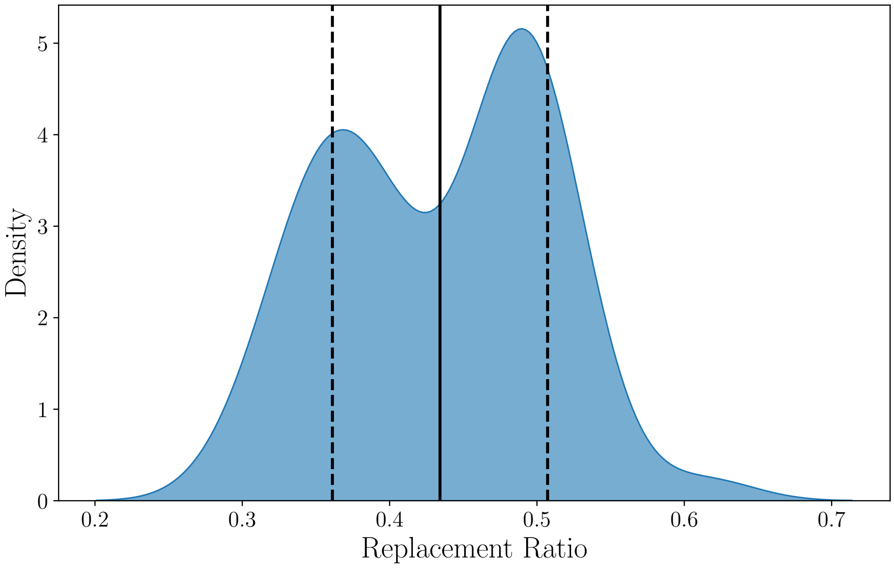


Figure A.9: Density of replacement ratios across states

Note: The replacement ratio for each state is measured as the average ratio of weekly benefits over wages. This measure is provided by the Department of Labor's report on the UI system and is measured in the first quarter of 2024. The solid black line represents the mean, and the two dashed black lines show the values that are one standard deviation away from the mean.

B Proofs

B.1 Proof of Proposition 1

The incentive compatibility constraint IC1 defines the lowest incentive compatible wage $\underline{w}^H$ implicitly as the wage which satisfies

$$\psi(q^H(\underline{w}^H, z_t))W^L(\underline{w}^H, z_t) + \Big(1 - \psi(q^H(\underline{w}^H, z_t))\Big) S^L_{a^L-1}(z_t) = S^L_{a^L}(z_t)$$

Differentiating this equation with respect to $\underline{w}^H$, b^L and τ yields (I drop the arguments z_t for clarity)

$$\frac{d\underline{w}^H}{db^L} = \frac{1}{\frac{\partial\left[\psi(q^H(\underline{w}^H, z_t))W^L(\underline{w}^H, z_t) + \left(1 - \psi(q^H(\underline{w}^H, z_t))\right) S^L_{a^L-1}(z_t)\right]}{\partial \underline{w}^H}} \Bigg[\left(\frac{\partial S^L_{a^L}}{\partial b^L} + \frac{\partial S^L_{a^L}}{\partial \tau}\frac{d\tau}{db^L} \right)$$
$$- \left(\frac{\partial \left[\psi(q^H(\underline{w}^H))W^L(\underline{w}^H) + \left(1 - \psi(q^H(\underline{w}^H))\right) S^L_{a^L-1}\right]}{\partial b^L} \right.$$
$$\left. + \frac{\partial \left[\psi(q^H(\underline{w}^H))W^L(\underline{w}^H) + \left(1 - \psi(q^H(\underline{w}^H))\right) S^L_{a^L-1}\right]}{\partial \tau}\frac{d\tau}{db^L} \right) \Bigg]$$

The denominator,

$$\frac{\partial \left[\psi(q^H(\underline{w}^H, z_t))W^L(\underline{w}^H, z_t) + \Big(1 - \psi(q^H(\underline{w}^H, z_t))\Big) S^L_{a^L-1}(z_t)\right]}{\partial \underline{w}^H}$$

is negative, by definition of the lowest incentive compatible contract. A contract that offers a slightly higher wage than $\underline{w}^H$ provides a lower level of utility to workers of type L than their outside option $S^L_{a^L}(z_t)$, and thus a lower level of utility than contract $\underline{w}^H$. This is what makes the contracts with higher wages than $\underline{w}^H$ incentive compatible.

The numerator is the difference between the welfare gains from a higher level of insurance for the application portfolios $(w^L_1, ..., w^L_{a^L-1}, w^L_{a^L})$ and $(w^L_1, ..., w^L_{a^L-1}, \underline{w}^H)$. The first portfolio is the one chosen by workers of type L in equilibrium, when they apply to a^L jobs simultaneously. The second portfolio is the same one, except that the last application is sent to the lowest incentive compatible market. By definition, these two portfolios provide the same utility to workers of type L. However, the second portfolio bears more unemployment risk than the first one, which it offsets by providing a higher excepted wage. Hence, the welfare gains from a marginal increase in unemployment insurance are higher for the riskier portfolio, $(w^L_1, ..., w^L_{a^L-1}, \underline{w}^H)$. Therefore, the numerator is negative, and

$$\frac{d\underline{w}^H}{db^L} > 0$$

This result means that, as the unemployment benefit for workers of type L, b^L, increases, the incentive compatibility constraint IC1 tightens, i.e. the set of incentive compatible wages shrinkens.

It follows that the set on incentive compatible queues shrinkens as well, given that the lowest incentive compatible queue length is defined by

$$\underline{q}^H = q^H(\underline{w}^H, z_t) \tag{9}$$

where $q^H(w, z_t)$ is increasing in w. Indeed, this function represents the isoprofit curve along which firms post contracts that satisfy the free-entry condition. As such, any increase in the posted wage needs to be compensated for by an longer queue for firms expected profits to remain equal to zero. Therefore,

$$\frac{d\underline{q}^H}{db^L} = \frac{\partial \underline{q}^H}{\partial b^L}\frac{d\underline{w}^H}{db^L} > 0$$

B.2 Model with aggregate risk

B.2.1 Productivity

I introduce aggregate shocks by assuming that the common component of labor productivity, z_t, is indexed by t and follows an AR(1) process

$$z_t = \rho_z z_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_z)$$

where ρ_z is the persistence parameter of this process and σ_z is its volatility.

In period t, a job filled with a low-productivity worker produces z_t, while a job filled with a high-productivity worker produces $z_t(1+\xi)$, $\xi > 0$.

B.2.2 Contracts

A contract ϕ_t, which promises utility x in period t, is state-contingent and specifies the wages $\{w_{t+j}\}_{j=0}^{\infty}$ that the worker would obtain in any period $t+j$ and any aggregate state of the world ω_{t+j} Ω if the match survives idiosyncratic separation shocks. These contracts can be expressed recursively as $\phi = \{w, x'(\omega')\}$, that is a contract that specifies the current wage and the state-contingent promised utility for next period. I assume that these recursive contracts are written at the production phase of each period.

B.2.3 Firms

In terms of notations, I use tilde capital letters for the value functions that are expressed in terms of promised utility rather than wages. The expected payoff for a firm posting a vacancy, in period t, that promises utility x is

$$\widetilde{V}_t(x) = \mu(q_t(x)) \left[\gamma_t(x)\widetilde{J}_t^L(x) + (1-\gamma_t(x))\widetilde{J}_t^H(x)\right] - \kappa$$

where the value of a job filled with a worker of type $i \in \{H, L\}$, conditional on the aggregate state indexed by t and the utility x promised to the worker, is

$$\begin{aligned}\widetilde{J}_t^i(x) = \max_{\phi=\{w,\boldsymbol{x'(\omega')}\}} \quad & z_t\Big(1 + \xi\mathbb{1}[i = H]\Big) - w_t + \beta(1-\delta)\mathbb{E}_t\left[\widetilde{J}_{t+1}^i(x'(\omega_{t+1}))\right] \\ \text{s.t.} \quad & \widetilde{W}_t^i(\phi) \geq x\end{aligned}$$

Each active firm rewrites its contract at the production stage of every period. It chooses its worker's current wage, w_t, and a state-contingent promised utility for the next period,

$\boldsymbol{x'(\omega')}$, conditional on utility x promised in the previous period, to maximize its expected profits. $\boldsymbol{x'(z')}$ is a vector of promised utility for each possible value of ω_{t+1}.

B.2.4 Workers

The lifetime utility of a worker of type $i \in \{H, L\}$ employed with a contract ϕ is

$$\widetilde{W}_t^i(\phi) = v\Big((1-\tau)w_t\Big) + \beta\mathbb{E}_t\Big[\delta U_{t+1}^i + (1-\delta)x'(\omega_{t+1})\Big]$$

The lifetime utility of an unemployed worker of type $i \in \{H, L\}$ is

$$U_t^i = v(b^i + b_0^i) + \beta\mathbb{E}_t\left[\max_a\Big\{S_t^i(a) - c(a)\Big\}\right]$$

where the value of a portfolio of a applications, $S_t^i(a)$, is defined in a similar way than in the steady state case, that is

$$S_t^i(a) = \max_{\boldsymbol{x}\in\mathcal{X}_t^a} \quad \Bigg[\sum_{k=1}^{a} \underbrace{\Bigg(\prod_{j=k+1}^{a}\Big[1-\psi(q_t(x_j))\Big]\Bigg)\psi(q_t(x_k))}_{\Pr\left[i\text{-th lowest app = first success}\right]} x_k \ + \underbrace{\prod_{j=1}^{a}\Big[1-\psi(q_t(x_j))\Big]}_{\Pr\left[\text{no job offer}\right]} U_t^i\Bigg]$$

The recursive formulation of workers' application problem is

$$S_t^i(k) = \max_{x\in\mathcal{X}_t} \quad \psi(q_t(x))x + \Big[1-\psi(q_t(x))\Big]S_t^i(k-1)$$

for every $k \in \{1, ..., a_t^i\}$, and where $S_t^i(0) = U_t^i$.

B.3 Optimal contracts under two-sided commitment

I assume that workers and firms can commit not separate until the match is exogenously destroyed. A job with contract ϕ, that promises utility x to its worker in period t, satisfies the constraint

$$\widetilde{W}_t^i(\phi) = x$$

Firms choose the lowest wage satisfying the promise-keeping constraint in order to maximize their current profits, which explains why this constraint is binding. Then, one can express the primsed utility x as

$$x = v\Big((1-\tau)w_t\Big) + \beta\mathbb{E}_t\Big[\delta U^i_{t+1} + (1-\delta)x'(\omega_{t+1})\Big]$$

which gives an expression for the wage

$$w_t = \frac{1}{1-\tau}v^{-1}\left(x - \beta\mathbb{E}_t\Big[\delta U^i_{t+1} + (1-\delta)x'(\omega_{t+1})\Big]\right)$$

I use this equation to substitute for w_t in the Bellman equation that defines the present discounted value of profits generated by a job that employs a worker of type i, such that it becomes a maximization over a single variable

$$\begin{aligned}\widetilde{J}^i_t(x) = \max_{\boldsymbol{x'(\omega')}} \quad & z_t\Big(1+\xi\mathbb{1}[i=H]\Big) - \frac{1}{1-\tau}v^{-1}\left(x - \beta\mathbb{E}_t\Big[\delta U^i_{t+1} + (1-\delta)x'(\omega_{t+1})\Big]\right) \\ & + \beta(1-\delta)\mathbb{E}_t\left[\widetilde{J}^i_{t+1}\left(x'(\omega_{t+1})\right)\right]\end{aligned}$$

There are as many first order conditions to this maximization problem as possible states of the world. The FOC with respect to $x'(\omega')$, i.e. the utility promised for one particular state ω', is

$$\begin{aligned}\frac{1}{1-\tau}(v^{-1})'\left(x - \beta\mathbb{E}_t\Big[\delta U^i_{t+1} + (1-\delta)x'(z_{t+1})\Big]\right)\beta(1-\delta)\mathbb{P}\mathrm{r}[\omega'|\omega] \\ + \beta(1-\delta)\mathbb{P}\mathrm{r}[\omega'|\omega]\frac{\partial\widetilde{J}^i_{t+1}\left(x'(\omega')\right)}{\partial x'(\omega')} = 0\end{aligned}$$

i.e.

$$\frac{1}{1-\tau}(v^{-1})'\left(x - \beta\mathbb{E}_t\Big[\delta U^i_{t+1} + (1-\delta)x'(z_{t+1})\Big]\right) + \frac{\partial\widetilde{J}^i_{t+1}\left(x'(\omega')\right)}{\partial x'(\omega')} = 0$$

Using the envelope theorem, we get

$$\frac{\partial\widetilde{J}^i_t(x)}{\partial x} = -\frac{1}{1-\tau}(v^{-1})'\left(x - \beta\mathbb{E}_t\Big[\delta U^i_{t+1} + (1-\delta)x'(\omega_{t+1})\Big]\right)$$

Combining this condition at $t+1$ and for $x = x'(\omega')$ with the FOC yields

$$v'\left((1-\tau)w_t\right) = v'\left((1-\tau)w_{t+1}(\omega_{t+1})\right)$$

for any $\omega_{t+1} \in \Omega$, and where $w_{t+1}(\omega_{t+1})$ is the wage in period $t+1$ if the state is ω_{t+1}.

Hence, the promised wage is a constant wage if $v(.)$ is strictly concave. Under full-commitment on both sides (employer and employee), firms optimally post contracts that offer a constant wage. Intuitively, firms are risk-neutral while workers are risk-averse, which gives the former an incentive to post a contract that offers a relatively low average wage but a perfect insurance against aggregate shocks that occur during the period of activity of the match.

B.4 Block Recursive Equilibrium

Hence, we can rewrite the value functions in terms of the wage w associated with each contract instead of promised utility x. I will use capital letters without tildes to denote these value functions, that is $\widetilde{V}_t(w) := \widetilde{V}_t(x)$, $J^i_t(w) := \widetilde{J}^i_t(x)$, $W^i_t(w) := \widetilde{W}^i_t(x)$, and

$$S^i_t(k) = \max_{w \in \mathcal{W}_t} \quad \psi(q_t(w))W^i_t(w) + \Big[1 - \psi(q_t(w))\Big] S^i_t(k-1)$$

for every $k \in \{1, ..., a^i_t\}$, and where $S^i_t(0) = U^i_t$.

The corresponding FOC, for $k \in \{1, ..., a^i_t\}$, is

$$\psi'(q_t(w))q'_t(w)\Big[W^i_t(w) - S^i_t(k-1)] + \psi(q_t(w))\frac{\partial W^i_t(w)}{\partial w} = 0 \tag{10}$$

If the equilibrium is separating, i.e. if $\gamma(w) \in \{0,1\}$, an entrant firm that targets low-productivity workers chooses a wage to post w by solving

$$\max_{w} \quad \mu(q_t(w))J^L_t(w) - \kappa$$

The corresponding FOC is

$$\mu'(q_t(w))q'_t(w)J(w,z) + \mu(q_t(w))\frac{\partial J^L_t(w)}{\partial w} = 0 \tag{11}$$

Following the market utility approach (Wright et al., 2021), I assume that firms expect queues to satisfy (10) for any wage they could post on markets that attract low-productivity workers, such that we can combine the two conditions to obtain

$$\frac{\mu'(q_t(w))\psi(q_t(w))\frac{\partial W^L_t(w)}{\partial w}J^L_t(w)}{\psi'(q_t(w))\Big[W^L_t(w) - S^L_t(k-1)]} = \mu(q_t(w))\frac{\partial J^L_t(w)}{\partial w} \tag{12}$$

Note that, given a value of unemployment U^L_t and a stochastic process for z_t, one can solve for $J^L_t(w)$ and $W^L_t(w)$ for any wage w. Indeed, $J^L_t(w)$ is a function of the aggregate state only through the productivity z_t, and $W^L_t(w)$ is a function of the aggregate state only through the value of unemployment.

Then, the free entry condition pins down the equilibrium queue $q_t(w)$ for any w,

$$\mu(q_t(w))J_t^L(w) = \kappa \tag{13}$$

Hence, conditional on a value of unemployment U_t^L and a stochastic process for z_t, one can solve for posted wages and associated queues by combining (13) and (12) starting from $k = 1$ up to any number of applications A. This process also determines the value functions $W_t^L(w)$, $J_t^L(w)$ and $S_t^L(k)$, $k \in \{1, ..., A\}$.

The optimal number of applications sent by low-productivity workers a_t^L is determined by

$$a_t^L = \arg\max_a \ \{S_t^L(a) - c(a)\}$$

This shows that one can solve for equilibrium wages and value functions for low-productivity markets without any other reference to the aggregate state than the aggregate productivity z_t and the value of unemployment U_t^L.

In order to solve for the same outcomes for high-productivity markets, one needs first solve for the unconstrained outcomes using exactly the same equations than for low-productivity workers, but with a productivity which is now $z_t(1+\xi)$ and a value of unemployment U_t^H. This yields a set of wages $(\widehat{w}_{1,t}^H, ..., \widehat{w}_{A,t}^H)$, queues $(\widehat{q}_{1,t}^H, ..., \widehat{q}_{A,t}^H)$, and portfolio values $(\widehat{S}_{1,t}^H, ..., \widehat{S}_{A,t}^H)$ that would be the equilibrium outcome for high-productivity workers if firms were perfectly informed about their applicants.

To construct the constrained equilibrium with asymmetric information, I use the equilibrium values of $S_t^L(a^L)$ and $S_t^L(a^L - 1)$ to check whether the lowest wage posted in H-type markets under perfect information, $\widehat{w}_{1,t}^H$, and its associated queue $\widehat{q}_{1,t}^H$, satisfy the incentive compatibility constraints (IC1) and (IC2). If they do, these constraints are not binding and the equilibrium is the same as the one under perfect information.

However, if $(\widehat{w}_{1,t}^H, \widehat{q}_{1,t}^H)$ violates incentive compatibility constraints, the first equilibrium wage posted on H-type markets is the lowest one that satisfies both incentive compatibility constraints and free entry on this market. The wage-queue combination needs to be on the isoprofit curve defined by the free entry condition

$$\mu(q_t(w))J_t^H(w) = \kappa \tag{14}$$

The first incentive compatibility constraint, (IC1), can be represented as the wage-queue combinations that satisfy the following equation

$$\psi(q_t(w))W_t^L(w) + [1 - \psi(q_t(w))]S_t^L(a^L - 1) = S_t^L(a^L) \tag{15}$$

The second incentive compatibility constraint is represented similarly as the wage-queue pairs that satisfy

$$\psi(q_t(w))W_t^L(w) + [1 - \psi(q_t(w))]S_t^L(a^L) - [c(a^L + 1) - c(a^L)] = S_t^L(a^L) \tag{16}$$

Each of these two curves intersect with the isoprofit curve defined by (14) at on the interval of wages $(\widehat{w}_{1,t}^H, \infty)$ exactly once. As firms choose the lowest incentive compatible wage, they choose the highest of the two wages corresponding to these intersections as the equilibrium wage posted to attract the first application of high-productivity workers, $w_{1,t}^H$, and its associated queue (defined by the intersection point) $q_{1,t}^H$.

This new market for H-type workers' first application determines a new value $S_t^H(1)$ for a portfolio of one application. One needs to recompute the unconstrained wages for high-productivity markets for $k \in \{2, ..., A\}$, i.e. $(\widehat{w}_{2,t}^H, ..., \widehat{w}_{A,t}^H)$, as well as their associated queues $(\widehat{q}_{2,t}^H, ..., \widehat{q}_{A,t}^H)$. Then, it is possible to check whether $(\widehat{w}_{2,t}^H, \widehat{q}_{2,t}^H)$ is incentive compatible. If it is, H-type markets are characterized by wages $(w_{1,t}^H, \widehat{w}_{2,t}^H, ..., \widehat{w}_{A,t}^H)$ and queues $(q_{1,t}^H, \widehat{q}_{2,t}^H, ..., \widehat{q}_{A,t}^H)$. If it is not incentive compatible, the same process starts again to find the equilibrium wage and queues for the second H-type application, using the intersection of the free entry condition with incentive compatibility constraints (IC1) and (IC2).

This is how the separating equilibrium is constructed. Importantly, these equilibrum outcomes are found withtout any other reference to the aggregate state than through the aggregate productivity z_t and values of unemployment U_t^L and U_t^H. In particular, the distribution of workers across employment states and wages is not needed. This crucial element proves the block recursivity of this equilibrium. I solve for U_t^L and U_t^H through value function iteration in an outer loop.

B.5 Tightening of IC1 as the number of applications increases

The first incentive compatibility constraint is

$$\psi(q)W^L(w, z_t) + [1 - \psi(q)]S_{a^L-1}^L(z_t) \leq S_{a^L}^L(z_t)$$

Need to show that the set of $(w, q) \in \mathbb{R}^2$ satisfying this constraint for $a^L + 1$ applications is a subset of the $(w, q) \in \mathbb{R}^2$ that satisfy the constraint for a^L applications.

Regardless of the number of applications, contracts posted in markets targeting high-type workers need to satisfy the zero-profit condition, such that q can be expressed as a function of w, i.e. $q^H(w, z_t)$. Hence, the constraint for a^L applications, can be rewrittent as

$$\psi(q^H(w, z_t))W^L(w, z_t) + [1 - \psi(q^H(w, z_t))]S_{a^L-1}^L(z_t) \leq S_{a^L}^L(z_t)$$

Function $q^H(w, z_t)$ is continuous and strictly increasing in w. Intuitively, a firm posting a contract with a high wage needs to be compensated with a long queue, i.e. a high job-filling rate, to reach the same level of expected profits.

Then, it is enough to prove that the lowest wage satisfying IC1 when workers of type L send a^L applications, denoted $\underline{w}(a^L)$, and implicitly defined by

$$\psi(q^H(\underline{w}(a^L), z_t))W^L(\underline{w}(a^L), z_t) + [1 - \psi(q^H(\underline{w}(a^L), z_t))]S^L_{a^L-1}(z_t) = S^L_{a^L}(z_t) \tag{17}$$

is lower than the lowest wage satisfying IC1 when workers of type L send a^L+1 applications, denoted $\underline{w}(a^L + 1)$, and implicitly defined by

$$\psi\Big(q^H(\underline{w}(a^L+1), z_t)\Big)W^L\Big(\underline{w}(a^L+1), z_t\Big) + \Big[1 - \psi\Big(q^H(\underline{w}(a^L+1), z_t)\Big)\Big]\, S^L_{a^L}(z_t) = S^L_{a^L+1}(z_t) \tag{18}$$

An indifference curve in the (w, q) space for the k-th lowest application, $k \geq 1$, sent by a worker of type L is

$$\psi(q)W^L(w, z_t) + [1 - \psi(q)]S^L_{k-1}(z_t) = S^L_k(z_t)$$

The slope of this curve can be determined by differentiating this equation with respect to q and w, i.e.

$$\psi'(q)W^L(w, z_t)dq + \psi(q)\frac{\partial W^L(w, z_t)}{\partial w}dw - \psi'(q)S^L_{k-1}(z_t)dq = 0$$

or

$$\frac{dq}{dw} = \frac{\psi(q)\frac{\partial W^L(w,z_t)}{\partial w}}{\psi'(q)\Big(S^L_{k-1}(z_t) - W^L(w, z_t)\Big)}$$

When $W^L(w, z_t) > S^L_{k-1}(z_t)$, which is always the case when w is a potential target for an additional application and $S^L_{k-1}(z_t)$ is its fallback option (can prove that from the fact that the cost of an extra application is strictly positive), $dq/dw > 0$ along the indifference curve, and this slope is increasing in $S^L_{k-1}(z_t)$.

Consequently, two consecutive indifference curves, C(k) and C(k+1), corresponding to applications k and $k + 1$, defined by the pairs of $(w, q) \in \mathbb{R}^2$ such that

$$\psi(q)W^L(w, z_t) + [1 - \psi(q)]S^L_{k-1}(z_t) = S^L_k(z_t) \tag{C(k)}$$

and

$$\psi(q)W^L(w,z_t) + [1-\psi(q)]S_k^L(z_t) = S_{k+1}^L(z_t) \tag{C(k+1)}$$

intersect only once in the (w,q) space, because $S_{k-1}^L(z_t) \neq S_k^L(z_t)$. Let $(w_{k,k+1}, q_{k,k+1})$ denote the wage and queue length that correspond to this intersection. Let q_0^k and q_0^{k+1} denote the images of a given wage w_0 on indifference curves C(k) and C(k+1), that is the queue lengths which satisfy the last two equations for wage w_0. Since $S_{k-1}^L(z_t) < S_k^L(z_t)$ for every $k \geq 1$, $q_0^k > q_0^{k+1}$ if $w_0 < w_{k,k+1}$, and $q_0^k < q_0^{k+1}$ if $w_0 > w_{k,k+1}$.

As mentioned above, both indifference curves have a positive slope in the (w,q) space, i.e. are such that dq/dw along the curves. Moreover, this slope is steeper for C(k+1) than for C(k). We also know that C(k+1) goes through the point (w_{k+1}^L, q_{k+1}^L), which denotes the optimal combination of wage and queue length that workers target with their $k+1$-th application. Similarly, the indifference curve C(k) goes through the point (w_k^L, q_k^L). These pairs are equilibrium outcomes, located on firms' isoprofit curve for L-type workers. Queue lengths are increasing with wages along this isoprofit curve, which implies that $q_{k+1}^L > q_k^L$ because $w_{k+1}^L > w_k^L$. Hence, C(k) is increasing at a lower rate than C(k+1) and goes through a point, (w_k^L, q_k^L), that is to the south-west of (w_{k+1}^L, q_{k+1}^L), which belongs to C(k+1). These properties imply that the unique wage $w_{k,k+1}$ at which C(k) and C(k+1) intersect necessarily belongs to the interval (w_k^L, w_{k+1}^L).

Let $f_k : \mathbb{R} \to \mathbb{R}$ denote the function that associates to every wage w a queue length q such that $(w,q) \in C(k)$, i.e. such that the point (w,q) belongs to the indifference curve $C(k)$. It follows from the above analysis that, for any wage w in the interval (w_{k+1}^L, ∞), $f_{k+1}(w) > f_k(w)$.

The two wages we are interested in, $\underline{w}(a^L)$ and $\underline{w}(a^L+1)$, belong respectively to indifference curves C(a^L) and C(a^L+1). From the *lemons condition*, we know that both wages are higher than any wage that would be offered on a L-type market, in particular $w_{a^L+1}^L$. Thus, both wages are strictly greater than the wage at which C(a^L) and C(a^L+1) intersect. It follows that $\underline{w}(a^L)$ and $\underline{w}(a^L+1)$ belong to the interval of wages $(w_{a^L+1}^L, \infty)$ such that $f_{a^L+1}(w) > f_{a^L}(w)$ for any w in that interval.

Let $g^H : \mathbb{R} \to \mathbb{R}$ denote the function that associates to every wage w a queue length q such that $(w,q) \in IP-H$, i.e. such that the point (w,q) belongs to the isoprofit curve for H-type markets. g^L is increasing in w.

Wages $\underline{w}(a^L)$ and $\underline{w}(a^L+1)$ are respectively the intersections of functions $f_{a^L}(w)$ and $f_{a^L+1}(w)$ with $g^H(w)$ on the interval $(w_{a^L+1}^L, \infty)$. As f_{a^L+1} lies strictly above f_{a^L} on this interval, a strictly increasing function like $g^H(w)$ necessarily intersects with f_{a^L} at a lower wage than with $f_{a^L+1}(w)$. Therefore, $\underline{w}(a^L) < \underline{w}(a^L+1)$.