How much work experience do you need to get your first job?

The macroeconomic implications of bias against labor market entrants

Shisham Adhikari
University of California - Davis

Athanasios Geromichalos
University of California - Davis

Ateş Gürsoy
University of California - Davis

Ioannis Kospentaris
Athens University of Economics and Business

First Version: October 2023
This Version: June 2024

ABSTRACT

The first step in a worker's career is often particularly hard. Many firms seeking workers require experience in a related field, so a vicious circle is created, whereby an entry level job is required in order to get an entry level job. Consequently, entrant workers have lower job-finding rates and longer unemployment durations than the unemployed who have looked for a job in the past. To study the welfare implications of these observations, we consider a version of the DMP model where firms who match with entrant workers have to incur training costs. As a result, firms are biased against entrant workers, who, in turn, stay unemployed for a prolonged period of time, exposing themselves to a persistent skill loss shock. We use a calibrated version of the model to quantitatively assess the effectiveness of four government interventions, whose common goal is to reduce bias against entrant workers. We find that the most effective intervention takes the form of a subsidy that induces firms to rank entrants higher than experienced workers and that this policy brings the economy very close to the constrained efficient outcome.

JEL Classification: E24, E60, J24, J64
Keywords: search and matching, unemployment, labor market entrants, training, skills

Email: shadhikari@ucdavis.edu, ageromich@ucdavis.edu, agursoy@ucdavis.edu, ikospentaris@aueb.gr.

We are grateful to Efi Adamopoulou, Marios Angeletos, Andri Chassamboulli, Nicolas Caramp, Miro Gabrovski, Manolis Galenianos, Christos Genakos, Andreas Gulyas, Loukas Karabarbounis, Myrto Kalouptsidi, Minki Kim, Moritz Kuhn, Florian Madison, Guido Menzio, Victor Ortego-Marti, Theodore Papageorgiou, Giovanni Peri, Krzysztof Pytka, Guillaume Rocheteau, Nikos Thedoropoulos, as well as seminar participants at UC Davis, the University of Cyprus, the University of Mannheim, the West Coast Search and Matching Spring 2023 Workshop, the Second Athens Winter Econ Meeting, and CRETE 2023 for useful comments and suggestions.

1 Introduction

The first step in a worker's career is often particularly hard. Entrant workers have much lower job-finding rates and longer unemployment durations than the unemployed who have looked for a job in the past (Figure 1). Moreover, positions targeted to workers who have just entered the labor market tend to become a rarity. Indicatively, 35% of entry level positions posted since 2017 on LinkedIn required years of experience in a related field (see Section 2.1). So a vicious circle is created, whereby an entry level job is required in order to get an entry level job. These observations raise a number of important questions. First, why would firms choose to exclude from the applicant pool workers who are inexperienced, but may turn out to be extremely able? Second, what are the aggregate welfare implications of this bias against inexperienced workers, given that, by definition, *all* workers in the economy start their careers as inexperienced? This last question becomes especially important once we consider the recent literature in labor economics arguing that market conditions during the start of a worker's career have long lasting effects (Von Wachter, 2020). Finally, one wonders whether there is room for welfare improving government interventions and what form these interventions should take.

To study these questions, we augment the classic Diamond-Mortensen-Pissarides (DMP) model in several directions. First, we assume that firms that hire entrant/ inexperienced workers have to incur training expenses. This creates bias against these workers which we capture with a Petrongolo and Pissarides (2001) matching function with ranking. Second, we assume that entrants who stay unemployed for an extended time period are more likely to suffer skill loss. The combination of these two channels creates novel welfare trade offs, since the inherent bias against entrant workers increases their unemployment duration which, in turn, tends to lower their productivity. In this environment, an obvious market failure arises. Firms that hire inexperienced workers provide a *benefit to society* by helping these workers stay unemployed for a shorter period of time, thus reducing their exposure to the skill loss shock. (A shock which leaves a permanent "scar", thus affecting the entrant workers' productivity when they are hired by other firms later in their career.) However, firms cannot fully internalize this societal contribution, and ultimately choose to discriminate against entrant workers, causing a social welfare loss.

Given this market failure, there is obvious scope for government intervention. Since the root of the inefficiency is the inability of firms to fully recoup the training costs, which results in hiring bias against entrants, we consider government interventions whose common goal is to alleviate the bias and reduce the inexperienced workers' exposure to the skill loss shock. We use a calibrated version of the model to quantitatively study the effectiveness of three government interventions. The first intervention, which we dub "unbiased matching", bans discrimination against inexperienced workers by law. In the second intervention ("government subsidies"), the government raises taxes to subsidize firms that hire inexperienced workers. Finally, the third intervention ("internships") also bans discrimination but additionally explores the possibility that the compensation of entrant workers is determined exogenously by the government.

We find that all three government interventions improve aggregate welfare.This is true even though in all of them the aggregate unemployment rate is higher than in the benchmark economy with ranking. To explain the economics behind this result, let us begin with the first intervention, i.e., unbiased matching. Without the ability to discriminate against entrants, firms are effectively forced to incur larger training expenses. As a result, firm entry is discouraged and equilibrium unemployment is higher compared to the baseline economy. Despite this unintended consequence on aggregate unemployment, entrant workers have shorter unemployment spells and are less likely to suffer skill loss. The productivity gains from the latter channel are so large that aggregate welfare increases by 0.58%. The economics behind the second intervention, i.e., government subsidies, is similar, but the welfare increase is smaller (around 0.51%) because the tax needed to finance training subsidies reduces match surplus and further distorts entry.

The third intervention, i.e., internships, achieves all the benefits of the other two, since it also involves unbiased matching, but it suffers less from the downside of discouraged entry. With internships, entrant workers' wages are exogenous and treated as parameters whose level varies. For high enough wages, entrant workers are compensated almost as much as they would under Nash bargaining, and welfare levels are similar to those of Intervention 1. On the other extreme, if entrant wages are too low, firms realize they can hire these workers almost for free, which leads to inefficiently large levels of training and vacancy creation costs. In total, aggregate welfare has an inverse-U shape with respect to the entrants wage level, and it achieves its maximum (0.63% greater than the baseline) when entrants' wages are at intermediate levels. That is, a carefully designed internship scenario achieves the highest welfare among these interventions.

The common thread among the interventions considered so far is that they raise aggregate productivity by abolishing bias against entrant workers and, as a result, lowering their exposure to skill depreciation. In light of this finding, we consider an additional, fourth, intervention, in which the government "goes all the way" subsidizing the hiring of entrant workers so heavily that firms actually prefer to match with these workers rather than the experienced ones. The results of this intervention are striking: it improves aggregate welfare by 1.8%, an improvement almost three times larger than that in the internships scenario. This intervention has the largest impact among all four interventions because it directly confronts the problem of inexperienced workers spending a lot of time in unemployment upon entry.[1] Finally, we examine how close this fourth intervention can bring the economy to its efficient level by comparing it with the allocation of a social planner who also favors entrant workers. The comparison reveals that the fourth intervention brings the economy arbitrarily close to the constrained efficient outcome. The social planner implements a Hosios (1990)-type condition for our environment; however, it turns out that implementing the "correct" ranking (i.e., favoring entrant workers) is an order of magnitude more important than fine-tuning the level of vacancy creation.

As we have already explained, our baseline model adopts the Petrongolo and Pissarides (2001) specification, because it fits the stylized fact that motivates this paper: entrant workers have lower job-finding rates and longer unemployment spells. However, we also explore alternative approaches that have been used in the literature to study models of one-sided heterogeneity, most prominently competitive search. The main result of this analysis is that the alternative matching protocols fail to capture the aforementioned stylized fact. The reason is that under these protocols firms are able to secure an unrealistically large fraction of the (future) benefits generated by the training they provide to entrant workers. However, there is an extensive empirical literature (surveyed in Section 2.1) demonstrating that firms do not recoup the benefits of their training investments, thus providing suboptimal levels of training. Consequently, the alternative matching protocols predict that the job-finding rate of entrants will be higher than that of experienced unemployed workers, which directly contradicts the data presented in Figure 1.[2]

The rest of the paper is organized as follows. Section 2 describes the model, including a discussion of some key assumptions and a literature review, and Section 3 analyzes the baseline model with matching bias against inexperienced workers. Section 4 considers three government interventions intended to improve welfare and Section 5 describes the calibration strategy. Section 6 introduces the fourth intervention, discusses the quantitative effects of all policy interventions considered, and solves the social planner's problem. Section 7 explores two alternative matching specifications and Section 8 concludes.

[1] It should be pointed out that programs that target entrant workers are particularly popular in Europe. The Youth Employment Initiative, a large EU-sponsored program, directly finances young workers' apprenticeships, traineeships, job placements, and further education within 4 months of leaving school or becoming unemployed. Our model predicts that interventions of this kind generate sizeable welfare gains.

[2] Our analysis also uncovers a surprising result: the competitive search specification attains lower welfare than the baseline model, when the ranking order favors entrant workers. While we consider this a side result of the paper, it may be of interest to researchers who work with search models, since it confounds the standard expectation that the competitive search model attains maximum welfare.

2 The Model

We consider an extension of the Diamond-Mortensen-Pissarides framework. Time is continuous with an infinite horizon, and all agents discount future at rate r. The labor force is normalized to 1. Workers exit the labor market (retire) at Poisson rate δ, and each retired worker is immediately replaced by a new labor market entrant. New workers enter the labor market as unemployed. The retirement of workers, and their subsequent replacement by an entrant, is crucial in our model, because these new entrants will be a special category, and the length of time they spend being unemployed will have long-term consequences. There is a large mass of *ex ante* homogeneous firms who can enter the labor market with one vacancy. As is standard, the measure of active firms in equilibrium is determined by free entry.

Firms who decide to enter the labor market and search for workers must pay a flow recruiting cost c. Existing jobs get terminated at the job destruction rate λ. Generally, job matches produce an amount p of the numeraire good, but this productivity will be affected by the worker's specific type. Firms who have hired entrant (inexperienced) workers must pay a flow training cost κ until the match dissolves. This is our way of modeling the real world observation that firms are often biased against workers who do not have any working experience. Perhaps these new entrants were brilliant students, but they still need training to become *productive workers*. Thus, one possible interpretation is that the κ term is the number of hours other colleagues must spend with the inexperienced workers "showing them the ropes". (For a further justification of this assumption, see Section 2.1.) After losing their very first job workers become automatically experienced, and firms who shall hire them in the future will not need to pay the κ cost ever again.[3]

Since the effective productivity of an inexperienced worker is lower than that of an experienced worker (with κ representing the differential), firms are biased against inexperienced workers, and we will capture this by using the matching process of Petrongolo and Pissarides (2001). (However, we also consider alternative matching processes; details to follow.) As a result, entrant workers will tend to stay in the pool of unemployment for longer periods of time, which is precisely what we see in the data (Figure 1). Another crucial, and empirically relevant assumption we will make is that entrants who stay un-

[3] Our "training cost" story is not the only way to capture the fact that firms are biased against inexperienced workers. Another possibility is that employers have asymmetric information about a worker's quality. The asymmetric information problem would be less severe in the case of a worker who has previously held at least one job, since the firm could ask for a reference letter from a former employer. While our approach is certainly not the only way to go, it has two big advantages. First, it leads to a simple and tractable model. Second, it is extremely relevant and easier to quantify, as there is a large literature studying training costs in the labor market. See Section 2.1 for the relevant references.

employed for a prolonged period of time are in risk of suffering skill loss. We assume that this skill loss is permanent, and we often refer to this phenomenon as the "scar". To keep the model tractable, we will assume that the skill loss of entrant workers takes place stochastically, at Poisson rate γ; when that shock hits an inexperienced worker her productivity declines by an amount $\tilde{\kappa}$, and that skill/productivity loss follows that specific worker for the rest of her life. The skill loss and the "scar" assumptions are further discussed in Section 2.1.

To fix ideas and offer the reader a mnemonic rule that will help them comprehend the notation that follows, we now provide a description of the various worker types. We will refer to new entrants who just replaced a retired worker as type-0 workers (i.e, they have "0 working experience".) Type-0 workers enter the model as unemployed. If they find their first job quickly, they will become employed type-0 workers, and, as discussed, their effective productivity will be given by $p - \kappa$. After leaving their first job, type-0 workers permanently become type-1 workers, and their productivity at any future job will be equal to p. However, if type-0 workers stay unemployed for a long period of time, they are more likely to be hit by the skill-loss shock; if this happens (before they could find their first job), they will turn into type-$\tilde{0}$ (unemployed) workers. When these types find their first job, they will become type-$\tilde{0}$ employed workers, and their productivity will be equal to $p - \kappa - \tilde{\kappa}$, since these are inexperienced workers who need training (hence the $-\kappa$), *and* they have also suffered skill loss (hence the $-\tilde{\kappa}$). Since in the baseline model skill loss is permanent, when type-$\tilde{0}$ workers find and, eventually, lose their first job, they will turn into type-$\tilde{1}$ workers. This means that at any future job their productivity will be equal to $p - \tilde{\kappa}$. (These workers are now experienced, but the skill-loss scar remains.) This is our tractable way of capturing the empirically relevant observation that early-career conditions have long-lasting effects on workers' productivity and earnings.

To sum up, at any point in time there are $2^3 = 8$ types of workers. First, workers can be unemployed or employed. Next, they can be inexperienced or experienced, where an experienced worker is defined as one who has held (and lost) at least one job in the past. Finally, workers can be scarred or not scarred, depending on whether they got hit by the skill loss shock when they first entered the labor market. Notation-wise, the number 0 (1) will denote an inexperienced (experienced) worker, and the "tilde" symbol will denote (variables related to) a worker who suffered skill loss during her youth. For example, $\tilde{f}_0$ is the *job-finding rate* of an inexperienced worker who suffered skill loss, $\tilde{f}_1$ is the *job-finding rate* of an experienced worker who suffered skill loss at youth, w_0 is the *wage* of an inexperienced worker who has not suffered skill loss, u_1 is the measure of experienced (non-scarred) *unemployed* workers, $\tilde{e}_1$ is the measure of scarred experienced *employed* workers, and so on.

Given the description of the environment so far, it is obvious that the core of the model consists of homogeneous firms searching for workers of different types. Importantly, these types are not inherent to each worker; they depend on the worker's history, including the length of their unemployment spell during youth and whether they previously held a job. Deviating from the standard DMP model with homogeneous workers necessitates taking a stance on how firms meet with these different types.[4] A natural first attempt is to consider a model of competitive search, since these models were developed to deal with one-sided market heterogeneity. (See for example Montgomery 1991 and Rogerson, Shimer, and Wright 2005). However, it turns out that this modeling choice has counterfactual implications about the job-finding rates of different worker types. In the remainder of this section, we describe an alternative matching process that fits the job-finding data better, and in Section 2.1 we provide a justification for this modeling choice. The interested reader can find a detailed discussion of the model with competitive search in Section 7.

The matching process we adopt is a generalization of the Blanchard and Diamond (1994) *matching with ranking* proposed by Petrongolo and Pissarides (2001). This specification allows us to capture the fact that firms are biased against inexperienced and scarred workers, as these workers are less productive. This matching function exhibits bias against a certain type of unemployed workers who are considered less desirable. The main, and very simple idea of that matching technology is that the more desirable/productive types of workers get matched first, without being crowded out by the inferior types. When that "first round" of matching has concluded, the less desirable types of workers get a chance to match with firms. It should be pointed out that in Petrongolo and Pissarides (2001) there are only two types of unemployed workers, while in our paper there are four (type 0, $\tilde{0}$, 1, and $\tilde{1}$). However, their simple idea that more desirable unemployed workers get to match first, can be easily carried over to our analysis.

The firms' ranking of workers is based on workers' productivity, which, in turn, is affected by their experience and whether they are scarred. Type-1 workers are the most productive and, hence, the most desirable. It is also quite clear that type-$\tilde{0}$ workers, who are inexperienced and scarred, are the least desirable. In principle, it is not obvious whether firms would prefer workers of type 0 or workers of type $\tilde{1}$, because their ranking depends

[4] A more imaginative way to make this point is to paraphrase Lucas's quote about complete and incomplete markets. (In turn, Lucas was paraphrasing Tolstoy's Anna Karenina principle about happy and unhappy families.) "Homogeneous agent economies are all alike, but each heterogeneous agent economy is heterogeneous in its own individual way" (Lucas Jr, 1989).

on the magnitude of training costs (κ) versus the skill loss ($\tilde{\kappa}$). Since our calibration in Section 5 implies that $\tilde{\kappa} < \kappa$, we assume that firms prefer workers of type $\tilde{1}$ to type 0. In sum, firms rank workers in the following order: type $1 \succ$ type $\tilde{1} \succ$ type $0 \succ$ type $\tilde{0}$.

It turns out that the job-finding rates of the various worker types, implied by the Petrongolo and Pissarides (2001) matching process, can be conveniently expressed as functions of the *queue lengths* of the various types. Thus, we define

$$b_1 \equiv \frac{u_1}{v}; \quad \tilde{b_1} \equiv \frac{\tilde{u_1}}{v}; \quad b_0 \equiv \frac{u_0}{v}; \quad \tilde{b_0} \equiv \frac{\tilde{u_0}}{v}, \tag{1}$$

where v is the measure of vacancies in the economy. Notice that the each queue length is simply the unemployment-vacancy ratio for that particular worker type (which is the inverse of the market tightness, typically used in the baseline DMP model). Extending the Petrongolo and Pissarides (2001) methodology under a Cobb-Douglas specification in our framework, implies the following job-finding rates for each worker type:

$$f_1 = \frac{m(u_1, v)}{u_1} = {b_1}^{\alpha-1},$$

$$\tilde{f_1} = \frac{m(u_1 + \tilde{u_1}, v) - m(u_1, v)}{\tilde{u_1}} = \frac{(b_1 + \tilde{b}_1)^{\alpha} - b_1^{\alpha}}{\tilde{b}_1},$$

$$f_0 = \frac{m(u_1 + \tilde{u_1} + u_0, v) - m(u_1 + \tilde{u_1}, v)}{u_0} = \frac{(b_1 + \tilde{b}_1 + b_0)^{\alpha} - (b_1 + \tilde{b}_1)^{\alpha}}{b_0},$$

$$\tilde{f_0} = \frac{m(u_1 + \tilde{u_1} + u_0 + \tilde{u_0}, v) - m(u_1 + \tilde{u_1} + u_0, v)}{\tilde{u_0}} = \frac{(b_1 + \tilde{b}_1 + b_0 + \tilde{b}_0)^{\alpha} - (b_1 + \tilde{b}_1 + b_0)^{\alpha}}{\tilde{b}_0}.$$

The details of these derivations have been relegated to Appendix A. (There, we also report the rates at which firms meet the various types of workers.) Notice how the "ranking" manifests itself in the various job-finding rates. The job-finding rate of the most desirable workers, those of type 1, is a function only of the queue length for that particular type. But take the next most desirable group, type-$\tilde{1}$ workers: this type's job-finding rate is a function of their own queue length, $\tilde{b_1}$, as well as the queue length of the types that are "above" them in the ranking, b_1. This means, intuitively, that type-$\tilde{1}$ workers are crowded out by each other and by type-1 workers. In similar spirit, the least desirable type-$\tilde{0}$ workers are crowded out by all other types, including their own.

We close the model with a few more standard assumptions. After the matching has concluded and firms have met the various types of workers, the two parties negotiate over the wage using Nash Bargaining. We will let $\eta \in [0, 1]$ denote the bargaining power of the worker. All unemployed workers enjoy a benefit z, which we think of as the utility

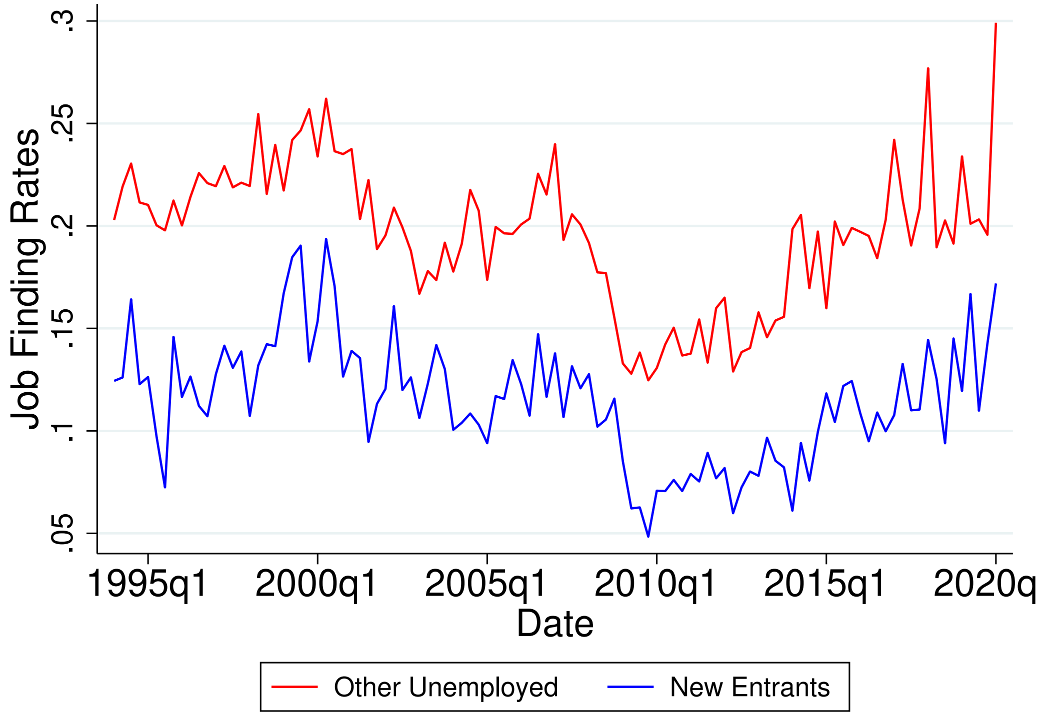

Figure 1. Job-finding rates for unemployed workers 16-24 years old. Calculations based on monthly data of the Current Population Survey for 1994 to 2020.

of leisure and/or the value of home production.[5] Notice that, with the exception of productivity and the consequent differences in job-finding rates, all the other parameters of the model $(\eta, z, \lambda, r, \delta)$ are independent of the worker's type. This is intentional, since we want the results to be driven *only* by differences in the workers' experience and whether they suffered skill loss during their youth, which is the focus of our paper.

2.1 Justification of Modeling Choices and Literature Review

Several elements of our model are relatively non-standard and combined together in a common framework for the first time. Therefore, it is important to provide justification for some of the model's novel ingredients. In particular, we discuss the following modeling choices: (i) biased matching based on Blanchard and Diamond (1994) and Petrongolo and Pissarides (2001), (ii) firm-provided training, and (iii) persistent skill loss during workers'

[5] While the range of realistic values for z is discussed in detail in Section 5, z is not in principle required to be lower than the productivity of every worker type. Consider for example a type-0 worker. That worker may well choose to work for a firm even if we had $z \geq p - \kappa$. The reason is that working for a firm would allow this worker to become experienced and secure a higher wage in the future.

first unemployment spell. We comment on each one of them in order.

First, the rationale for biased matching in favor of experienced workers is straightforward. As can be seen in Figure 1, the job-finding rates of entrants are much lower than the job-finding rates of the experienced unemployed workers between 16 and 24 years old.[6] The generalization of the Blanchard and Diamond (1994) ranking by Petrongolo and Pissarides (2001) provides a natural and tractable way to deliver job-finding rates consistent with this fact. Moreover, there are numerous pieces of anecdotal evidence indicating that many firms rank workers without previous experience lower than more experienced applicants. For example, an analysis of almost four million job postings on Linked-In since late 2017 showed that 35% of postings for entry-level positions asked for years of prior relevant work experience.[7] As we have already mentioned, adopting a model of random matching with ranking is not the only modeling option available, but it is the most appropriate for our purposes. In Section 7, we provide a detailed discussion of two alternative matching protocols (competitive search and random search with segmented submarkets) and show that they fail to capture the stylized fact illustrated in Figure 1.

Second, a crucial feature of our framework is the assumption that firms need to devote resources to train entrant workers. These resources capture the fact that the experienced workers have to take time away from production to teach the necessary traits to the entrants. Examples of these traits include the ability to work in teams, follow instructions, understand and complete a task, or how to network. There is a plethora of recent empirical papers documenting the importance of firm-provided training in the labor market. Ma, Nakab, and Vidart (2022), in a cross-country study, document that firm-provided training is a key determinant of workers' human capital. (See Herkenhoff, Lise, Menzio, and Phillips (2024) for a complementary interpretation of the workers' human capital component.) Faccini and Yashiv (2022), using German and Swiss data, estimate that training costs (in the form of opportunity costs incurred by managers and coworkers incurred to make new hires as productive as experienced workers) are the dominant source of all hiring costs. Bertheau, Larsen, and Zhao (2023), using linked survey-administrative data from Denmark, find that one-third of employers consider the time to train new recruits as a major obstacle.

Moreover, there is a literature that highlights the empirical relevance of the externality identified in this paper, namely, the idea that firms underhire inexperienced workers, since they are not fully compensated for the social benefit of the training they provide

[6] The difference is almost the same for workers between 16 and 64 years old as well. However, almost 85% of entrants is between 16 and 24, hence we present the data for this age group.

[7] https://www.bbc.com/worklife/article/20210916-why-inexperienced-workers-cant-get-entry-level-jobs

(Becker, 2009; Acemoglu, 1997; Acemoglu and Pischke, 1999).[8] The literature sometimes refers to this as the "future employer externality"; see Acemoglu (1997) and Lentz and Roys (2024). An important empirical finding in this line of work is that firms provide general training which is not fully offset by lower wages; see Acemoglu and Pischke (1998), Loewenstein and Spletzer (1998), and Autor (2001). Pallais (2014) aptly surveys this literature and concludes that "...neither the theoretical nor the empirical literature shows that firms recoup the full value of their training investments resulting in their providing the optimal level of training" (p. 3568).

Third, there is compelling empirical evidence documenting that skill loss during the early unemployment spells has persistent negative effects on a worker's career. Arellano-Bover (2022) shows that early-career unemployment shocks have negative effects on measured cognitive skills several decades later. Similarly, Dinerstein, Megalokonomou, and Yannelis (2022), using quasi-experimental variation in unemployment duration at the beginning of teachers' careers in Greece, document strong negative effects of the length of unemployment on teachers' performance measured by students' test scores. More generally, there is a large empirical literature, summarized by Von Wachter (2020), that highlights the persistence of the effects of labor market conditions upon entry for young workers on multiple outcomes later in their careers.

In the model, we have assumed that the unemployment spells after the first one have no effect on a worker's productivity. We have made this assumption for various reasons. First, the consequences of unemployment for the skills of experienced workers is well-studied elsewhere and its inclusion would only make the model unnecessarily complicated.[9] More importantly, there is evidence that skill loss may be of limited importance in older ages. For instance, Cohen, Johnston, and Lindner (2023) find no indication for a decline in skills over the unemployment spell in the overall population in Germany and for the older workers in the US. In the authors' own words, "This suggests that the negative consequences of unemployment might be a more relevant concern *at younger ages*" (p. 5, emphasis added).

To conclude, we should highlight that it is the *interaction* of firm provided training with skill loss during early unemployment that creates novel implications in our paper. As we explained above, each one of these mechanisms on its own is well-studied and its implications are well-understood. In our model, however, the combination of the two el-

[8] Moen and Rosén (2004) and Lentz and Roys (2024) provide conditions for the efficient provision of training to workers.

[9] For references on skill loss in unemployment, see Pissarides (1992), Ljungqvist and Sargent (1998), Coles and Masters (2000), Ortego-Marti (2016), Flemming (2020), and Kospentaris (2021), among others.

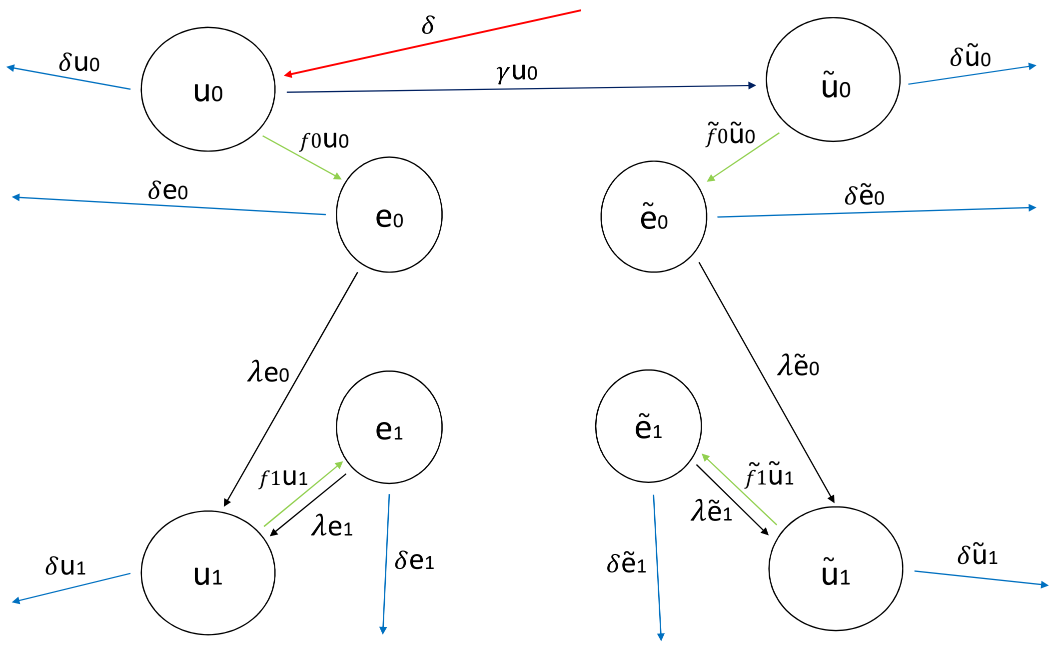

Figure 2. Worker flows in and out of the various states.

ements creates a novel intertemporal trade-off: to save on training costs, firms hire fewer entrant workers today, which, in turn, lowers the workers' future productivity. To the best of our knowledge, we are the first to study the theoretical and quantitative implications of this trade-off.

3 Analysis of the Model: Matching with Ranking

3.1 Beveridge curves

Having described the economic environment, we are ready to proceed with the analysis of the model, starting with the derivation of the Beveridge curves. For this task, it is useful to inspect Figure 2, which illustrates the worker flows in and out of every state. While at first glance the figure may look complicated (workers could be in one of eight possible states), the logic is simple. New entrants come into the labor market as type-0 unemployed workers at rate δ; this is indicated by the red arrow at the top of the graph. Then, at any state of the world, workers could get hit by the retirement shock and exit the labor market; these are the light blue arrows pointing away from the eight "bubbles" representing the various states. There are also four black arrows starting from employment and pointing to unemployment bubbles; clearly, these are flows initiated by the job destruction shock. Notice that job destruction always leads to the bubble of "experienced unemployed", which may be scarred ($\tilde{u_1}$) or not scarred (u_1). Green arrows indicate workers who found a job and are moving from unemployment to employment. As discussed, the rate at which this transition takes place is different for each worker type and governed by the Petrongolo and Pissarides (2001) matching process. Finally, the dark blue arrow starting at u_0 and pointing to $\tilde{u_0}$ captures the crucial group of workers who stayed in the "inexperienced unemployed" pool for too long and got hit by the skill loss shock.

Equating the flows in and out of each state, and after some algebra, we can show that the steady state measure of workers in the various states are as follows:

$$
\begin{aligned}
u_0 &= \frac{\delta}{\delta + \gamma + f_0}, \\
\tilde{u_0} &= \frac{\gamma}{\delta + \tilde{f}_0} \cdot \frac{\delta}{\delta + \gamma + f_0}, \\
e_0 &= \frac{f_0}{\delta + \lambda} \cdot \frac{\delta}{\delta + \gamma + f_0}, \\
\tilde{e_0} &= \frac{\tilde{f}_0}{\delta + \lambda} \cdot \frac{\gamma}{\delta + \tilde{f}_0} \cdot \frac{\delta}{\delta + \gamma + f_0}, \\
u_1 &= \frac{\lambda f_0}{(\gamma + \delta + f_0)(\delta + \lambda + f_1)}, \\
e_1 &= \frac{f_1}{\delta + \lambda} \cdot \frac{\lambda f_0}{(\gamma + \delta + f_0)(\delta + \lambda + f_1)}, \\
\tilde{u_1} &= \frac{\lambda \gamma \tilde{f}_0}{(\delta + \tilde{f}_0)(\gamma + \delta + f_0)(\delta + \lambda + \tilde{f}_1)}, \\
\tilde{e_1} &= \frac{\tilde{f}_1 \lambda \gamma \tilde{f}_0}{(\delta + \lambda)(\delta + \tilde{f}_0)(\gamma + \delta + f_0)(\delta + \lambda + \tilde{f}_1)}.
\end{aligned}
$$

3.2 Value Functions

We move to the steady state value functions, and we start with the firms. We will let q denote the arrival rate of workers to firms, and we will follow the usual notation, e.g, $\tilde{q_0}$ will stand for the arrival rate of an inexperienced worker who has suffered skill loss. The various q's are derived in Appendix A. The value function of a vacant firm is given by

$$
rV = -c + q_0(J_0 - V) + \tilde{q_0}(\tilde{J}_0 - V) + q_1(J_1 - V) + \tilde{q_1}(\tilde{J}_1 - V).
$$

Of course, free entry implies that in equilibrium we must have $V = 0$, therefore, we can state the free entry condition as

$$c = q_0 J_0 + \tilde{q_0}\tilde{J}_0 + q_1 J_1 + \tilde{q_1}\tilde{J}_1. \tag{2}$$

We also have four value functions for productive firms in the various states, i.e., for firms who matched with the four different types of workers (type 0, $\tilde{0}$, 1, and $\tilde{1}$). These are given as follows:

$$rJ_0 = p - \kappa - w_0 - \lambda J_0 - \delta J_0, \tag{3}$$
$$r\tilde{J}_0 = p - \kappa - \tilde{\kappa} - \tilde{w_0} - \lambda \tilde{J}_0 - \delta \tilde{J}_0, \tag{4}$$
$$rJ_1 = p - w_1 - \lambda J_1 - \delta J_1, \tag{5}$$
$$r\tilde{J}_1 = p - \tilde{\kappa} - \tilde{w_1} - \lambda \tilde{J}_1 - \delta \tilde{J}_1. \tag{6}$$

Next, consider the value functions of workers in the various states. Let U (W) denote the value function of an unemployed (employed) worker. The remaining notation is standard. (For example, $\tilde{W}_1$ is the value function of a worker who is employed, has had some work experience, but was hit by the skill loss shock during her youth.) The value functions for unemployed workers in the various states are given by:

$$rU_0 = z + f_0(W_0 - U_0) + \gamma(\tilde{U_0} - U_0) - \delta U_0, \tag{7}$$
$$r\tilde{U_0} = z + \tilde{f}_0(\tilde{W}_0 - \tilde{U_0}) - \delta \tilde{U_0}, \tag{8}$$
$$rU_1 = z + f_1(W_1 - U_1) - \delta U_1, \tag{9}$$
$$r\tilde{U_1} = z + \tilde{f}_1(\tilde{W}_1 - \tilde{U_1}) - \delta \tilde{U_1}. \tag{10}$$

The value functions for employed workers in the various states are given by:

$$rW_0 = w_0 + \lambda(U_1 - W_0) - \delta W_0, \tag{11}$$
$$r\tilde{W}_0 = \tilde{w_0} + \lambda(\tilde{U_1} - \tilde{W}_0) - \delta \tilde{W}_0, \tag{12}$$
$$rW_1 = w_1 + \lambda(U_1 - W_1) - \delta W_1, \tag{13}$$
$$r\tilde{W}_1 = \tilde{w_1} + \lambda(\tilde{U_1} - \tilde{W}_1) - \delta \tilde{W}_1. \tag{14}$$

Notice that inexperienced workers who lose their first job now move to the pool of experienced unemployed workers. (That is precisely why the terms U_1 and $\tilde{U_1}$ appear on the right-hand side of equations (11) and (12)).

Having described the value functions of all economic agents in detail, we are now ready to study the bargaining problems in the various types of meetings.

3.3 Bargaining problems

Bargaining in a type-1 meeting

We begin with the description of the terms of trade in a meeting between a firm and an unemployed worker of type-1, which, as we shall see, is the simplest case. Solving the standard Nash bargaining problem implies that the following condition must be satisfied:

$$(1-\eta)(W_1 - U_1) = \eta J_1. \tag{15}$$

This condition simply states that each party will enjoy a fraction of the total surplus of the match, and that fraction will be equal to her bargaining power. (Recall that η is the bargaining power of the worker.) Replacing the value functions W_1 and J_1 from (13) and (5), respectively, allows us to write the wage of a type-1 worker as $w_1 = \eta p + (1-\eta)(r+\delta)U_1$. Substituting U_1 from equation (9) into this expression yields

$$w_1 = \eta p + (1-\eta)z + (1-\eta)f_1(W_1 - U_1) = \eta p + (1-\eta)z + \eta f_1 J_1,$$

where the second equality follows from (15). Substituting J_1 one more time from equation (5), and solving for w_1, delivers the final version of our "wage curve" for type-1 workers:

$$w_1 = \frac{\eta p(r+\lambda+\delta+f_1) + (1-\eta)z(r+\lambda+\delta)}{r+\lambda+\delta+\eta f_1}. \tag{16}$$

Clearly, this is a relationship between the wage for type-1 workers and their job arrival rate, which, in turn, depends on firm entry and market tightness.

Bargaining in a type-$\tilde{1}$ meeting

Next, consider the bargaining problem between a firm and a worker who is experienced but suffered skill loss during her youth. Once again, we must have:

$$(1-\eta)(\tilde{W}_1 - \tilde{U}_1) = \eta \tilde{J}_1.$$

As in the case of type-1 workers, we can replace the functions $\tilde{W}_1$ and $\tilde{J}_1$ from (14) and (6) into the last expression. Following identical steps, and after some algebra, one can easily derive the wage curve for type-$\tilde{1}$ workers:

$$\tilde{w_1} = \frac{\eta(p-\tilde{\kappa})(r+\lambda+\delta+\tilde{f}_1)+(1-\eta)z(r+\lambda+\delta)}{r+\lambda+\delta+\eta\tilde{f}_1}. \tag{17}$$

Once again, we obtain a relationship between the wage for type-$\tilde{1}$ workers and their job arrival rate, which, in turn, depends on firm entry and market tightness.

Bargaining in type-$\tilde{0}$ meeting

We now move to the bargaining problem between a firm and an inexperienced worker who suffered skill loss. In this case the surplus sharing rule is given by

$$(1-\eta)(\tilde{W}_0-\tilde{U}_0)=\eta\tilde{J}_0.$$

As is standard, we first replace the value functions $\tilde{W}_0$ and $\tilde{J}_0$ from (12) and (4), respectively, which allows us to write the wage of a type-$\tilde{0}$ worker as

$$\tilde{w_0}=\eta(p-\kappa-\tilde{\kappa})-\lambda(1-\eta)(\tilde{U_1}-\tilde{U_0})+(1-\eta)(r+\delta)\tilde{U_0}. \tag{18}$$

Unlike the previous cases, where the wage depended on the value function of unemployment for that specific type (only), here $\tilde{w_0}$ depends on both $\tilde{U_0}$ and $\tilde{U_1}$, and specifically on their difference $\tilde{U_1}-\tilde{U_0}$. To deal with this, subtract (8) from (10) to obtain

$$\tilde{U_1}-\tilde{U_0}=\frac{\tilde{f}_1(\tilde{W}_1-\tilde{U}_1)-\tilde{f}_0(\tilde{W}_0-\tilde{U}_0)}{r+\delta}. \tag{19}$$

To obtain a useful expression for the term $\tilde{W}_1-\tilde{U}_1$, that now appears in (19), subtract (10) from (14), to get

$$\tilde{W}_1-\tilde{U}_1=\frac{\tilde{w_1}-z}{r+\delta+\lambda+\tilde{f}_1}. \tag{20}$$

Substitute equation (20) into (19), and the resulting outcome into equation (18), and, after some some algebra, one can arrive at the wage curve for the type-$\tilde{0}$ worker, specifically:

$$\begin{aligned}\tilde{w_0}=\frac{1}{r+\delta+\eta\tilde{f}_0}\big[&\eta(p-\kappa-\tilde{\kappa})(r+\delta+\tilde{f}_0)+\\&+\frac{(r+\delta+\lambda)(r+\delta+\tilde{f}_1)}{r+\delta+\lambda+\tilde{f}_1}(1-\eta)z-\frac{\lambda(1-\eta)\tilde{f}_1}{r+\delta+\lambda+\tilde{f}_1}\tilde{w_1}\big].\end{aligned} \tag{21}$$

Inspection of the last wage curve reveals that, in this case, the wage for type-$\tilde{0}$ workers is not only a function of this type's job arrival rate (as was the case for type-1 and type-$\tilde{1}$ workers). The wage $\tilde{w_0}$ also depends on the wage that type-$\tilde{1}$ workers make. The intuition is clear. When a type-$\tilde{0}$ worker meets a firm, working for that firm is the step that will allow her to move out of the "inexperienced" state, and earn the wage $\tilde{w_1}$ for the rest of her life. This is precisely, why the term $\tilde{w_1}$ enters equation (21) with a minus: a higher (future) wage $\tilde{w_1}$ induces the type-$\tilde{0}$ worker to be more eager to accept a lower (current) wage $\tilde{w_0}$, since that lower wage comes together with the opportunity of abandoning the bad "inexperienced" stage once and for all.

Bargaining in type-0 meeting

The last type of meeting is the one between a firm and an unskilled worker who has not yet been hit by the skill loss shock. The surplus sharing rule is given by

$$(1-\eta)(W_0 - U_0) = \eta J_0.$$

Following standard steps, substitute W_0 and J_0 from (11) and (3), respectively, to write the wage of a type-0 worker as

$$w_0 = \eta(p-\kappa) - (1-\eta)\lambda(U_1 - U_0) + (1-\eta)(r+\delta)U_0.$$

Just like in the case of type-$\tilde{0}$ workers (and unlike the cases of type-1 and type-$\tilde{1}$ workers), the wage for the types under consideration (also) depends on the differential term $U_1 - U_0$. Since the steps for deriving the final version of the wage curve are virtually identical to the case of type-$\tilde{0}$ workers presented above, we will skip the details and move directly to the final formula:

$$\begin{aligned} w_0 = {} & \frac{r+\delta+\gamma+f_0}{r+\delta+\gamma+\eta f_0}\eta(p-\kappa) + \frac{(r+\lambda+\delta)(r+f_1+\delta)(r+\gamma+\delta)}{(r+\delta)(r+\delta+\gamma+\eta f_0)(r+\delta+\lambda+f_1)}(1-\eta)z + \\ & + \frac{\gamma\tilde{f}_0\eta(p-\kappa-\tilde{\kappa}-\tilde{w_0})}{(r+\delta)(r+\delta+\gamma+\eta f_0)} - \frac{(1-\eta)\lambda f_1(r+\gamma+\delta)w_1}{(r+\delta)(r+\delta+\gamma+\eta f_0)(r+\delta+\lambda+f_1)}. \end{aligned} \tag{22}$$

The wage curve for type-0 workers admits an interpretation that is similar to the one following the wage curve for type-$\tilde{0}$ workers, i.e., equation (21). The wage that type-0 workers are willing to accept is not just a function of their job arrival rate: w_0 also depends on w_1 *and* $\tilde{w_0}$. Why w_0 depends on w_1 should now be obvious, given our earlier discussion: the type-0 worker realizes that if she agrees to work for that firm, she will be able to move out of the inexperienced state and earn the wage w_1 henceforth. Why does w_0 also depend on $\tilde{w_0}$? When the type-0 worker agrees to work for the firm with which she has matched, she realizes that she will never again be subject to the skill loss shock. Thus, what the worker *would* have made if she turned down the firm's offer and continued searching (and being subject to the skill loss shock), i.e., the wage $\tilde{w_0}$, enters the currently negotiated wage through her outside option.

3.4 Definition of Steady State Equilibrium

We conclude this section with a formal definition of equilibrium.

Definition 1. A steady state equilibrium in our model is a list of wages for the four types of workers $(w_0, \tilde{w_0}, w_1, \tilde{w_1})$, a measure of vacant firms v, and measures of unemployed and employed workers in the various states $(u_0, \tilde{u_0}, u_1, \tilde{u_1}, e_0, \tilde{e_0}, e_1, \tilde{e_1})$, satisfying the free entry condition (2), the four wage curves (16), (17), (21), and (22), and the eight Beveridge curves reported at the end of Section 3.1.[10]

4 Government Interventions

The discussion so far reveals that our model is characterized by a prominent market failure. Firms who hire entrant workers provide a *public good* to society by transforming them into experienced and, hence, (more) productive workers. However, the firms cannot internalize the societal benefits of this public service, and they rationally choose to discriminate against inexperienced workers. This bias increases the inexperienced workers' unemployment duration, which, in turn, raises their exposure to skill loss, a skill loss which permanently scars the workers, thus affecting their productivity when they are hired by other firms later in their lifetime.

It is obvious that in this environment there is scope for government intervention. Since the root of the inefficiency lies in the firms' decision to underhire inexperienced workers (see also Pallais (2014) and the references therein), we consider three possible government interventions whose common goal is to alleviate the bias against this group of workers and reduce their exposure to the skill loss shock, with the intention of increasing welfare. We start with a short description of the idea behind each of these three interventions, and then we analyze them in detail in the rest of this section. A *vis-à-vis* comparison of the effectiveness of each intervention will follow in Section 6.

[10] Implicit in this definition are the queue lengths $(b_0, \tilde{b_0}, b_1, \tilde{b_1})$, defined in equation (1) as functions of the various unemployment measures and v. In turn, these queue lengths are used to determine the various job-finding rates $(f_0, \tilde{f}_0, f_1, \tilde{f}_1)$, which appear in the Beveridge curves, as well as the firms' worker-finding rates $(q_0, \tilde{q_0}, q_1, \tilde{q_1})$ (described in Appendix A), which appear in the job creation curve.

Intervention 1: "Unbiased matching". We dub the first intervention "unbiased matching", as it describes the case where the government makes it *illegal* for firms to discriminate against any group of workers. Even though one could argue that this intervention is somewhat unrealistic (firms have the right to not hire less productive workers), it is still interesting, from a theoretical point of view, to study a benchmark model without biased matching, which has been a staple of the analysis so far, and appears to be the root of the ongoing inefficiency. Comparing the baseline model with the economy under Intervention 1 can tell us how much welfare can improve if the firms' bias against inexperienced workers was eradicated. In technical terms, the unbiased matching version of the model simply replaces the Petrongolo and Pissarides (2001) matching process with a standard Pissarides (2000) matching function where all workers meet firms at the same rate.

Intervention 2: "Government subsidies". Our second intervention is one where the government raises funds to subsidize firms who hire less productive workers. Subsidies are designed so that firms are effectively *indifferent* among the various types of workers, so that they *choose* to not discriminate against any type of workers. Thus, one can think of Intervention 2 as a market-based way of achieving what Intervention 1 achieves by law.[11]

Intervention 3: "Internships". The third intervention explores the possibility that the wage of entrant workers is not determined endogenously in the model, but chosen exogenously by the government. We dub this intervention "internships" because it assumes that inexperienced workers are not compensated based on their true productivity, but they are treated as trainees whose salary is predetermined and exogenous. We think this is a crucial element of an internship, although we realize that other important elements of real-world internships are absent from our model. In any case, the term "internship" here is just a tag that will summarize Intervention 3. This intervention allows firms to employ inexperienced workers at lower salaries, effectively lowering the bias against these workers and mitigating the inefficiency described in the beginning of this section.

4.1 Intervention 1: Unbiased matching

Beveridge Curves

Consider first the Beveridge curves for this version of our model. When it is illegal for firms to discriminate against certain types of workers, all workers match with firms at the same rate. All the Beveridge curves reported at the end of Section 3.1 are still valid,

[11] Despite this similarity, the two versions of the model do not lead to the same equilibrium outcomes, and the two interventions are not equally effective at improving welfare. For details see Section 6.

but the various matching rates are now equal. That is, $f_0 = \tilde{f}_0 = f_1 = \tilde{f}_1 = f$, where the common arrival rate f is now given by a standard (unbiased) matching process, i.e.,

$$f = \frac{m(u, v)}{u}, \tag{23}$$

with u representing the total mass of unemployed workers (of all types).

Value Functions

Next, we move to the value functions under the regime of Intervention 1, starting with the firms. With unbiased matching, the probability that a firm meets a worker of a specific type depends only on that type's relative representation in the pool of unemployed. Letting q denote the arrival rate of a(ny) worker to the typical firm, i.e., $q = m(u, v)/v$, the value function of a vacant firm is given by

$$rV = -c + q\left[\frac{u_0}{u}(J_0 - V) + \frac{\tilde{u_0}}{u}(\tilde{J}_0 - V) + \frac{u_1}{u}(J_1 - V) + \frac{\tilde{u_1}}{u}(\tilde{J}_1 - V)\right].$$

As is standard, free entry implies that in equilibrium $V = 0$, therefore, we can state the free entry condition as

$$c = \frac{q}{u}\left(u_0 J_0 + \tilde{u_0}\tilde{J}_0 + u_1 J_1 + \tilde{u_1}\tilde{J}_1\right). \tag{24}$$

Even though the process through which firms meet workers is different compared to the baseline model of Section 3, once a firm has met a specific type of worker, the value functions for productive firms in the various states remain the same, i.e., they are still given by equations (3)-(6).

Moving on to the workers, the value functions for unemployed workers in the various states are given by:

$$rU_0 = z + f(W_0 - U_0) + \gamma(\tilde{U}_0 - U_0) - \delta U_0, \tag{25}$$

$$r\tilde{U}_0 = z + f(\tilde{W}_0 - \tilde{U}_0) - \delta \tilde{U}_0, \tag{26}$$

$$rU_1 = z + f(W_1 - U_1) - \delta U_1, \tag{27}$$

$$r\tilde{U}_1 = z + f(\tilde{W}_1 - \tilde{U}_1) - \delta \tilde{U}_1. \tag{28}$$

Notice that these expressions are almost identical to the value functions (7)-(10) reported in Section 3.2, with the only difference being that the various arrival rates of that section have now been replaced by the common rate f, defined in equation (23).

The last set of value functions for this model specification concerns employed workers. Since employed workers have already matched with a firm, the different matching process assumed in this section will not affect their employment value functions, which are still given by equations (11)-(14) in Section 3.2.

Bargaining problems

As the discussion so far reveals, the only parts of the analysis that are affected by the adoption of the new (unbiased) matching process are those that take place *before* a firm and a worker match. Consequently, all the derivations of the wage curves in Section 3.3 remain valid, with the only difference being that the various arrival rates will now be replaced by the common arrival rate f. This observation also sheds some light on the economic insights of this intervention. The government, under Intervention 1, does not intervene in the labor market to change the way in which firms and workers produce or negotiate over the wages. It only intervenes by stating that discriminating against any type of worker, at the recruiting stage, is illegal. By doing this, the government ensures that *all* types of workers have the same matching rate, which, in turn, ensures that entrant/inexperienced workers will not stay unemployed for a prolonged period of time, thus risking a skill loss that will scar them for the rest of their career. We will discuss the effectiveness of this intervention, and how it compares to the alternatives, in Section 6.

Given that the derivations of the wage curves in Section 3.3 remain unaltered, we will not repeat them here, and we will only report the wage curves for the four types of workers, reminding the reader that they are identical to ones reported in equations (16), (17), (21), and (22), once one has replaced the various f's with the common arrival rate f defined in equation (23). More precisely, we have:

$$w_1 = \frac{\eta p(r+\delta+\lambda+f) + (1-\eta)z(r+\delta+\lambda)}{r+\delta+\lambda+\eta f}, \tag{29}$$

$$\tilde{w_1} = \frac{\eta(p-\tilde{\kappa})(r+\delta+\lambda+f) + (1-\eta)z(r+\delta+\lambda)}{r+\delta+\lambda+\eta f}, \tag{30}$$

$$\tilde{w_0} = \frac{r+\delta+f}{r+\delta+\eta f}\left[\eta(p-\kappa-\tilde{\kappa}) + \frac{r+\delta+\lambda}{r+\delta+\lambda+f}(1-\eta)z - \frac{\lambda(1-\eta)f}{(r+\delta+\lambda+f)(r+\delta+f)}\tilde{w_1}\right], \tag{31}$$

$$\begin{aligned} w_0 = &\frac{r+\delta+\gamma+f}{r+\delta+\gamma+\eta f}\eta(p-\kappa) + \frac{(r+\delta+\lambda)(r+\delta+f)(r+\delta+\gamma)}{(r+\delta)(r+\delta+\gamma+\eta f)(r+\delta+\lambda+f)}(1-\eta)z \\ &+ \frac{\gamma f\eta(p-\kappa-\tilde{\kappa}-\tilde{w_0})}{(r+\delta)(r+\delta+\gamma+\eta f)} - \frac{(1-\eta)\lambda f(r+\delta+\gamma)w_1}{(r+\delta)(r+\delta+\gamma+\eta f)(r+\delta+\lambda+f)}. \end{aligned} \tag{32}$$

Definition 2. A steady state equilibrium, under Intervention 1, is a list of wages for the four types of workers $(w_0, \tilde{w_0}, w_1, \tilde{w_1})$, a measure of vacancies v, and measures of unemployed and employed workers in the various states $(u_0, \tilde{u_0}, u_1, \tilde{u_1}, e_0, \tilde{e_0}, e_1, \tilde{e_1})$, satisfying the free entry condition (24), the four wage curves (29), (30), (31), and (32), and eight Beveridge curves, which are the ones reported at the end of Section 3.1, after one replaces the various f's with the common arrival rate f defined in equation (23).

4.2 Intervention 2: Government subsidies

Unlike Intervention 1, where the government could impose unbiased matching by law, here the government raises funds to subsidize firms who hire less productive workers. Specifically, a firm who hires a type-0 worker will receive a (flow) subsidy σ_0, a firm who hires a type-$\tilde{0}$ worker will receive a subsidy $\tilde{\sigma_0}$, and a firm who hires a type-$\tilde{1}$ worker will receive a subsidy $\tilde{\sigma_1}$. To raise funding for these subsidies, every active firm pays a flat (lump-sum) tax equal to τ. The aforementioned subsidies are designed so that firms are *effectively indifferent* among the various types of workers. Thus, one can think of Intervention 2 as a market-based (as opposed to legislative) way of achieving unbiased matching.

Beveridge Curves

Even though under Intervention 2 this happens for different reasons (subsidies rather than anti-discriminating laws), the end result is that firms are indifferent among the various types of workers. This simply means that all workers face identical job-finding rates, and, consequently, all the Beveridge curves remain the same as the ones described in Section 4.1. This, in turn, means that relevant Beveridge curves are the ones reported at the end of Section 3.1, but with $f_0 = \tilde{f_0} = f_1 = \tilde{f_1} = f$ (where f was defined in equation (23)).

Value Functions

Next, we move to the value functions under the regime of Intervention 2, starting with the firms. Using our standard notation, the value functions of firms who have matched with the various types of workers are given by

$$rJ_0 = p - \kappa - w_0 + \sigma_0 - \tau - \lambda J_0 - \delta J_0, \tag{33}$$

$$r\tilde{J_0} = p - \kappa - \tilde{\kappa} - \tilde{w_0} + \tilde{\sigma_0} - \tau - \lambda \tilde{J_0} - \delta \tilde{J_0}, \tag{34}$$

$$rJ_1 = p - \tau - w_1 - \lambda J_1 - \delta J_1, \tag{35}$$

$$r\tilde{J_1} = p - \tilde{\kappa} - \tilde{w_1} + \tilde{\sigma_1} - \tau - \lambda \tilde{J_1} - \delta \tilde{J_1}. \tag{36}$$

However, recall that here the various subsidies are designed to make firms indifferent among the various types of workers. This implies that $J_0 = \tilde{J_0} = J_1 = \tilde{J_1} = J$. This greatly simplifies the value function of a vacant firm, which is now given by

$$rV = -c + q(J - V),$$

where, as before, $q = m(u, v)/v$, is the worker-finding rate of the typical firm. Free entry implies that in equilibrium we must have

$$c = qJ. \tag{37}$$

Since all types of workers under Intervention 2 have identical job-finding rates (albeit for different reasons), the value functions for unemployed workers reported in Section 4.1, i.e., equations (25)-(28), remain valid. The value functions for employed workers are also identical to the ones reported in Section 4.1, which, in turn, are identical to the value functions given by equations (11)-(14) in Section 3.2.

Before we move to the bargaining problems, we present an auxiliary result that will significantly simplify our task.

Lemma 1. *Under Intervention 2 all workers types receive the same wage, i.e.,* $w_0 = \tilde{w}_0 = w_1 = \tilde{w}_1 = w$.

Proof. See Appendix A. □

The proof of the lemma has been relegated to the appendix, but the statement is quite intuitive. Since the government's intervention makes firms indifferent among the various types of workers, it turns out that all the workers will make the same wage in equilibrium. It should also be obvious why Lemma 1 simplifies the analysis: under Intervention 2, we will only have to solve one bargaining problem, instead of four.

Bargaining problem(s)

Since the various J_i terms are all equal to each other, and since the bargaining problem in each type of meeting would imply $(1 - \eta)(W_i - U_i) = \eta J_i$, we must have that the various $W_i - U_i$ terms are also equal to each other. Thus, instead of solving four distinct bargaining problems, in this specification of the model we only need to solve one bargaining problem. As one can see in detail in the proof of Lemma 1, the various $W_i - U_i$ terms are all equal to

$$W - U = \frac{w - z}{r + \lambda + \delta + f},$$

where w is the common wage established in Lemma 1. As for the J term, we can simply replace it from the free entry condition, i.e., equation (37). Then, the standard bargaining protocol, prescribing that $(1-\eta)(W-U)=\eta J$, here implies that

$$w = z + \frac{\eta}{1-\eta}\frac{c(r+\delta+\lambda+f)}{q}, \tag{38}$$

which is our (unique) wage curve under Intervention 2.

The size of the tax and the various subsidies

Before we proceed to the definition of equilibrium under Intervention 2, we must characterize the size of the various subsidies and the flat tax. Exploiting equations (33)-(36), and keeping in mind that $J_0 = \tilde{J}_0 = J_1 = \tilde{J}_1 = J$, we can deduce that

$$p-\kappa-w_0-\tau+\sigma_0 = p-\kappa-\tilde{\kappa}-\tilde{w_0}-\tau+\tilde{\sigma_0} = p-\tau-w_1 = p-\tilde{\kappa}-\tau-\tilde{w_1}+\tilde{\sigma_1}.$$

But since Lemma 1 has established that all the wages must be equal (and since all firms pay the same flat tax τ), we have:

$$\sigma_0 = \kappa; \quad \tilde{\sigma_0} = \kappa+\tilde{\kappa}; \quad \tilde{\sigma_1} = \tilde{\kappa}. \tag{39}$$

Again, this is intuitive. The only way in which the government can make firms indifferent among the various types of workers is by fully covering the "cost" associated with hiring a less productive type (i.e., anyone other than type-1).

The last item we need to specify is the flat tax rate. A balanced government budget constraint implies that

$$\tau = \frac{\kappa e_0 + (\kappa+\tilde{\kappa})\,\tilde{e_0} + \tilde{\kappa}\tilde{e_1}}{e_0+\tilde{e_0}+e_1+\tilde{e_1}}. \tag{40}$$

Definition 3. A steady state equilibrium, under Intervention 2, is a (common) wage, w, for all types of workers, a list of subsidies $(\sigma_0, \tilde{\sigma_0}, \tilde{\sigma_1})$, a flat tax, τ, paid by all active firms, a measure of vacancies v, and measures of unemployed and employed workers in the various states $(u_0, \tilde{u_0}, u_1, \tilde{u_1}, e_0, \tilde{e_0}, e_1, \tilde{e_1})$. The three types of subsidies are described in equation (39), and the tax satisfies equation (40). The remaining equilibrium variables satisfy the free entry condition (37), the wage curve (38), and the eight Beveridge curves reported at the end of Section 3.1, after one replaces the various f's with the common arrival rate f defined in equation (23).

4.3 Intervention 3: Internships

Our third and last intervention is the one dubbed "internships". Here, we explore the possibility that the wage of entrant/inexperienced workers is not determined endogenously in the model (i.e., by Nash bargaining), but it is chosen exogenously by the government. Let us denote the wages of type-0 and type-$\tilde{0}$ workers by w_0 and $\tilde{w_0}$, respectively. For now, we treat them as exogenous parameters, but in Section 6 we will discuss how changes in these two terms affect equilibrium welfare. All experienced workers (of type-1 and type-$\tilde{1}$) will continue to make wages that are determined by Nash bargaining.

A natural question that arises is whether firms still have an incentive to discriminate against certain types of workers. Think, for example, of type-0 workers. These workers must be trained by the firms (their productivity is $p - \kappa$), but their wage is given exogenously by w_0, which could be very low (perhaps even zero). Whether firms would prefer to hire a type-1 to a type-0 worker depends on the size of κ and w_0. In fact, as long as κ is not too large, the government could always choose w_0 to be low enough, so that firms prefer to match with a type-0 worker and discriminate against type-1 workers. Since, for now, we have decided to treat w_0 and $\tilde{w_0}$ as parameters whose value can change (thus tilting the firms' preferences towards the various worker types), we will strive for the maximum degree of flexibility, by assuming that firms *do not discriminate* against any type of worker.[12] This assumption also promotes tractability and allows a direct comparison of Intervention 3 with Interventions 1 and 2. (Recall that under Interventions 1 and 2 firms do not discriminate against any types of workers, either by law or by choice.)

Beveridge Curves

Given our modeling choice to assume no bias in matching, all types of workers face identical job-finding rates, and the Beveridge curves remain the same as in Sections 4.1 and 4.2. This, in turn, means that relevant Beveridge curves are the ones reported at the end of Section 3.1, but with $f_0 = \tilde{f}_0 = f_1 = \tilde{f}_1 = f$ (where f was defined in equation (23)).

Value Functions

Next, we move to the value functions under the regime of Intervention 3, starting with the firms. Since we assume unbiased matching, the probability that a firm meets a worker of a specific type depends only on that type's relative representation in the pool of unemployed. Letting q denote the arrival rate of a(ny) worker to the typical firm, as we did under Interventions 1 and 2, the value function of a vacant firm is given by

$$rV = -c + q\left[\frac{u_0}{u}(J_0 - V) + \frac{\tilde{u_0}}{u}(\tilde{J}_0 - V) + \frac{u_1}{u}(J_1 - V) + \frac{\tilde{u_1}}{u}(\tilde{J}_1 - V)\right].$$

[12] Our framework would allow us to explore many different versions, including the less standard case where firms discriminate against type-1 workers; this would be relevant if inexperienced workers are not too unproductive, and they are very cheap. Intervention 4, which we study in Section 6.4, is in that spirit.

Free entry implies that $V = 0$, therefore, we can state the free entry condition as

$$c = \frac{q}{u}\left(u_0 J_0 + \tilde{u_0}\tilde{J}_0 + u_1 J_1 + \tilde{u_1}\tilde{J}_1\right). \tag{41}$$

The value functions for productive firms in the various states are still given by equations (3)-(6), but there is an important difference. In Section 3, equations (3) and (4), the terms w_0 and $\tilde{w_0}$ represented endogenous variables; here the value functions appear identical, but these terms represent exogenous policy parameters.

Moving on to the workers, the value functions for unemployed workers in the various states are still given by equation (25)-(28) in Section 4.1, and the value functions for employed workers in the various states are still described by equations (11)-(14) in Section 3. Again, it is useful to point out that the only (conceptual) difference is that the terms w_0 and $\tilde{w_0}$ appearing in equations (11) and (12) are endogenous variables, while here these same terms are understood to be exogenous policy parameters.

Bargaining problems

As we have explained, under Intervention 3 we only need to solve the bargaining problem in the two types of meetings with experienced workers. Consider first a meeting between a firm and a type-1 worker. As is standard, the Nash protocol requires that $(1-\eta)(W_1 - U_1) = \eta J_1$. Replacing W_1 from (13) and J_1 from (5), and following the standard steps, we conclude that the wage curve for type-1 workers, under Intervention 3, is given by

$$w_1 = \frac{\eta p(r+\delta+\lambda+f) + (1-\eta)z(r+\delta+\lambda)}{r+\delta+\lambda+\eta f}. \tag{42}$$

Next, consider a meeting between a firm and a type-$\tilde{1}$ worker. Once again, the Nash bargaining protocol requires $(1-\eta)(\tilde{W_1} - \tilde{U_1}) = \eta\tilde{J_1}$. As in the case of type-1 workers, we can replace the functions $\tilde{W_1}$ and $\tilde{J_1}$ from (14) and (6) into the surplus splitting rule. Following identical steps, and after some algebra, one can easily derive the wage curve for type-$\tilde{1}$ workers, under Intervention 3:

$$\tilde{w_1} = \frac{\eta(p-\tilde{\kappa})(r+\delta+\lambda+f) + (1-\eta)z(r+\delta+\lambda)}{r+\delta+\lambda+\eta f}. \tag{43}$$

Notice that, the two wage curves are identical expect for the fact that they adjust for the worker's productivity, i.e., p versus $p - \tilde{\kappa}$.

Definition 4. A steady state equilibrium, under Intervention 3, consists of two wages for experienced workers $(w_1, \tilde{w_1})$, a measure of vacancies v, and measures of unemployed and employed workers in the various states $(u_0, \tilde{u_0}, u_1, \tilde{u_1}, e_0, \tilde{e_0}, e_1, \tilde{e_1})$, satisfying the free entry condition (41), the two wage curves (42) and (43), and eight Beveridge curves, which are the ones reported at the end of Section 3.1, after one replaces the various f's with the common arrival rate f defined in equation (23).

5 Calibration

We calibrate the benchmark model with ranking at a monthly frequency. Several parameters are set exogenously to their direct empirical counterparts or by following the literature. We normalize the match output p to 1 and set the discount rate r to 0.0042, consistent with an annual interest rate of 5%. We set the elasticity of the aggregate matching function with respect to unemployment α to 0.65, towards the upper end of the estimates reported in Petrongolo and Pissarides (2001). Following Shimer (2005), the workers' bargaining weight η is also set equal to the elasticity of the aggregate matching function.

Finally, we set the skill loss shock intensity γ to 1/6, which implies that an unemployed entrant spends on average six months in unemployment before their skills depreciate. We chose the six months interval for two reasons, one conventional and one substantial: first, the definition of "long-term unemployment" according to the Bureau of Labor Statistics is consecutive unemployment of 27 weeks and over. Second, and more substantial, it is well known from the duration dependence literature that the job-finding probability strongly decreases for the first six months in unemployment and flattens out afterwards (see Jarosch and Pilossoph 2019 and Kospentaris 2021 among many others). Hence, based on the job-finding duration profile, it seems that the six months threshold is a discrete event for the transition from short- to long-term unemployment and we treat it as such in our calibration.

Parameter	Description	Value	Source
r	Discount Rate	0.0042	Annual Interest Rate of 5%
p	Match Output	1	Normalization
α	Matching Function Elasticity	0.65	Petrongolo and Pissarides (2001)
η	Worker Bargaining Power	0.65	Shimer (2005)
γ	Skill Loss Intensity	1/6	Duration Dependence Literature

Table 1: Exogenously Set Parameters.

The remaining six parameters are calibrated through the model and their values are reported in Table 2. The vacancy creation cost c, the worker exit/entry rate δ, and the separation rate λ are chosen to make the model consistent with the following labor market moments, respectively: i) the aggregate unemployment rate ($u_0 + \tilde{u_0} + u_1 + \tilde{u_1}$), ii) the fraction of entrants in the unemployment pool ($(u_0 + \tilde{u_0})/(u_0 + \tilde{u_0} + u_1 + \tilde{u_1})$), and iii) the fraction of long-term unemployed among entrants ($\tilde{u_0}/(u_0 + \tilde{u_0})$). Next, we follow Hall and Milgrom (2008) and set the opportunity cost of employment z to 71% of average worker productivity. Regarding the size of skill loss $\tilde{\kappa}$, we employ the estimates of Ortego-Marti (2016, 2017) which imply a monthly 1.22% drop in worker wages while the worker is unemployed.

Parameter	**Description**	**Value**
c	Vacancy Cost	1.3984
δ	Worker Exit Rate	0.0023
λ	Separation Rate	0.0379
z	Unemployment Value	0.6517
$\tilde{\kappa}$	Skill Loss Scar	0.0676
κ	Training Cost	0.1228

Table 2: Internally Calibrated Parameters.

Finally, to discipline the training cost parameter κ we use numbers reported by training professionals for US businesses. The specialist publication *Training Magazine* asks businesses from several industries about their expenses devoted to employee training and reports the results in their Training Industry reports.[13] The annual training expenses per employee were \$ 1,075 in 2017 and reached \$ 1,207 in 2022. To be conservative, we chose κ to match annual training expenses of \$ 1,000 per employee, which is 0.75% of the US GDP. This means that total training expenses are of a similar order of magnitude as total vacancy creation costs which are usually estimated to be 1-2% of GDP (see, e.g., Michaillat and Saez 2021). Using a different calibration strategy, Masui (2023) also estimates training costs to be close to vacancy creation costs, which provides a sanity check for our strategy. As can be seen in Table 3, the model exactly matches the calibration target (the difference between model-implied and data moments is in the order of 10^{-8}).

6 Quantitative Results

In this section, we present the quantitative effects of the government interventions analyzed in Section 4. We focus on the main aggregate variables of interest: i) the aggregate output minus total vacancy costs (Y, which is also our main measure of welfare, since

[13] Available here: `https://trainingmag.com/2022-training-industry-report/`.

Target	Data	Model
Unemployment Rate	5.8%	5.8%
Unemployed Entrants/Unemployed	9%	9%
Long-term Unemployed Entrants/ Unemployed Entrants	28%	28%
Value of Non-Employment/Average Productivity	71%	71%
Wage Loss for Six Months Unemployment	7.1%	7.1%
Training Expenses/GDP	0.75%	0.75%

Table 3: Matching the Calibration Targets.

agents are risk neutral), ii) the aggregate unemployment rate (u), iii) the job-finding rate of each worker type ($f_1, \tilde{f}_1, f_0, \tilde{f}_0$), and iv) the wage of each worker type ($w_1, \tilde{w_1}, w_0, \tilde{w_0}$). The quantitative results are presented in Table 5 as percentage differences from the baseline ranking economy and analyzed in the rest of this section.

Y	u	f_1	$\tilde{f}_1$	f_0	$\tilde{f}_0$	w_1	$\tilde{w_1}$	w_0	$\tilde{w_0}$
0.9315	5.8%	0.80	0.48	0.43	0.42	0.9904	0.9201	0.4488	0.8192

Table 4: Baseline Economy: Matching with Ranking.

Before analyzing the results of each government intervention, we summarize the results of the baseline economy in Table 4 (i.e., the model of Section 2 evaluated at the calibrated parameters of Section 5). A few features of the benchmark economy are worth noting. First, the order of the job-finding rates follows the order of the productivity exhibited by the different worker types ($p > p - \tilde{\kappa} > p - \kappa > p - \kappa - \tilde{\kappa}$). There is, however, a discrete jump between the job-finding rate of type-1 workers and the other unemployed, which are relatively close in magnitude. That is, firms strongly prefer to match with experienced workers without a scar than with the other unemployed types. Second, entrant workers of type 0 have particularly low wages compared to the other types. Strikingly, they receive lower wages than their scarred counterparts of type $\tilde{0}$ who also have lower productivity. This is consistent with the intuition we gave in Section 3.3 that entrant workers are eager to leave their current state and, as a result, willing to accept very low wages to incentivize firms to hire them. This force is even stronger for workers of type 0 than $\tilde{0}$ because they have more to gain by leaving the entry state as soon as possible to enjoy a career as non-scarred experienced workers.

Variable	Unbiased	Subsidies	Internships	Type-0 Bias	Planner
Y	0.58%	0.51%	0.63%	1.831%	1.833%
u	31.50%	39.14%	19.51%	24.33%	34.04%
f_1	-39.12%	-42.73%	-32.50%	-36.28%	-41.10%
$\tilde{f}_1$	2.51%	-3.57%	13.64%	-27.91%	-33.44%
f_0	12.32%	5.65%	24.52%	454.83%	411.43%
$\tilde{f}_0$	14.21%	7.43%	26.61%	-21.05%	-27.20%
w_1	-0.54%	-3.09%	-0.41%	-1.20%	-
$\tilde{w_1}$	0.03%	4.32%	0.14%	-1.19%	-
w_0	81.58%	113.87%	70.06%	103.01%	-
$\tilde{w_0}$	-2.53%	17.17%	-20.44%	-1.53%	-

Table 5: Quantitative Effects of Interventions and the Constrained Optimum. Each number is the percentage difference between the value of the variable in the equilibrium with a particular intervention versus the value of the variable in the baseline economy of matching with ranking.

6.1 Unbiased Matching

In this scenario, it is illegal for firms to rank workers of different types. Essentially, there is a common job-finding rate for all workers, given by a standard Cobb-Douglas matching function (equation 23). Intuitively, this intervention forces firms to incur larger training expenses, since they have no way of discriminating against entrants. As a result, entry drops and the aggregate unemployment rate is more than 30% higher than the baseline economy with ranking (second column of Table 5). At the same time, this intervention features a substantially higher welfare level than the ranking economy. The reason is the sizeable improvement in the job-finding prospects of entrant workers: f_0 and $\tilde{f}_0$ are 12% and 14% higher than their baseline levels, respectively. This means that entrant workers spend less time in unemployment and thus suffer less from lower skill loss compared to the baseline model. This trumps the productivity losses due to larger training costs and results in a 0.58% increase in aggregate welfare. Finally, the wages of type-0 entrants also improve considerably making them the big winners of this intervention.[14]

[14] In terms of levels, wages follow the order of worker productivity ($w_1 > \tilde{w_1} > w_0 > \tilde{w_1}$) but recall that the wage of type-0 workers is by far the lowest in the ranking economy. Hence, in percentage terms, workers of type 0 see the biggest percentage improvement in wages.

6.2 Government Subsidies

As explained in Section 4, this intervention is the market-oriented way of implementing what the ranking ban achieves through legislation: unbiased matching. To incentivize firms to not rank workers in hiring, the government subsidizes training and skill loss costs (see equation 39) and balances its budget with a uniform lump-sum tax τ equal to 2.85% of average match productivity. Given the similarity of outcomes between Interventions 1 and 2, the economics behind the two reforms is also similar: incentivizing firms to hire entrants increases training costs, lowers entry but saves on skill loss and raises aggregate welfare (third column of Table 5). The fact that the government takes away part of the surplus in tax revenue explains the relatively smaller effect on welfare and the job-finding rates of entrants compared to Intervention 1. It is of interest, however, that the wage effects of the two interventions differ: the common wage of Intervention 2 results in larger gains for workers of types $\tilde{1}$, 0, and $\tilde{0}$, as well as larger losses for type 1. In other words, there is a trade-off in terms of job-finding rates and wages which depends on whether unbiased matching is achieved through legislation (with smaller wage but larger job-finding effects) or market-based incentives (with larger wage but smaller job-finding effects).

6.3 Internships

The third intervention we look at is "internships": entrant workers' wages are exogenously fixed and discrimination against entrants is not allowed.[15] To ease the presentation, we fix $\tilde{w}_0 = z$ and show how welfare changes for various values of w_0 (the results in the fourth column of Table 5 are for the value of w_0 that maximizes aggregate welfare). None of the economics of the intervention rests on this assumption though; that is, fixing $\tilde{w}_0$ at some other level or varying both $\tilde{w}_0$ and w_0 would deliver the same insights.

Figure 3 graphs the aggregate welfare for different values of w_0. As the entrants' wage increases, aggregate welfare initially increases but after some critical value decreases, yielding an inverse-U shape. For relatively large values of w_0, the intervention yields familiar insights: entrants are expensive, firms are discouraged to enter to avoid training costs, and wage increases lead to larger skill loss and lower aggregate welfare. For relatively low values of w_0, however, the intervention creates the opposite problem: entrants are too cheap and entry is inefficiently high. As a result, when w_0 increases, ag-

[15] As explained in Section 4.3, there are several reasons behind this choice. Most importantly, it is the most transparent formulation and allows a direct comparison with the first two interventions. It should be noted, however, that our model can easily handle internships with endogenous ranking.

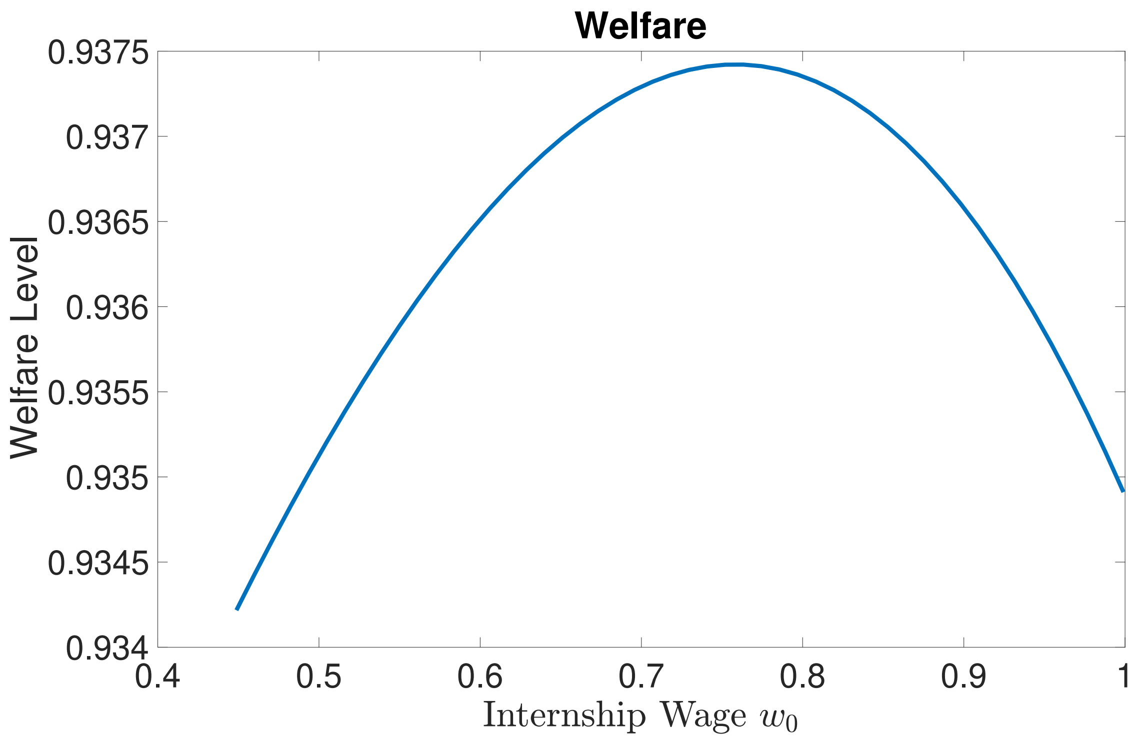

Figure 3. Aggregate Welfare for Various Levels of Entrant Wages.

gregate welfare also increases as the economy saves in vacancy creation and training costs (this is the well-known reasoning of Hosios 1990). In total, there is a welfare maximizing level of w_0, which is 70% higher than the corresponding level in the baseline economy.

The most important takeaway is that a carefully designed "internship" achieves the highest welfare level among all three interventions because it delivers the largest increase in the job-finding rate of type-0 workers. At the same time, every worker type gains from internships other than the experienced type-1 workers. All in all, these results offer a rationalization of internships in terms of raising aggregate welfare, which may partially explain the popularity of actual internship positions found in modern labor markets.

6.4 A fourth intervention

The discussion in Section 6 thus far conveys that government interventions that eradicate bias against entrant workers can achieve significant welfare improvements through the reduced unemployment duration of these workers, which lowers their exposure to skill depreciation shocks and raises aggregate productivity. In light of this finding, one wonders about the effects of an intervention that "goes all the way" and actually induces firms to be biased *in favor* of entrant workers. A simple way to achieve this is to consider a government that subsidizes type-0 workers generously enough to tilt the firms' ranking in favor of type-0 over type-1 workers. In this environment, matching is still characterized by the Petrongolo and Pissarides (2001) specification, but, due to the government's decision to (heavily) subsidize type-0 workers, that worker type is now ranked number one in the firms' preferences. Specifically, we consider a government that provides a subsidy $\sigma_0 = \kappa(1+\rho)$ to every firm that hires a type-0 worker.[16] We treat ρ as a parameter that captures the government's generosity, but, importantly, any $\rho > 0$ implies that firms now rank type-0 workers first. Thus, the firms' ranking of workers is: type $0 \succ$ type $1 \succ$ type $\tilde{1} \succ$ type $\tilde{0}$. The arrival rates, value functions, and various equilibrium conditions under Intervention 4 are described in detail in Appendix B for the interested reader. We dub this intervention "type-0 bias" and present its effects in the fifth column of Table 5.[17]

The results of this intervention are striking: it improves aggregate welfare by 1.831%, an improvement three times larger than the one achieved with Internships. The reason for this sizeable improvement is of course the treatment of type-0 workers, whose job-finding rate increases by an order of magnitude (from 0.43 in the baseline economy to 2.41). On the contrary, all other worker types lose from the intervention, since both their job-finding rates and wages substantially decrease, while aggregate unemployment also increases. The success of Intervention 4 lies on the fact that it directly confronts the problem of type-0 workers spending a lot of time in unemployment upon entry. As a result, it has the largest impact on the economy's aggregate productivity among all four interventions. It is important to mention that even though programs of this kind are not popular in the US, they are actually quite common in Europe. The Youth Employment Initiative, for example, is one of the main financial resources of the European Union to support young people who are not in education or employment. The total budget of the Youth Employment Initiative was €8.9 billion for the period 2014-2020 to directly finance young workers' apprenticeships, traineeships, job placements, and further education within 4 months of leaving school or becoming unemployed.[18] Based on our results, interventions of this kind are expected to have large welfare gains, even though they may raise the unemployment rates of other worker groups.

[16] Since the subsidy pays for itself, we assume that every firm who is actively producing must pay a flat tax, so $\tau = \kappa(1+\rho)e_0/(e_0 + \tilde{e_0} + e_1 + \tilde{e_1})$, which is the direct analogue of equation (40).

[17] For this exercise we have set $\rho = 5\%$ because it is small enough to not distort entry incentives too much but still sizeable to incentivize firms to hire type-0 workers.

[18] The details of the program are available here: `https://ec.europa.eu/social/main.jsp?langId=en&catId=89&newsId=9793&furtherNews=yes`.

6.5 The Social Planner's Allocation

Having showed that Intervention 4 (type-0 bias) yields sizeable welfare gains, it is natural to ask how close it brings the economy to its efficient level. However, given the non-standard features of our model, characterizing this "efficient level" is not obvious. The most straightforward way to go about describing the efficient allocation is to repeat the "textbook" exercise (i.e., Chapter 8 of Pissarides 2000), where the social planner chooses the aggregate measure of vacancies, while being constrained by the matching technology, which in our case is the Petrongolo and Pissarides (2001) matching with ranking. Importantly, however, the ranking the planner is subject to ought to be *aligned* with the ranking described under Intervention 4, since this is the spirit of our exercise.[19] The social planner's problem and its solution are formally presented in the following proposition.

Proposition 1. *The social planner solves the following problem:*

$$\max_{v} \int_0^\infty e^{-rt}[e_0(p-k) + e_1 p + \tilde{e}_1(p-\tilde{k}) + \tilde{e}_0(p-k-\tilde{k}) + (u_0+u_1+\tilde{u}_1+\tilde{u}_0)z - cv]\, dt,$$

subject to

$$\dot{u}_0 = \delta - u_0(\delta+\gamma+f_0), \qquad \dot{e}_0 = f_0 u_0 - (\delta+\lambda)e_0,$$
$$\dot{u}_1 = \lambda(e_0+e_1) - (\delta+f_1)u_1, \qquad \dot{e}_1 = f_1 u_1 - (\delta+\lambda)e_1,$$
$$\dot{\tilde{u}}_1 = \lambda(\tilde{e}_0+\tilde{e}_1) - (\delta+\tilde{f}_1)\tilde{u}_1, \qquad \dot{\tilde{e}}_0 = \tilde{f}_0\tilde{u}_0 - (\delta+\lambda)\tilde{e}_0,$$
$$\dot{\tilde{u}}_0 = \gamma u_0 - (\delta+\tilde{f}_0)\tilde{u}_0, \qquad \dot{\tilde{e}}_1 = \tilde{f}_1\tilde{u}_1 - (\delta+\lambda)\tilde{e}_1,$$

where

$$f_0 = b_0^{a-1},$$
$$f_1 = \frac{(b_0+b_1)^a - b_0^a}{b_1},$$
$$\tilde{f}_1 = \frac{(b_0+b_1+\tilde{b}_1)^a - (b_0+b_1)^a}{\tilde{b}_1},$$
$$\tilde{f}_0 = \frac{(b_0+b_1+\tilde{b}_1+\tilde{b}_1)^a - (b_0+b_1+\tilde{b}_1)^a}{\tilde{b}_0}.$$

[19] Studying the problem of a planner who is subject to a different ranking of worker types (say, the one described in the baseline model) would imply that the planner does not have the tools to address the main inefficiency in the model, i.e., the fact that firms do not internalize the social benefit of hiring inexperienced workers. Since the goal of this section is to see how close Intervention 4 brings the economy to the constrained efficient allocation, making sure that the two specifications share the same ranking is the only meaningful way to carry out this comparison.

The steady state solution to the planner's problem is summarized by unemployment levels for each type, $(u_0, u_1, \tilde{u}_0, \tilde{u}_1)$*, employment levels for each type,* $(e_0, e_1, \tilde{e}_0, \tilde{e}_1)$*, and a measure of vacancies,* v*. The planner's solution is uniquely determined by the eight Beveridge curves reported at the end of Section 3.1, where the arrival rates are stated above, and an equation that governs the optimal level of vacancies which is given below:*

$$f_0 u_0(\mu_{e_0} - \mu_{u_0}) + \tilde{f}_0 \tilde{u}_0(\mu_{\tilde{e}_0} - \mu_{\tilde{u}_0}) + f_1 u_1(\mu_{e_1} - \mu_{u_1}) + \tilde{f}_1 \tilde{u}_1(\mu_{\tilde{e}_1} - \mu_{\tilde{u}_1}) = \frac{vc}{1-a} e^{-rt}. \quad (44)$$

The various μ *terms are the co-state variables in the social planner's problem, and the differentials* $(\mu_{e_0} - \mu_{u_0})$, $(\mu_{\tilde{e}_0} - \mu_{\tilde{u}_0})$, $(\mu_{e_1} - \mu_{u_1})$*, and* $(\mu_{\tilde{e}_1} - \mu_{\tilde{u}_1})$ *are described in detail in the proof.*

Proof of Proposition 1. See Appendix C □

Inspection of Proposition 1 reveals that the social planner chooses the number of vacancies to maximize social welfare, defined as production appropriately weighed by the measure of each employed type, plus enjoyment of leisure by all unemployed types, net of vacancy costs. The first eight constraints indicate that the planner must respect the laws of motion of the measures of the various worker types. The four latter constraints ensure that the job finding rates of the four unemployed types are governed by the Petrongolo and Pissarides (2001) matching process, where, however, the ranking has been tweaked to give priority to type-0 workers (thus matching the ranking of Intervention 4). A natural interpretation of equation (44) follows the equalization of marginal social cost with marginal social benefit of vacancy creation. The right hand-side of equation (44) captures the effective discounted cost of creating an additional vacancy. The left hand-side of this equation illustrates the social gain an additional vacancy generates by assisting each worker type to transition from unemployment to employment. Specifically, the term $f_i u_i$ is the mass of unemployed of type i that transition to employment and the difference of the shadow prices captures the social benefit of that transition.

The quantitative implications of the planner's allocation are shown in the last column of Table 5. The main takeaway from these results is that the planner implements what Intervention 4 does: raises the job-finding rate of type-0 workers by an order of magnitude to save their skills from depreciation. To make this happen, the planner has to give priority to entrants compared to the other unemployed, something Intervention 4 achieved by compensating firms for ranking type-0 workers above the other unemployed. As a result, the job-finding rates of the other worker types are susbtantially lower in the constrained efficient allocation than in the baseline economy. Table 5 demonstrates that Intervention 4 brings the economy particularly close to the constrained efficient outcome. Intuitively, the social planner implements a Hosios-type condition for our environment; however, it becomes apparent that implementing the "correct" ranking is an order of magnitude more important than fine-tuning the level of vacancy creation.

7 Discussion of Alternative Matching Models

As discussed in Section 2, a key ingredient of our model is that homogeneous firms search for different types of workers. Importantly, these types are not inherent to each worker; they depend on the worker's history, including the length of their unemployment spell during youth. Since, by design, our model departs from the standard DMP model with homogeneous workers, we need to take a stance on how firms meet these different worker types. In Section 2.1, we argued that adopting the "matching with ranking" approach of Petrongolo and Pissarides (2001) is an appropriate choice, since it allows us to capture a stylized fact that is key for the paper's motivation: job-finding rates of entrant workers are much lower than the job-finding rates of experienced unemployed workers (Figure 1). Now that we have described the main results under the baseline (matching with ranking) specification, it is a good time to take a step back and illustrate why some other candidate specifications would not have worked as well. We do so for a model of competitive search with wage posting. At the end of the section we also shortly discuss the specification of random search (and bargaining) with segmented markets.

Given our environment with homogeneous firms and heterogeneous workers, adopting a competitive search model seems like a natural choice. After all, studying search markets with one-sided heterogeneity was one of the main reasons this literature was developed. (See for example the pioneering work of Montgomery 1991.) Surprisingly, it turns out that such a modeling choice is not appropriate for our research question, since it produces *counterfactual* predictions regarding the job-finding rates of different worker types. Following the standard approach in the literature (e.g., Rogerson et al. 2005), we consider a model where firms can enter one of four different submarkets (one for each type of unemployed) by posting a wage to attract workers. Within each submarket there is a standard Cobb-Douglas matching function, $m(u_i, v_i) = u_i^a v_i^{1-a}$, and the arrival rates for firms and workers are determined by the market tightness $\theta_i \equiv v_i/u_i$. Any match within a given submarket must take place at the posted wage, w_i, and the four v_i terms are effectively determined by free entry. Proposition 2 describes the competitive search equilibrium, and it is followed by a discussion of the main results of this section.

Proposition 2. *a) The equilibrium allocation for the competitive search specification of our model,* *summarized by four market tightnesses,* $(\theta_0, \tilde{\theta}_0, \theta_1, \tilde{\theta}_1)$*, satisfies the following four conditions:*

$$(r+\delta+\lambda)\,\theta_1^a + a\theta_1 = (1-a)\frac{p-z}{c}, \tag{45}$$

$$(r+\delta+\lambda)\,\tilde{\theta}_1^{\;a} + a\tilde{\theta}_1 = (1-a)\frac{p-\tilde{\kappa}-z}{c}, \tag{46}$$

$$(r+\delta+\lambda)\tilde{\theta}_0^{\;a} + \alpha\tilde{\theta}_0 - \frac{a\lambda}{r+\delta}(\tilde{\theta}_1-\tilde{\theta}_0) = (1-a)\frac{p-\kappa-\tilde{\kappa}-z}{c}, \tag{47}$$

$$(r+\delta+\lambda)\theta_0^{\;a} + \alpha\theta_0 - \frac{a\lambda}{r+\delta}(\theta_1-\theta_0) + \frac{a\gamma(r+\delta+\lambda)}{(r+\delta)(r+\delta+\gamma)}(\tilde{\theta}_0-\theta_0) = \frac{(1-a)(p-\kappa-z)}{c}. \tag{48}$$

b) The competitive search equilibrium allocation coincides with the allocation of a planner who chooses the measure of vacancies in each submarket (v_i) *to maximize social welfare.*

Proof of Proposition 2. See Appendix D.1 □

Part (a) of the proposition describes the competitive search equilibrium allocation. Part (b) reveals that despite the non-standard environment, with heterogeneity driven by individual worker history, our model admits its own version of a very well-known result: the competitive search allocation coincides with the solution of the (appropriately chosen) social planner's problem (Moen, 1997; Shimer, 1996).[20] When the planner chooses the optimal v_1 and, effectively, θ_1, (equation 45), she balances out the benefit of a marginal increase in the tightness of that market (workers match faster and produce $p > z$) with the cost of such an increase (the additional recruiting cost c and the forgone unemployment benefit z). Importantly, the optimal θ_1 depends *only* on the economic conditions (productivity, recruiting costs, etc) that characterize the submarket for workers of type 1. An analogous statement holds for equation (46) and the optimal $\tilde{\theta}_1$.

Inspection of equations (47) and (48) indicates that this is not the case for the submarkets of type-0 and $\tilde{0}$ workers. For example, the optimal $\tilde{\theta}_0$ depends on the economic conditions in submarket $\tilde{0}$ (the terms $p-\kappa-\tilde{\kappa}, z, c$), but also on the conditions characterizing the submarket for type-$\tilde{1}$ workers, captured through the presence of $\tilde{\theta}_1$ in equation (47). Intuitively, the planner realizes that increasing $\tilde{\theta}_0$ has the additional benefit of moving workers from state $\tilde{0}$ into (the more productive) state $\tilde{1}$ at a faster rate. This reasoning also explains why the optimal θ_0 (equation 48) involves θ_1 and $\tilde{\theta}_0$. The term θ_1 appears in (48) because, like before, the planner realizes that an increase in θ_0 has the additional benefit of switching workers into state 1 at a faster rate. As for $\tilde{\theta}_0$, it appears in (48) because

[20] The term " appropriately chosen" is crucial. The planner's problem here is different from the problem of the planner of Section 6.5, since in each of these sections the model contains a different matching process, and that process must be respected by the planner.

the planner realizes that an increase in θ_0 will move workers out of state 0 more quickly, thus, reducing their exposure to the skill loss shock. Notice that the discussion so far is centred around the *planner's* choices, but, from part (b) of the proposition, we know that the competitive search *equilibrium* implements the exact same allocation.

The more surprising result of this section is illustrated in the first row of Table 6. Since the pioneering work of Moen (1997) and Shimer (1996), competitive search equilibrium has been intertwined with the (constrained) efficient allocation. Our model delivers an interesting finding, which, to the best of our knowledge has not been pointed out in the literature until now. In a model with one-sided heterogeneity, the specification that assumes competitive search and wage posting (by the homogeneous side of the market, in our case, the firms) can deliver *lower* welfare than the Petrongolo and Pissarides (2001) specification, with an appropriately chosen "ranking" of the matching process. Since the focus of our paper is mainly on adopting the Petrongolo and Pissarides (2001) specification and examining quantitatively how different government interventions can improve welfare, we consider this a side result of the paper, and we do not delve into a deeper analysis. However, we think it is a result that should be of interest to researchers who work with search models, since it confounds the standard expectation that the competitive search model should attain maximum welfare.

The most important result of this section is illustrated in the middle rows of Table 6. Under our parameterization, the competitive search specification predicts that the job finding rate of entrant (type-0) workers is 0.83, the highest by far among all worker types. The question that naturally arises is "why does the competitive search model deliver this counterfactual result?", i.e., why do type-0 workers find jobs so fast under this model specification? After all, type-0 workers are less productive than type-1 workers, and one would expect that this would discourage firms from entering the type-0 submarket.

A more careful inspection of the problem (see also Appendix D.1) reveals that the result in question is not all that paradoxical. Firms who contemplate entering submarket 0 understand that if they do so, they will help type-0 workers find a job faster. By doing so they help type-0 workers escape the skill loss scar, which, in turn, generates an additional match surplus. The key observation is that, under competitive search, firms are able to grasp their fair share of that surplus, which basically means that the lower *current* productivity of type-0 workers does *not deter* firms from entering that specific submarket. We find this subtle result particularly interesting, because what drives firm entry into the type-0 submarket is their ability to grasp part of the *future* earnings of workers who are currently less productive (type-0), but are aided in leaving that state faster by the high entry of firms in submarket 0. In other words, the high firm entry in submarket 0, and

Variable	Matching with Ranking	Competitive Search	Intervention 4 (Type-0 Bias)
Y	0.9315	0.9418	0.9486
u	5.8%	7.8%	7.2%
f_1	0.80	0.47	0.51
$\tilde{f}_1$	0.48	0.44	0.34
f_0	0.43	0.83	2.41
$\tilde{f}_0$	0.42	0.42	0.33
w_1	0.9904	0.9846	0.9785
$\tilde{w}_1$	0.9201	0.9191	0.9091
w_0	0.4488	0.8330	0.9111
$\tilde{w}_0$	0.8192	0.7969	0.8067

Table 6: The Implications of Alternative Models.

the willingness of type-0 workers to share their (greater, due to that high entry) future earnings, manifested through a low w_0, work together in general equilibrium.

Although we find this result intriguing from a theoretical perspective, we should repeat that it is counterfactual. In the competitive search equilibrium firms are eager to enter submarket 0 because they are able to secure a healthy fraction of the (future) benefits generated by the training they provide to type-0 workers. However, as we explain in Section 2.1, there is an extensive empirical literature demonstrating that firms do not recoup the benefits of their training investments, and, as a result, they provide suboptimal levels of training. (See Pallais 2014 for a survey of this evidence.) Consequently, the competitive search model also predicts that the job-finding rate of type-0 workers will be higher than that of types 1 and $\tilde{1}$, which directly contradicts the data presented in Figure 1.

In Appendix D.2, we explore another possible model specification used in the literature to study environments with one-sided heterogeneity, namely, a model of random search, bargaining, and segmented markets.[21] The most important takeaway of this analysis is that this model specification delivers the same counterfactual prediction about type-0 worker job-finding rates as the competitive search specification, thus making it (also) an inadequate choice for our research question. The intuition for this result is analogous to the one provided earlier. When firms consider entering submarket 0, they realize that their entry will help type-0 workers avoid the skill loss scar, thus generating an additional match *surplus*. The splitting of this surplus is not anymore implemented through

[21] To be more precise, firms are free to enter any one of four submarkets, one for each type of unemployed worker. Within each submarket, there is a standard Cobb-Douglas matching function, $m(u_i, v_i) = u_i^a v_i^{1-a}$, and the various v_i terms are determined by free entry. In any given match within any submarket, the wage is determined through Nash bargaining between the firm and the unemployed worker. For examples of models using this structure, see Mortensen and Pissarides (1999) and Krause and Lubik (2006).

wage posting (as in the case of competitive search); it is implemented through Nash bargaining, but that turns out to be irrelevant. As long as firms have some bargaining power, they will be able to grasp a share of that future surplus, and the lower *current* productivity of type-0 workers will not deter them from entering submarket 0. Simply put, the economic force driving firms to enter submarket 0 is virtually identical to the one operating under the competitive search specification (even though the surplus-splitting protocol differs), and this leads to the same counterfactual prediction.

Why does the Petrongolo and Pissarides (2001) matching with ranking succeed in explaining the stylized fact of Figure 1 when the other two candidates fail to do so? The Petrongolo and Pissarides (2001) specification allows us to model firms' ranking of worker types according to their *current* productivity. In contrast, the alternative matching specifications determine firms' ranking of worker types according to the workers' *future* career prospects. Specifically, the two alternative specifications allow firms to grasp an unrealistically large fraction of the future surplus generated because of the higher present job-finding rate of entrant workers. In a sense, firms under the Petrongolo and Pissarides (2001) specification behave somewhat myopically because they do not fully realize the future potential surplus to be gained if they match with type-0 workers faster. Interestingly, it turns out that in environments where the switching of productivity types is endogenous, this type of "myopia" can improve the model's ability to fit the data.

8 Conclusion

We develop a DMP model where entrant workers must be trained by firms, and if they stay unemployed for a prolonged period of time they are subject to a permanent skill loss, which we dub the "scar". Depending on their history (how much time they spent in unemployment during their youth and whether they previously held a job), unemployed workers attain one of four types that differ in productivity, and homogeneous firms search for these different worker types. Given this environment, we have a number of ways to model the matching process. We explore three possibilities, but we choose to adopt the Petrongolo and Pissarides (2001) "matching with ranking" process, where firms exhibit bias against entrant/inexperienced workers. We favor the adoption of this model specification, as it is the one that best fits the stylized fact that motivates this paper: entrant workers have lower job-finding rates and longer unemployment spells. In the process of justifying this modeling choice, we show that a specification that assumes competitive search fails to capture this fact.

In our model, firms who hire entrant workers provide a *public good* by reducing these workers' unemployment spells, thus mitigating their exposure to the skill loss shock and reducing the mass of "scarred" workers in the unemployment pool (and the probability that other firms will bump into them in the future). However, firms cannot fully internalize this societal contribution, and ultimately choose to discriminate against entrant workers, causing a social welfare loss. Given this market failure, there is obvious scope for government intervention, aimed to alleviating bias against entrant workers. In a calibrated version of the model, we quantify the effectiveness of three government interventions: "unbiased matching", "government subsidies", and "internships". We find that all three interventions improve aggregate welfare, even though in all of them the aggregate unemployment rate is higher than in the benchmark economy with ranking/bias. The key behind this result is that all government interventions effectively induce firms to incur larger training expenses, thus discouraging entry and increasing unemployment. Despite this unintended consequence on aggregate unemployment, entrant workers have shorter unemployment spells and, as a result, are less likely to suffer skill loss. The productivity gains from the latter channel are so large that aggregate welfare ultimately increases.

Taking this reasoning to its extreme, we consider a fourth intervention in which the government subsidizes the hiring of entrant workers so heavily that firms actually rank these workers higher than the experienced ones. This intervention delivers a substantial improvement in aggregate welfare (almost three times larger than the other three interventions) because it directly confronts the problem of inexperienced workers spending a lot of time in unemployment upon entry. Hence, our analysis predicts that programs that promote the training and job-placement of young workers, such as the Youth Employment Initiative in the EU, generate sizeable welfare gains. Finally, we examine how close this fourth intervention can bring the economy to its efficient level. To that end, we characterize the problem of a social planner who chooses the measure of vacancies, but is subject to the Petrongolo and Pissarides (2001) matching process, with bias in favor of entrant workers. The exercise reveals that the fourth intervention brings the economy arbitrarily close to the constrained efficient outcome. Intuitively, the social planner implements a Hosios-type condition for our environment; however, our analysis illustrates that implementing the "correct" ranking (i.e., giving priority to entrant workers) is an order of magnitude more important than fine-tuning the level of vacancy creation.

References

Acemoglu, D. (1997). Training and innovation in an imperfect labour market. *The Review of Economic Studies 64*(3), 445–464.

Acemoglu, D. and J.-S. Pischke (1998). Why do firms train? theory and evidence. *The Quarterly journal of economics 113*(1), 79–119.

Acemoglu, D. and J.-S. Pischke (1999). Beyond becker: Training in imperfect labour markets. *The economic journal 109*(453), 112–142.

Arellano-Bover, J. (2022). The effect of labor market conditions at entry on workers' long-term skills. *Review of Economics and Statistics 104*(5), 1028–1045.

Autor, D. H. (2001). Why do temporary help firms provide free general skills training? *The Quarterly Journal of Economics 116*(4), 1409–1448.

Becker, G. S. (2009). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago press.

Bertheau, A., B. Larsen, and Z. Zhao (2023). What makes hiring difficult? evidence from linked survey-administrative data. Technical report.

Blanchard, O. J. and P. Diamond (1994). Ranking, unemployment duration, and wages. *The Review of Economic Studies 61*(3), 417–434.

Cohen, J. P., A. C. Johnston, and A. S. Lindner (2023). Skill depreciation during unemployment: Evidence from panel data. Technical report, National Bureau of Economic Research.

Coles, M. and A. Masters (2000). Retraining and long-term unemployment in a model of unlearning by not doing. *European Economic Review 44*(9), 1801–1822.

Dinerstein, M., R. Megalokonomou, and C. Yannelis (2022). Human capital depreciation and returns to experience. *American Economic Review 112*(11), 3725–3762.

Faccini, R. and E. Yashiv (2022). The importance of hiring frictions in business cycles. *Quantitative Economics 13*(3), 1101–1143.

Flemming, J. (2020). Skill accumulation in the market and at home. *Journal of Economic Theory 189*, 105099.

Hall, R. E. and P. R. Milgrom (2008). The limited influence of unemployment on the wage bargain. *American economic review 98*(4), 1653–74.

Herkenhoff, K., J. Lise, G. Menzio, and G. M. Phillips (2024). Production and learning in teams. *Econometrica 92*(2), 467–504.

Hosios, A. J. (1990). On the efficiency of matching and related models of search and unemployment. *The Review of Economic Studies 57*(2), 279–298.

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

Kospentaris, I. (2021). Unobserved heterogeneity and skill loss in a structural model of duration dependence. *Review of Economic Dynamics 39*, 280–303.

Krause, M. U. and T. A. Lubik (2006). The cyclical upgrading of labor and on-the-job search. *Labour Economics 13*(4), 459–477.

Lentz, R. and N. Roys (2024). Training and search on the job. *Review of Economic Dynamics*.

Ljungqvist, L. and T. J. Sargent (1998). The european unemployment dilemma. *Journal of political Economy 106*(3), 514–550.

Loewenstein, M. A. and J. R. Spletzer (1998). Dividing the costs and returns to general training. *Journal of labor Economics 16*(1), 142–171.

Lucas Jr, R. E. (1989). The effects of monetary shocks when prices are set in advance. Technical report, University of Chicago.

Ma, X., A. Nakab, and D. Vidart (2022). Human capital investment and development: The role of on-the-job training. Technical report.

Masui, M. (2023). Provision of firm-sponsored training to temporary workers and labor market performance. *Macroeconomic Dynamics 27*(4), 966–997.

Michaillat, P. and E. Saez (2021). Beveridgean unemployment gap. *Journal of Public Economics Plus 2*, 100009.

Moen, E. R. (1997). Competitive search equilibrium. *Journal of political Economy 105*(2), 385–411.

Moen, E. R. and Å. Rosén (2004). Does poaching distort training? *The Review of Economic Studies 71*(4), 1143–1162.

Montgomery, J. D. (1991). Equilibrium wage dispersion and interindustry wage differentials. *The Quarterly Journal of Economics 106*(1), 163–179.

Mortensen, D. T. and C. A. Pissarides (1999). Unemployment responses to 'skill-biased'technology shocks: the role of labour market policy. *The Economic Journal 109*(455), 242–265.

Ortego-Marti, V. (2016). Unemployment history and frictional wage dispersion. *Journal of Monetary Economics 78*, 5–22.

Ortego-Marti, V. (2017). Loss of skill during unemployment and tfp differences across countries. *European Economic Review*.

Pallais, A. (2014). Inefficient hiring in entry-level labor markets. *American Economic Review 104*(11), 3565–3599.

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

Pissarides, C. A. (1992). Loss of skill during unemployment and the persistence of employment shocks. *The Quarterly Journal of Economics 107*(4), 1371–1391.

Pissarides, C. A. (2000). *Equilibrium unemployment theory*. MIT press.

Rogerson, R., R. Shimer, and R. Wright (2005). Search-theoretic models of the labor market: A survey. *Journal of economic literature 43*(4), 959–988.

Shimer, R. (1996). *Essays in search theory*. Ph. D. thesis, Massachusetts Institute of Technology.

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

Von Wachter, T. (2020). The persistent effects of initial labor market conditions for young adults and their sources. *Journal of Economic Perspectives 34*(4), 168–194.

A Appendix: Proofs of Results

Proof. Derivation of the job-finding rates for each type of workers.
Using the formula for a CRS Cobb-Douglas matching function and the usual notations for the variables, the job-finding rate for a worker of type 1 is given by:

$$f_1 = \frac{m(u_1, v)}{u_1} = \frac{u_1^{\alpha} v^{1-\alpha}}{u_1^{\alpha} u_1^{1-\alpha}} = {b_1}^{\alpha-1}.$$

Now, for the type-$\tilde{1}$ workers, one can use the Petrongolo and Pissarides (2001) logic that type-1 workers "move first", and, only then, type-$\tilde{1}$ workers get a chance to match. The authors of that paper capture this idea by subtracting the matches of only type-1 workers (the "first movers") from the total matches of type-1 and type-$\tilde{1}$ workers, i.e.,

$$\begin{aligned}
\tilde{f_1} &= \frac{m(u_1 + \tilde{u_1}, v) - m(u_1, v)}{\tilde{u_1}} && \text{(A.1)}\\
&= \frac{(u_1 + \tilde{u_1})^{\alpha} v^{1-\alpha}}{\tilde{u_1}} - \frac{u_1^{\alpha} v^{1-\alpha}}{\tilde{u_1}} && \text{(A.2)}\\
&= \left(\frac{u_1 + \tilde{u_1}}{v}\right)^{\alpha} \cdot \frac{v}{\tilde{u_1}} - \left(\frac{u_1}{v}\right)^{\alpha} \cdot \frac{v}{\tilde{u_1}} && \text{(A.3)}\\
&= \frac{(b_1 + \tilde{b}_1)^{\alpha} - b_1^{\alpha}}{\tilde{b}_1}.
\end{aligned}$$

The derivations for the job-finding rates of type-0 and type-$\tilde{0}$ workers follow similar steps, hence, they are omitted. But the logic of matching with ranking based on Petrongolo and Pissarides (2001) is exactly the same, adjusted for a richer set of types.

Based on the Petrongolo and Pissarides (2001) method, and following similar steps, we can also derive the various worker-finding rates for firms. (These are the q expressions that appear in the job creation curves.) Specifically, we have:

$$\begin{aligned}
q_1 &= b_1^{\alpha}, && \text{(A.4)}\\
\tilde{q_1} &= (b_1 + \tilde{b_1})^{\alpha} - b_1^{\alpha}, && \text{(A.5)}\\
q_0 &= (b_1 + \tilde{b_1} + b_0)^{\alpha} - (b_1 + \tilde{b_1})^{\alpha}, && \text{(A.6)}\\
\tilde{q_0} &= (b_1 + b_0 + \tilde{b_1} + \tilde{b_0})^{\alpha} - (b_1 + b_0 + \tilde{b_1})^{\alpha}.
\end{aligned}$$

□

Proof. Proof of Lemma 1. Since the various J_i terms are all equal to each other, and since the bargaining problem in each type of meeting will imply $(1-\eta)(W_i - U_i) = \eta J_i$, we must have that the various $W_i - U_i$ terms are also equal to each other. Let us subtract U_1, given by equation (27), from W_1, given by equation (13); this will give us

$$W_1 - U_1 = \frac{w_1 - z}{r + \lambda + \delta + f}. \tag{A.7}$$

Similarly, let us subtract $\tilde{U_1}$, given by equation (28), from $\tilde{W_1}$, given by equation (14); this will give us

$$\tilde{W_1} - \tilde{U_1} = \frac{\tilde{w_1} - z}{r + \lambda + \delta + f}. \tag{A.8}$$

Direct comparison of equations (A.7) and (A.8), implies that $w_1 = \tilde{w_1}$.

Next, subtract $\tilde{U_0}$, given by equation (26), from $\tilde{W_0}$, given by equation (12), to obtain

$$(r + \delta + f)(\tilde{W_0} - \tilde{U_0}) = \tilde{w_0} - z + \lambda(\tilde{U_1} - \tilde{W_0}). \tag{A.9}$$

To obtain a useful expression for the term $\tilde{U_1} - \tilde{W_0}$ in the last equation, subtract $\tilde{W_0}$, given by equation (12), from $\tilde{U_1}$, given by equation (28). This will give us

$$(r + \delta + \lambda)(\tilde{U_1} - \tilde{W_0}) = z - \tilde{w_0} + f(\tilde{W_1} - \tilde{U_1}),$$

which we can now use to replace the term $\tilde{U_1} - \tilde{W_0}$ in equation (A.9). After this substitution, one should find that

$$(r + \delta + f)(\tilde{W_0} - \tilde{U_0}) = (\tilde{w_0} - z)\left(1 - \frac{\lambda}{r + \delta + \lambda}\right) + \frac{\lambda f}{r + \delta + \lambda}(\tilde{W_1} - \tilde{U_1}).$$

But since we have already established that $(\tilde{W_0} - \tilde{U_0}) = (\tilde{W_1} - \tilde{U_1})$, the last equation can be re-written as

$$\left(r + \delta + f - \frac{\lambda f}{r + \delta + \lambda}\right)(\tilde{W_0} - \tilde{U_0}) = \tilde{w_0} - z + \frac{r + \delta}{r + \delta + \lambda},$$

which, after some algebra, simplifies to

$$\tilde{W_0} - \tilde{U_0} = \frac{\tilde{w_0} - z}{r + \delta + \lambda + f}. \tag{A.10}$$

Direct comparison of equations (A.7), (A.8), and (A.10), implies that $w_1 = \tilde{w_1} = \tilde{w_0}$.

The last thing is to show that the remaining wage w_0 is also equal to the wages of the other three types. To that end, start by subtracting subtract U_0, given by equation (25), from W_0, given by equation (11), to obtain

$$(r+\delta+f)(W_0-U_0)=w_0-z+\lambda(U_1-W_0)+\gamma(U_0-\tilde{U}_0). \tag{A.11}$$

Again, we can find more useful expressions for the terms U_1-W_0 and $U_0-\tilde{U}_0$ in the last equation. Subtracting W_0, given by equation (11), from U_1, given by equation (27), we get

$$(r+\delta+\lambda)(U_1-W_0)=z-w_0+f(W_1-U_1).$$

As for the term $U_0-\tilde{U}_0$, it is easy to show that

$$(r+\delta+\gamma)(U_0-\tilde{U}_0)=0 \implies U_0=\tilde{U}_0.$$

Using these two facts back into equation (A.11) allows to write

$$(r+\delta+f)(W_0-U_0)=w_0-z+\frac{\lambda}{r+\delta+\lambda}[z-w_0+f(W_1-U_1)].$$

But since we have established that $(W_0-U_0)=(W_1-U_1)$, the last equation can be re-written as

$$\left(r+\delta+f-\frac{\lambda f}{r+\delta+\lambda}\right)(W_0-U_0)=(w_0-z)\cdot\frac{r+\delta}{r+\delta+\lambda},$$

which, after some algebra, simplifies to

$$W_0-U_0=\frac{w_0-z}{r+\delta+\lambda+f}. \tag{A.12}$$

Comparison of equations (A.7), (A.8), (A.10), and (A.12) implies that $w_1=\tilde{w_1}=\tilde{w_0}=w_0$, which concludes the proof. □

B Appendix: The details of Intervention 4

This appendix contains the details of the analysis under "Intervention 4", i.e., a government subsidy *only* to type-0 workers, which is generous enough to make them the preferred types for firms.

Arrival rates and Beveridge curves

As we discussed in the main body of the paper, we will still have a Petrongolo and Pissarides (2001) matching process, but the ranking will now favor type-0 workers. (The relative ranking of the remaining types will stay unchanged.) Recalling the definition of the various queue lengths, from equation (1), one can easily show that the new job-finding rates for the four worker types are as follows:

$$
\begin{aligned}
f_0 &= \frac{m(u_0, v)}{u_0} = {b_0}^{\alpha-1}, \\
f_1 &= \frac{m(u_0+u_1, v) - m(u_0, v)}{u_1} = \frac{(b_0+b_1)^\alpha - b_0^\alpha}{b_1}, \\
\tilde{f}_1 &= \frac{m(u_0+u_1+\tilde{u_1}, v) - m(u_0+u_1, v)}{\tilde{u_1}} = \frac{(b_0+b_1+\tilde{b_1})^\alpha - (b_0+b_1)^\alpha}{\tilde{b_1}}, \\
\tilde{f}_0 &= \frac{m(u_0+u_1+\tilde{u_1}+\tilde{u_0}, v) - m(u_0+u_1+\tilde{u_1}, v)}{\tilde{u_0}} = \frac{(b_0+b_1+\tilde{b}_1+\tilde{b}_0)^\alpha - (b_0+b_1+\tilde{b}_1)^\alpha}{\tilde{b}_0}.
\end{aligned}
$$

The formulas admit an interpretation similar to the one provided in Section 2. Although the job-finding rates have changed to reflect the new ranking of workers, the formulas describing the Beveridge curves in Section 3.1 are still valid.

Next, following similar steps, we can derive the various worker-finding rates for firms. These are given by:

$$
\begin{aligned}
q_0 &= b_0^\alpha, \\
q_1 &= (b_0+b_1)^\alpha - b_0^\alpha, \\
\tilde{q_1} &= (b_0+b_1+\tilde{b_1})^\alpha - (b_0+b_1)^\alpha, \\
\tilde{q_0} &= (b_0+b_1+\tilde{b_1}+\tilde{b_0})^\alpha - (b_0+b_1+\tilde{b_1})^\alpha.
\end{aligned}
$$

Value functions and free entry

Since in equilibrium we must have $V = 0$, we can state the free entry condition as

$$c = q_0 J_0 + \tilde{q_0}\tilde{J}_0 + q_1 J_1 + \tilde{q_1}\tilde{J}_1, \tag{A.13}$$

where the value functions for productive firms in the various states are as follows:

$$
\begin{aligned}
rJ_0 &= p + \kappa\rho - \tau - w_0 - \lambda J_0 - \delta J_0, \\
r\tilde{J}_0 &= p - \kappa - \tilde{\kappa} - \tau - \tilde{w}_0 - \lambda\tilde{J}_0 - \delta\tilde{J}_0, \\
rJ_1 &= p - \tau - w_1 - \lambda J_1 - \delta J_1, \\
r\tilde{J}_1 &= p - \tilde{\kappa} - \tau - \tilde{w}_1 - \lambda\tilde{J}_1 - \delta\tilde{J}_1.
\end{aligned}
$$

Clearly, the difference of these value functions with the ones provided in Section 3 is that firms who hire type-0 workers are better off, due to the government subsidy (equal to

$\sigma_0 = \kappa(1+\rho)$), and that all active firms pay the flat tax τ.[22]

The value functions of unemployed workers in the various states are still given by equations (7)-(10), even though the respective job-finding rates have now changed. Similarly, the value functions of employed workers in the various states are still described by equations (11)-(14).

Bargaining and wage curves

Deriving the various wage curves follows identical steps to the analysis in Section 3.3. We will therefore skip the details and proceed to the final formulas. The careful reader will observe that the following wage curves look identical to the ones provided in Section 3.3, after adjusting the formulas to reflect the new productivities.[23] More precisely, we have:

$$w_1 = \frac{\eta(p-\tau)(r+\delta+\lambda+f_1) + (1-\eta)z(r+\delta+\lambda)}{r+\delta+\lambda+\eta f_1},$$

$$\tilde{w_1} = \frac{\eta(p-\tilde{\kappa}-\tau)(r+\delta+\lambda+\tilde{f}_1) + (1-\eta)z(r+\delta+\lambda)}{r+\delta+\lambda+\eta\tilde{f}_1},$$

$$\tilde{w_0} = \frac{1}{r+\delta+\eta\tilde{f}_0}\left[\eta(p-\kappa-\tilde{\kappa}-\tau)(r+\delta+\tilde{f}_0) + \frac{(r+\delta+\lambda)(r+\delta+\tilde{f}_1)}{r+\delta+\lambda+\tilde{f}_1}(1-\eta)z - \frac{\lambda(1-\eta)\tilde{f}_1}{r+\delta+\lambda+\tilde{f}_1}\tilde{w_1}\right].$$

$$\begin{aligned} w_0 = {} & \frac{r+\delta+\gamma+f_0}{r+\delta+\gamma+\eta f_0}\eta(p+\kappa\rho-\tau) + \frac{(r+\delta+\lambda)(r+\delta+f_1)(r+\delta+\gamma)}{(r+\delta)(r+\delta+\gamma+\eta f_0)(r+\delta+\lambda+f_1)}(1-\eta)z \\ & + \frac{\gamma\tilde{f}_0\eta(p-\kappa-\tilde{\kappa}-\tau-\tilde{w_0})}{(r+\delta)(r+\delta+\gamma+\eta f_0)} - \frac{(1-\eta)\lambda f_1(r+\delta+\gamma)w_1}{(r+\delta)(r+\delta+\gamma+\eta f_0)(r+\delta+\lambda+f_1)}. \end{aligned}$$

Definition 5. A steady state equilibrium under Intervention 4 is a list of wages for the four types of workers $(w_0, \tilde{w_0}, w_1, \tilde{w_1})$, a measure of vacant firms v, and measures of unemployed and employed workers in the various states $(u_0, \tilde{u_0}, u_1, \tilde{u_1}, e_0, \tilde{e_0}, e_1, \tilde{e_1})$, satisfying the free entry condition (A.13), the four wage curves described in this section, and the eight Beveridge curves reported at the end of Section 3.1. Implicit in this definition is the flat tax paid by all active firms, given by $\tau = \kappa(1+\rho)e_0/e$.

[22] Which, in equilibrium, will be equal to $\tau = \kappa(1+\rho)e_0/e$, in order to keep the government budget constraint balanced.

[23] Namely, the fact that all firms must now pay a tax, and that firms that match with type-0 workers receive a subsidy which makes the pre-tax effective productivity of these workers equal to $p+\kappa\rho$.

C Appendix: The social planner's problem under matching with ranking and type-0 bias

The Hamiltonian function for the planner's problem is the following:

$$\begin{aligned}
\mathcal{H} = & e^{-rt}\big(e_1 p + e_0(p-\kappa) + \tilde{e}_1(p-\tilde{\kappa}) + \tilde{e}_0(p-\tilde{\kappa}-\kappa) + (u_1+u_0+\tilde{u}_1+\tilde{u}_0)z - cv\big) \\
& + \mu_{u_0}\big(\delta - u_0(\delta+\gamma+f_0)\big) + \mu_{\tilde{u}_0}\big(\gamma u_0 - \tilde{u}_0(\delta+\tilde{f}_0)\big) + \mu_{u_1}\big(\lambda(e_0+e_1) - u_1(\delta+f_1)\big) \\
& + \mu_{\tilde{u}_1}\big(\lambda(\tilde{e}_0+\tilde{e}_1) - \tilde{u}_1(\tilde{f}_1+\delta)\big) + \mu_{e_0}\big(u_0 f_0 - e_0(\delta+\lambda)\big) + \mu_{\tilde{e}_0}\big(\tilde{f}_0\tilde{u}_0 - (\delta+\lambda)\tilde{e}_0\big) \\
& + \mu_{e_1}\big(f_1 u_1 - (\lambda+\delta)e_1\big) + \mu_{\tilde{e}_1}\big(\tilde{f}_1\tilde{u}_1 - \tilde{e}_1(\lambda+\delta)\big).
\end{aligned}$$

It will be useful to define some objects to make the notation simpler.

$g_1 \equiv b_0^{a-1}$,

$g_2 \equiv (b_0+b_1)^{a-1}$,

$g_3 \equiv (b_0+b_1+b_2)^{a-1}$,

$g_4 \equiv (b_0+b_1+b_2+b_3)^{a-1}$,

where $b_0 = \frac{u_0}{v}, b_1 = \frac{u_1}{v}, b_2 = \frac{\tilde{u}_1}{v}$, and, $b_3 = \frac{\tilde{u}_0}{v}$.

We get the first order-conditions from the above Hamiltonian:

$$\begin{aligned}
\{v\} : & f_0 u_0(\mu_{e_0}-\mu_{u_0}) + \tilde{f}_0\tilde{u}_0(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_0}) + f_1 u_1(\mu_{e_1}-\mu_{u_1}) + \tilde{f}_1\tilde{u}_1(\mu_{\tilde{e}_1}-\mu_{\tilde{u}_1}) \\
& = \frac{vc}{1-a}e^{-rt},
\end{aligned} \tag{A.14}$$

$$\begin{aligned}
\{u_0\} : & ze^{-rt} + a(\mu_{e_0}-\mu_{u_0})f_0 + a(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_0})(g_4-g_3) + a(\mu_{e_1}-\mu_{u_1})(g_2-g_1) \\
& + a(\mu_{\tilde{e}_1}-\mu_{\tilde{u}_1})(g_3-g_2) + \gamma\mu_{\tilde{u}_0} - (\delta+\gamma)\mu_0 = -\dot{\mu}_{u_0},
\end{aligned} \tag{A.15}$$

$$\begin{aligned}
\{u_1\} : & ze^{-rt} + a(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_0})(g_4-g_3) + a(\mu_{\tilde{e}_1}-\mu_{\tilde{u}_1})(g_3-g_2) + a(\mu_{e_1}-\mu_{u_1})g_2 \\
& - \delta\mu_{u_1} = -\dot{\mu}_{u_1},
\end{aligned} \tag{A.16}$$

$$\{\tilde{u}_1\} : ze^{-rt} + a(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_0})(g_4-g_3) + a(\mu_{\tilde{e}_1}-\mu_{\tilde{u}_1})g_3 - \delta\mu_{\tilde{u}_1} = -\dot{\mu}_{\tilde{u}_1}, \tag{A.17}$$

$$\{\tilde{u}_0\} : ze^{-rt} + a(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_0})g_4 - \delta\mu_{\tilde{u}_0} = -\dot{\mu}_{\tilde{u}_0}, \tag{A.18}$$

$$\{e_0\} : (p-\kappa)e^{-rt} + \mu_{u_1} - \mu_{e_0}(\lambda+\delta) = -\dot{\mu}_{e_0}, \tag{A.19}$$

$$\{e_1\} : pe^{-rt} + \lambda\mu_{u_1} - \mu_{e_1}(\lambda+\delta) = -\dot{\mu}_{e_1}, \tag{A.20}$$

$$\{\tilde{e}_1\} : (p-\tilde{\kappa})e^{-rt} + \mu_{\tilde{u}_1}\lambda - \mu_{\tilde{e}_1}(\lambda+\delta) = -\dot{\mu}_{\tilde{e}_1}, \tag{A.21}$$

$$\{\tilde{e}_0\} : (p-\kappa-\tilde{\kappa})e^{-rt} + \mu_{\tilde{u}_1}\lambda - \mu_{\tilde{e}_0}(\lambda+\delta) = -\dot{\mu}_{\tilde{e}_0}. \tag{A.22}$$

Subtracting (A.14) from (A.21) gives

$$e^{-rt}(p-\tilde{\kappa}-z) - a(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_0})(g_4-g_3) = (r+\delta+\lambda+ag_3)(\mu_{\tilde{e}_1}-\mu_{\tilde{u}_1}). \tag{A.23}$$

Similarly, subtracting (A.18) from (A.22) yields

$$e^{-rt}(p-\kappa-\tilde{\kappa}-z) + \lambda(\mu_{\tilde{u}_1}-\mu_{\tilde{e}_0}) = (r+\delta+ag_4)(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_0}). \tag{A.24}$$

To get an expression for $(\mu_{\tilde{u}_1}-\mu_{\tilde{e}_0})$, subtract (A.17) from (A.22) to obtain

$$\begin{aligned}(r+\delta+\lambda)(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_1}) =& e^{-rt}(p-\kappa-\tilde{\kappa}-z) - a(\mu_{\tilde{e}_1}-\mu_{\tilde{u}_1})g_3 \\ & - a(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_0})(g_4-g_3).\end{aligned} \tag{A.25}$$

Next, using (A.25) in (A.24), we get

$$\begin{aligned}(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_0}) = & \left(r+\delta+ag_4\frac{r+\delta}{r+\delta+\lambda} + \frac{a\lambda}{r+\delta+\lambda}g_3\right)^{-1} \\ & \left(e^{-rt}(p-\kappa-\tilde{\kappa}-z)\frac{r+\delta}{r+\delta+\lambda} + \frac{\lambda}{r+\delta+\lambda}ag_3(\mu_{\tilde{e}_1}-\mu_{\tilde{u}_1})\right).\end{aligned} \tag{A.26}$$

Note that equations (A.26) and (A.23) consist of only two unknowns, namely, $(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_0})$ and $(\mu_{\tilde{e}_1}-\mu_{\tilde{u}_1})$. Solving for these expressions, we get

$$\begin{aligned}(\mu_{\tilde{e}_1}-\mu_{\tilde{u}_1}) = & \frac{e^{-rt}}{r+\delta+\lambda+ag_3}\Bigg((p-\tilde{\kappa}-z) \\ & - a(g_4-g_3)\frac{(r+\delta)(p-\kappa-\tilde{\kappa}-z) + \frac{a\lambda g_3}{r+\delta+\lambda+ag_3}(p-\tilde{\kappa}-z)}{(r+\delta+\lambda+ag_4)(r+\delta+\frac{a\lambda g_3}{r+\delta+\lambda+ag_3})}\Bigg)\end{aligned} \tag{A.27}$$

and

$$(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_0}) = e^{-rt}\frac{(r+\delta)(p-\kappa-\tilde{\kappa}-z) + \frac{a\lambda g_3}{r+\delta+\lambda+ag_3}(p-\tilde{\kappa}-z)}{(r+\delta+\lambda+ag_4)(r+\delta+\frac{a\lambda g_3}{r+\delta+\lambda+ag_3})}. \tag{A.28}$$

Next, we subtract (A.16) from (A.20) and, rearranging, we get a formula for $(\mu_{e_1}-\mu_{u_1})$:

$$\begin{aligned}(\mu_{e_1}-\mu_{u_1}) =& \frac{1}{r+\delta+\lambda+ag_2}\big(e^{-rt}(p-z) - a(g_3-g_2)(\mu_{\tilde{e}_1}-\mu_{\tilde{u}_1}) \\ & - a(g_4-g_3)(\mu_{\tilde{e}_0}-\mu_{\tilde{u}_0})\big).\end{aligned} \tag{A.29}$$

Note that above expression is a closed-form solution given equations (A.28) and (A.27), thus there is no reason to plug them in here. Now, we only need to solve for $(\mu_{e_0} - \mu_{u_0})$. Towards that, by subtracting (A.15) from (A.19) we have

$$\begin{aligned}\lambda(\mu_{e_0} - \mu_{u_1}) =& e^{-rt}(p - \kappa - z) + \gamma(\mu_{u_0} - \mu_{\tilde{u}_0}) - (r + \delta + af_0)(\mu_{e_0} - \mu_{u_0}) \\ & - a(g_2 - g_1)(\mu_{e_1} - \mu_{u_1}) - a(g_3 - g_2)(\mu_{\tilde{e}_1} - \mu_{\tilde{u}_1}) \\ & - a(g_4 - g_3)(\mu_{\tilde{e}_0} - \mu_{\tilde{u}_0}).\end{aligned} \tag{A.30}$$

Now, derive $(\mu_{e_0} - \mu_{u_1})$ by subtracting (A.16) from (A.19) and derive $(\mu_{u_0} - \mu_{\tilde{u}_0})$ by subtracting (A.18) from (A.15). Plugging both solutions in (A.30) results in the following expression for $(\mu_{e_0} - \mu_{u_0})$:

$$\begin{aligned}(\mu_{e_0} - \mu_{u_0}) =& \frac{r + \delta + \gamma}{(af_0 + r + \delta + \gamma)(r + \delta)}\Bigg[\frac{r + \delta}{r + \delta + \lambda}(p - \kappa - z)e^{-rt} \\ & + (\mu_{e_1} - \mu_{u_1})\left(\frac{\lambda}{r + \delta + \lambda}ag_2 - a(g_2 - g_1)\frac{r + \delta}{r + \delta + \lambda}\right) \\ & + a(g_3 - g_2)(\mu_{\tilde{e}_1} - \mu_{\tilde{u}_1})\left(\frac{\lambda}{r + \delta + \lambda} - \frac{r + \delta}{r + \delta + \gamma}\right) \\ & + a(g_4 - g_3)(\mu_{\tilde{e}_0} - \mu_{\tilde{u}_0})\left(\frac{\lambda}{r + \delta + \lambda} - \frac{r + \delta}{r + \delta + \gamma}\right) \\ & - \frac{a\gamma}{r + \delta + \gamma}g_4(\mu_{\tilde{e}_0} - \mu_{\tilde{u}_0})\Bigg].\end{aligned} \tag{A.31}$$

Using (A.31), (A.29), (A.28) and (A.27) altogether in (A.14), we get an equation that governs the optimal level of vacancies only including parameters of the model and other equilibrium variables. We do not state it here. Moreover, since we are interested in the steady state solution of planner's problem, the eight Beveridge curves in the end of section 3.1 must hold as well. Thereby, we end up with nine equations in nine variables which uniquely solves for the planner's problem.

D Appendix: Discussion of alternative matching models

D.1 Proof of Proposition 2.

We prove parts (a) and (b) jointly. Specifically, we first prove part (b), i.e., we show that the social planner's allocation coincides with equations (45)-(48) reported in part (a). We do so because the social planner's problem is significantly easier. Once done with this step, we will describe the competitive search equilibrium allocation and prove step by step that it coincides with the allocation chosen by the planner for all four submarkets.

Step 1: The social planner's problem

First we present and solve the problem of a social planner who takes as given the matching process within the submarkets and chooses the optimal firm entry for each one of them.[24] Using the definition of market tightness in each submarket, i.e., $\theta_i \equiv v_i/u_i$, allows us to write the planner's problem as one where the planner chooses θ's, as opposed to v's.[25] Then the planner's problem is to:

$$\max_{\theta_0,\tilde{\theta_0},\theta_1,\tilde{\theta_1}} \Omega \equiv \int_0^\infty e^{-rt}[e_1 p + e_0(p-\kappa) + \tilde{e_1}(p-\tilde{\kappa}) + \tilde{e_0}(p-\kappa-\tilde{\kappa}) \\ + (u_1 + u_0 + \tilde{u_1} + \tilde{u_0})z - c(\theta_1 u_1 + \theta_0 u_0 + \tilde{\theta}_1\tilde{u_1} + \tilde{\theta}_0\tilde{u_0})]dt,$$

subject to

$$\begin{aligned}
\dot{u}_0 &= \delta - u_0(\delta + \gamma + f_0), & \dot{\tilde{u_0}} &= \gamma u_0 - (\delta + \tilde{f}_0)\tilde{u_0}, \\
\dot{u}_1 &= \lambda(e_0 + e_1) - (\delta + f_1)u_1, & \dot{\tilde{u_1}} &= \lambda(\tilde{e_0} + \tilde{e_1}) - (\delta + \tilde{f}_1)\tilde{u_1}, \\
\dot{e_0} &= f_0 u_0 - (\delta + \lambda)e_0, & \dot{\tilde{e_0}} &= \tilde{f}_0\tilde{u_0} - (\delta + \lambda)\tilde{e_0}, \\
\dot{e_1} &= f_1 u_1 - (\delta + \lambda)e_1, & \dot{\tilde{e_1}} &= \tilde{f}_1\tilde{u_1} - (\delta + \lambda)\tilde{e_1}.
\end{aligned}$$

Since matching in each submarket is characterized by a CRS Cobb-Douglas matching function, it is understood that

$$f_1 = \left(\frac{m(u_1, v_1)}{u_1} =\right)\theta_1^{1-a}; \quad \tilde{f}_1 = \tilde{\theta}_1^{\,1-a}; \quad f_0 = \theta_0^{1-a}; \quad \tilde{f}_0 = \tilde{\theta}_0^{\,1-a}.$$

Similar to Chapter 8 of Pissarides (2000), we set up the Hamiltonian of this dynamic problem with the various $\dot{\mu}$ terms standing for the co-state variables. Then, we derive the first order conditions by setting the derivatives of the Hamiltonian with respect to the market tightnesses (the "control variables") equal to zero, and the derivatives with respect to the various employment states equal to the negative of the evolution of the respective co-state variables. Following these steps, we get the following 12 equations for

[24] As we discuss in the main text after Proposition 2, this planner solves a problem that is very different than the planner considered in Section 6.5.

[25] As is well-known from Chapter 8 of Pissarides (2000), this is an equivalent but easier problem to solve.

the first order conditions:

$$\frac{\partial \mathcal{H}}{\partial \theta_0} = 0 \iff ce^{-rt} = (1-\alpha){\theta_0}^{-\alpha}(\mu_{e_0} - \mu_{u_0}), \tag{A.32}$$

$$\frac{\partial \mathcal{H}}{\partial \theta_1} = 0 \iff ce^{-rt} = (1-\alpha){\theta_1}^{-\alpha}(\mu_{e_1} - \mu_{u_1}), \tag{A.33}$$

$$\frac{\partial \mathcal{H}}{\partial \tilde{\theta_0}} = 0 \iff ce^{-rt} = (1-\alpha)\tilde{\theta_0}^{-\alpha}(\mu_{\tilde{e_0}} - \mu_{\tilde{u_0}}), \tag{A.34}$$

$$\frac{\partial \mathcal{H}}{\partial \tilde{\theta_1}} = 0 \iff ce^{-rt} = (1-\alpha)\tilde{\theta_1}^{-\alpha}(\mu_{\tilde{e_1}} - \mu_{\tilde{u_1}}), \tag{A.35}$$

$$\frac{\partial \mathcal{H}}{\partial u_0} = -\dot{\mu}_{u_0} \iff e^{-rt}(z - c\theta_0) - \mu_{u_0}(\delta + \gamma + f_0) + \mu_{\tilde{u_0}}\gamma + \mu_{e_0}f_0 = -\dot{\mu_{u_0}}, \tag{A.36}$$

$$\frac{\partial \mathcal{H}}{\partial u_1} = -\dot{\mu}_{u_1} \iff e^{-rt}(z - c\theta_1) - \mu_{u_1}(\delta + f_1) + \mu_{e_1}f_1 = -\dot{\mu_{u_1}}, \tag{A.37}$$

$$\frac{\partial \mathcal{H}}{\partial \tilde{u_0}} = -\dot{\mu}_{\tilde{u_0}} \iff e^{-rt}(z - c\tilde{\theta_0}) - \mu_{\tilde{u_0}}(\delta + \tilde{f_0}) + \mu_{\tilde{e_0}}\tilde{f_0} = -\dot{\mu}_{\tilde{u_0}}, \tag{A.38}$$

$$\frac{\partial \mathcal{H}}{\partial \tilde{u_1}} = -\dot{\mu}_{\tilde{u_1}} \iff e^{-rt}(z - c\tilde{\theta_1}) - \mu_{\tilde{u_1}}(\delta + \tilde{f_1}) + \mu_{\tilde{e_1}}\tilde{f_1} = -\dot{\mu}_{\tilde{u_1}}, \tag{A.39}$$

$$\frac{\partial \mathcal{H}}{\partial e_0} = -\dot{\mu}_{e_0} \iff e^{-rt}(p - \kappa) + \mu_{u_1}\lambda - \mu_{e_0}(\delta + \lambda) = -\dot{\mu}_{e_0}, \tag{A.40}$$

$$\frac{\partial \mathcal{H}}{\partial e_1} = -\dot{\mu}_{e_1} \iff e^{-rt}p + \mu_{u_1}\lambda - \mu_{e_1}(\delta + \lambda) = -\dot{\mu}_{e_1}, \tag{A.41}$$

$$\frac{\partial \mathcal{H}}{\partial \tilde{e_0}} = -\dot{\mu}_{\tilde{e_0}} \iff e^{-rt}(p - \kappa - \tilde{\kappa}) + \mu_{\tilde{u_1}}\lambda - \mu_{\tilde{e_0}}(\delta + \lambda) = -\dot{\mu}_{\tilde{e_0}}, \tag{A.42}$$

$$\frac{\partial \mathcal{H}}{\partial \tilde{e_1}} = -\dot{\mu}_{\tilde{e_1}} \iff e^{-rt}(p - \kappa) + \mu_{\tilde{u_1}}\lambda - \mu_{\tilde{e_1}}(\delta + \lambda) = -\dot{\mu}_{\tilde{e_1}}. \tag{A.43}$$

For the first equation (45), subtracting equation (A.41) from (A.37) gives:

$$\dot{\mu}_{e_1} - \dot{\mu}_{u_1} = (\delta + \lambda + f_1)(\mu_{e_1} - \mu_{u_1}) - e^{-rt}(p - z + c\theta_1). \tag{A.44}$$

From equation (A.33), we have

$$\mu_{e_1} - \mu_{u_1} = \frac{ce^{-rt}}{(1-\alpha){\theta_1}^{-\alpha}} \implies \dot{\mu}_{e_1} - \dot{\mu}_{u_1} = \frac{-cre^{-rt}}{(1-\alpha){\theta_1}^{-\alpha}}.$$

Plugging this to equation (A.44) along with some algebra yields equation (45).

Then, the second equation (46) can be obtained by subtracting equation (A.43) from (A.39) and using similar steps. In this case, we use equation (A.35) to obtain useful expressions for $\mu_{\tilde{e_1}} - \mu_{\tilde{u_1}}$ and $\dot{\mu}_{\tilde{e_1}} - \dot{\mu}_{\tilde{u_1}}$ to get the final equation.

For the third equation (47), the steps are slightly more involved because when we subtract (A.42) from (A.38), we obtain:

$$\dot{\mu}_{\tilde{e_0}} - \dot{\mu}_{\tilde{u_0}} = (\delta + \tilde{f}_0)(\mu_{\tilde{e_0}} - \mu_{\tilde{u_0}}) + \lambda(\mu_{\tilde{e_0}} - \mu_{\tilde{u_1}}) - e^{-rt}(\tilde{p} - z + c\tilde{\theta}_0). \tag{A.45}$$

Here, $\dot{\mu}_{\tilde{e_0}} - \dot{\mu}_{\tilde{u_0}}$ and $\mu_{\tilde{e_0}} - \mu_{\tilde{u_0}}$ can be obtained from equation (A.34) using similar steps as before. For $\mu_{\tilde{e_0}} - \mu_{\tilde{u_1}}$, use equation (A.39) to get:

$$e^{-rt}(z - c\tilde{\theta}_1) + \tilde{f}_1(\mu_{\tilde{e_1}} - \mu_{\tilde{u_1}}) = (r + \delta)\mu_{\tilde{u_1}}. \tag{A.46}$$

Similarly from equation (A.42), we get:

$$e^{-rt}(p - \kappa - \tilde{\kappa}) + \mu_{\tilde{u_1}}\lambda = \mu_{\tilde{e_0}}(r + \delta + \lambda). \tag{A.47}$$

Subtracting equation (A.46) from (A.47) and substituting into (A.45) followed by some algebraic manipulations yields the final version of equation (47) involving $\tilde{\theta}_0$ and $\tilde{\theta}_1$.

The steps to derive the last equation (48) are similar to the steps above for the third equation. We first subtracted equation (A.40) from (A.36). Similar to before, we subtract the expressions obtained from equations (A.37) and (A.40) to get $\mu_{e_0} - \mu_{u_1}$, and we subtract the expressions obtained from equations (A.36) and (A.38) to get $\mu_{\tilde{u_0}} - \mu_{u_0}$. Substituting all the expressions followed by some algebra yields equation (48).

Step 2: The competitive search equilibrium allocation

Having shown that the planner's allocation is characterized by equations (45)-(48), we will now move on to describing the competitive search equilibrium allocation, and we will prove that it coincides with the allocation chosen by the planner for all submarkets. We will start with the easier cases of submarkets 1 and $\tilde{1}$, and then move on to the more involved submarkets $\tilde{0}$ and 0.

Submarket 1

The competitive search problem can be expressed as follows:[26]

[26] Notice that it is more convenient to set up the competitive search problem(s) using the "queue length" term b_1, as opposed to the "market tightness" term θ_1. Of course, that shoud not create any confusion, since we have already established that $b_i = 1/\theta_i$ in any submarket. We will switch between the two terms at will, depending on what is most convenient given the application at hand.

$$
\begin{aligned}
\max_{b_1, w_1} \quad & -c + b_1^a \cdot \frac{p - w_1}{r + \lambda + \delta} \\
\text{s.t.} \quad & U_1(r + \delta + f_1) = z + f_1 \frac{w_1 + \lambda U_1}{r + \lambda + \delta} \\
\iff & (r+\delta)U_1 \left[b_1 + \frac{b_1^a}{r + \lambda + \delta} \right] - b_1 z = \frac{b_1^a w_1}{r + \lambda + \delta}.
\end{aligned}
$$

Using the last version of the constraint, allows us to re-write the objective function as:

$$
\max_{b_1} \quad -c + \frac{b_1^a p}{r + \lambda + \delta} + b_1 z - (r+\delta)U_1 \left[b_1 + \frac{b_1^a}{r + \lambda + \delta} \right]. \tag{A.48}
$$

Taking the first order condition with respect to b_1 yields:

$$
\frac{apb_1^{a-1}}{r + \lambda + \delta} + z - (r+\delta)U_1 \left[b_1 + \frac{b_1^a}{r + \lambda + \delta} \right] = 0
$$

$$
\implies \frac{apb_1^{a-1}}{r + \lambda + \delta} \left[p - (r+\delta)U_1 \right] = (r+\delta)U_1 - z \tag{A.49}
$$

To solve b_1 only as a function of parameters, we need to go back to the objective function and impose $V_1 = 0$. We will impose that the (A.48) expression is equal to 0, but in that expression, we will replace $(r+\delta)U_1$ from (A.49). To that end, notice that from (A.49),

$$
(r+\delta)U_1 \left[1 + \frac{ab_1^{a-1}}{r + \lambda + \delta} \right] = z + \frac{pab_1^{a-1}}{r + \lambda + \delta}
$$

$$
\implies (r+\delta)U_1 = \frac{(r + \lambda + \delta)z + pab_1^{a-1}}{r + \lambda + \delta + ab_1^{a-1}}. \tag{A.50}
$$

We will use this into (A.48) and set equal to 0 to get:

$$
\begin{aligned}
c &= b_1 z + \frac{pb_1^a}{r + \lambda + \delta} - b_1 \left(\frac{r + \lambda + \delta + b_1^{a-1}}{r + \lambda + \delta} \right) \frac{(r + \lambda + \delta)z + pab_1^{a-1}}{r + \lambda + \delta + ab_1^{a-1}} \\
\implies c &= b_1 z \frac{(1-a)b_1^{a-1}}{r + \lambda + \delta + ab_1^{a-1}} + \frac{pb_1^a}{r + \lambda + \delta} \frac{(r + \lambda + \delta)(1-a)}{r + \lambda + \delta + ab_1^{a-1}} \\
\implies c &= \frac{b_1^a (1-a)(p-z)}{r + \lambda + \delta + ab_1^{a-1}}.
\end{aligned}
$$

Using the fact that $\theta_1 = 1/b_1$ and rearranging leads to

$$
(r + \lambda + \delta)\theta_1^a + a\theta_1 = (1-a)\frac{(p-z)}{c},
$$

which is equation (45).

Submarket $\tilde{1}$

The proof for submarket $\tilde{1}$ is identical to the one for submarket 1, so we omit it and move directly to the more difficult problems for submarkets $\tilde{0}$ and 0.

Submarket $\tilde{0}$

Defining $\tilde{p} \equiv p - \kappa - \tilde{\kappa}$ for convenience, we solve the following problem:

$$\max_{\tilde{b}_0, \tilde{w}_0} \quad -c + \tilde{b}_0^a \cdot \frac{\tilde{p} - \tilde{w}_0}{r + \lambda + \delta}$$

$$\text{s.t.} \quad \tilde{U}_0(r + \delta + \tilde{b}_0^{a-1}) = z + \tilde{b}_0^{a-1} \frac{\tilde{w}_0 + \lambda \tilde{U}_1}{r + \lambda + \delta}$$

$$\iff \quad -\tilde{b}_0 \tilde{U}_0(r + \delta + \tilde{b}_0^{a-1}) + \frac{\tilde{b}_0^a \lambda \tilde{U}_1}{r + \lambda + \delta} = -z\tilde{b}_0 - \frac{\tilde{b}_0^a \tilde{w}_0}{r + \lambda + \delta}.$$

This allows us to re-write our problem as:

$$\max_{\tilde{b}_0} \quad -c + \frac{\tilde{b}_0^a \tilde{p}}{r + \lambda + \delta} - \tilde{U}_0(\tilde{b}_0(r + \delta) + \tilde{b}_0^a) + \frac{\tilde{b}_0^a \lambda \tilde{U}_1}{r + \lambda + \delta} + z\tilde{b}_0.$$

Taking the first order condition with respect to $\tilde{b}_0$ yields:

$$\frac{a\tilde{p}\tilde{b}_0^{a-1}}{r + \lambda + \delta} - (r + \delta)\tilde{U}_0 - \tilde{U}_0 a \tilde{b}_0^{a-1} + \frac{a\tilde{b}_0^{a-1} \lambda \tilde{U}_1}{r + \lambda + \delta} + z = 0$$

$$\implies \tilde{U}_0[r + \delta + a\tilde{b}_0^{a-1}] = z + \frac{a\tilde{b}_0^{a-1}}{r + \lambda + \delta}(\tilde{p} + \lambda \tilde{U}_1). \tag{A.51}$$

$\tilde{V}_0 = 0$ implies that:

$$c = \frac{\tilde{p}\tilde{b}_0^a}{r + \lambda + \delta} + z\tilde{b}_0 + \frac{\tilde{b}_0^a \lambda \tilde{U}_1}{r + \lambda + \delta} - \tilde{b}_0 \tilde{U}_0[r + \delta + \tilde{b}_0^{a-1}].$$

Plug this to rewrite the last expression in A.51 as:

$$
\begin{aligned}
c &= \frac{\tilde{p}\tilde{b}_0^a}{r+\lambda+\delta} + z\tilde{b}_0 + \frac{\tilde{b}_0^a \lambda \tilde{U}_1}{r+\lambda+\delta} - \tilde{b}_0 \cdot (r+\delta+a\tilde{b}_0^{a-1}) \cdot \frac{z + \frac{a\tilde{b}_0^{a-1}}{r+\lambda+\delta}(\tilde{p}+\lambda\tilde{U}_1)}{r+\delta+a\tilde{b}_0^{a-1}} \\
&= \frac{\tilde{b}_0^a}{r+\lambda+\delta}(\tilde{p}+\lambda\tilde{U}_1) + z\tilde{b}_0 - \tilde{b}_0 \frac{r+\delta+\tilde{b}_0^{a-1}}{(r+\delta+a\tilde{b}_0^{a-1})} z - \frac{\tilde{b}_0(r+\delta+\tilde{b}_0^{a-1})a\tilde{b}_0^{a-1}(\tilde{p}+\lambda\tilde{U}_1)}{(r+\lambda+\delta)(r+\delta+a\tilde{b}_0^{a-1})} \\
&= z\tilde{b}_0 \left[1 - \frac{r+\delta+\tilde{b}_0^{a-1}}{r+\delta+a\tilde{b}_0^{a-1}}\right] + \frac{\tilde{b}_0^a(\tilde{p}+\lambda\tilde{U}_1)}{r+\lambda+\delta}\left[1 - \frac{(r+\delta+\tilde{b}_0^{a-1})a}{r+\delta+a\tilde{b}_0^{a-1}}\right] \\
&= z\tilde{b}_0 \frac{\tilde{b}_0^{a-1}(a-1)}{r+\delta+a\tilde{b}_0^{a-1}} + \frac{\tilde{b}_0^a(\tilde{p}+\lambda\tilde{U}_1)}{r+\lambda+\delta} \cdot \frac{(r+\delta)(1-a)}{r+\delta+a\tilde{b}_0^{a-1}} \\
&= \frac{\tilde{b}_0^a(1-a)}{r+\delta+a\tilde{b}_0^{a-1}}\left[\frac{(\tilde{p}+\lambda\tilde{U}_1)(r+\delta)}{r+\lambda+\delta} - z\right].
\end{aligned}
$$

Use the fact that $\tilde{\theta}_0 = 1/\tilde{b}_0$ to obtain:

$$
\begin{aligned}
&(r+\delta+a\tilde{\theta}_0^{1-a})(r+\lambda+\delta)\tilde{\theta}_0^a = \frac{1-a}{c}[(\tilde{p}+\lambda\tilde{U}_1)(r+\delta) - (r+\delta+\lambda)z] \\
\implies &(r+\delta+a\tilde{\theta}_0^{1-a})(r+\lambda+\delta)\tilde{\theta}_0^a = \frac{1-a}{c}[(r+\delta)(\tilde{p}-z) + \lambda((r+\delta)\tilde{U}_1 - z)].
\end{aligned}
$$

Finally, divide by $(r+\delta)$ to obtain

$$
\begin{aligned}
&(1 + \frac{a\tilde{\theta}_0^{1-a}}{r+\delta})(r+\delta+\lambda)\tilde{\theta}_0^a = \frac{(1-a)(\tilde{p}-z)}{c} + \frac{\lambda(1-a)}{(r+\delta)c}[(r+\delta)\tilde{U}_1 - z] \\
&\underset{\checkmark}{(r+\delta+\lambda)\tilde{\theta}_0^a} + \underset{\checkmark}{a\tilde{\theta}_0} + \underset{\checkmark}{\frac{a\lambda}{r+\delta}\tilde{\theta}_0} - \frac{\lambda(1-a)}{c(r+\delta)}[(r+\delta)\tilde{U}_1 - z] = \underset{\checkmark}{\frac{(1-a)(p-\kappa-\tilde{\kappa}-z)}{c}}.
\end{aligned}
$$

In the last expression, all the checked terms are identical to equation (47). To conclude the proof, we need to show that the unchecked term is equal to $\frac{\lambda a}{r+\delta}\tilde{\theta}_1$. To prove that we can write an equation similar to (A.49) for submarket $\tilde{1}$, which is isomorphic:

$$
\begin{aligned}
(r+\delta)\tilde{U}_1 &= \frac{(r+\lambda+\delta+a\tilde{b}_1^{a-1})z + a\tilde{b}_1^{a-1}(p-\tilde{\kappa}-z)}{r+\lambda+\delta+a\tilde{b}_1^{a-1}} \\
(r+\delta)\tilde{U}_1 - z &= \frac{a\tilde{b}_1^{a-1}(p-\tilde{\kappa}-z)}{r+\lambda+\delta+a\tilde{b}_1^{a-1}} = \frac{a}{(1-a)\tilde{b}_1} \frac{\tilde{b}_1^a(1-a)(p-\tilde{\kappa}-z)}{r+\lambda+\delta+a\tilde{b}_1^{a-1}} \\
(r+\delta)\tilde{U}_1 - z &= \frac{a}{a-1}\tilde{\theta}_1 \, c.
\end{aligned}
$$

Plug this into the above unchecked term to get:

$$
\frac{\lambda(1-a)}{c(r+\delta)}[(r+\delta)\tilde{U}_1 - z] = \frac{\lambda(1-a)}{c(r+\delta)} \frac{a\tilde{\theta}_1 c}{(1-a)} = \frac{\lambda a}{r+\delta}\tilde{\theta}_1.
$$

This concludes the proof.

Submarket 0

We solve the following problem:

$$\begin{aligned}
&\max_{b_0,w_0} \quad -c + b_0^a \cdot \frac{p-\kappa-w_0}{r+\lambda+\delta} \\
&\text{s.t.} \quad U_0(r+\delta+\gamma+b_0^{a-1}) = z + b_0^{a-1}\frac{w_0+\lambda U_1}{r+\lambda+\delta} + \gamma\tilde{U}_0 \\
&\iff -\frac{b_0^a w_0}{r+\lambda+\delta} = -b_0U_0(r+\delta+\gamma+b_0^{a-1}) + zb_0 + \frac{b_0^a\lambda U_1}{r+\lambda+\delta} + \gamma b_0\tilde{U}_0.
\end{aligned}$$

We can use the last version of the constraint to re-write the problem as follows:

$$\max_{b_0} \quad -c + b_0^a\frac{p-\kappa}{r+\lambda+\delta} - U_0(b_0(r+\delta)+b_0^a) + zb_0 + b_0^a\frac{\lambda U_1}{r+\lambda+\delta} + \gamma b_0\tilde{U}_0.$$

Taking the first order condition with respect to $\tilde{b}_0$ yields:

$$\begin{aligned}
&ab_0^{a-1}\frac{p-\kappa}{r+\lambda+\delta} - (r+\delta+\gamma)U_0 - ab_0^{a-1}U_0 + z + \gamma\tilde{U}_0 + ab_0^{a-1}\frac{\lambda U_1}{r+\lambda+\delta} = 0 \\
&\implies U_0[r+\delta+\gamma+ab_0^{a-1}] = z + \gamma\tilde{U}_0 + \frac{ab_0^{a-1}}{r+\lambda+\delta}(p-\kappa+\lambda U_1).
\end{aligned}$$

$\tilde{V}_0 = 0$ implies that

$$c = b_0^a\frac{p-\kappa}{r+\lambda+\delta} + zb_0 + \gamma b_0\tilde{U}_0 + \frac{b_0^a\lambda U_1}{r+\lambda+\delta} - b_0(r+\delta+\gamma+b_0^{a-1})\frac{z+\gamma\tilde{U}_0+\frac{ab_0^{a-1}(p-\kappa+\lambda U_1)}{r+\lambda+\delta}}{r+\delta+\gamma+ab_0^{a-1}}.$$

Use the last expression to re-write the first-order condition as:

$$\begin{aligned}
c = {} & \frac{b_0^a(p-\kappa+\lambda U_1)}{r+\lambda+\delta} + b_0(z+\gamma\tilde{U}_0) - \frac{b_0(r+\delta+\gamma+b_0^{a-1})(z+\gamma\tilde{U}_0)}{r+\delta+\gamma+ab_0^{a-1}} \\
& - \frac{ab_0^a(r+\delta+\gamma+b_0^{a-1})(p-\kappa+\lambda U_1)}{(r+\lambda+\delta)(r+\delta+\gamma+ab_0^{a-1})} \\
\implies c = {} & \frac{b_0^a(p-\kappa+\lambda U_1)}{r+\lambda+\delta}\left[1-\frac{a(r+\delta+\gamma+b_0^{a-1})}{r+\delta+\gamma+ab_0^{a-1}}\right] + b_0(z+\gamma\tilde{U}_0)\left[1-\frac{r+\delta+\gamma+b_0^{a-1}}{r+\delta+\gamma+ab_0^{a-1}}\right]
\end{aligned}$$

$$\implies c = \frac{b_0^a(1-a)}{r+\delta+\gamma+ab_0^{a-1}}\left[\frac{r+\delta+\gamma}{r+\delta+\lambda}(p-\kappa+\lambda U_1) - (z+\gamma\tilde{U}_0)\right]. \tag{A.52}$$

Before we proceed, we repeat some useful relationships that have already been established in earlier parts of this proof:

$$(r+\delta)\tilde{U}_1 - z = \frac{a\tilde{\theta}_1 c}{1-a}, \tag{A.53}$$

$$(r+\delta)U_1 - z = \frac{a\theta_1 c}{1-a}, \tag{A.54}$$

$$(r+\delta)\tilde{U}_0 - z = \frac{a\tilde{\theta}_0 c}{1-a}. \tag{A.55}$$

The next step is go back to equation (A.52) and plug in $\theta_0 = 1/b_0$, as well as the facts stated in equations (A.53), (A.54), and (A.55). We obtain:

$$(r+\delta+\gamma+a\theta_0^{1-a})\theta_0^a(r+\delta+\lambda) = \frac{1-a}{c}[(r+\delta+\gamma)(p-\kappa+\lambda U_1) - (r+\delta+\lambda)(z+\gamma\tilde{U}_0)].$$

Divide by $(r+\delta+\gamma)$ to get:

$$\left(1+\frac{a\theta_0^{1-a}}{r+\delta+\gamma}\right)(r+\delta+\lambda)\theta_0^a = \frac{1-a}{c}\left[p-\kappa+\lambda U_1 - \frac{r+\delta+\lambda}{r+\delta+\gamma}(z+\gamma\tilde{U}_0)\right],$$

$$\implies (r+\delta+\lambda)\theta_0^a + \frac{r+\delta+\lambda}{r+\delta+\gamma}a\theta_0 = \frac{1-a}{c}\left[p-\kappa+\lambda\left(\frac{z}{r+\delta}+\frac{a\theta_1 c}{(1-a)(r+\delta)}\right) - \frac{r+\delta+\lambda}{r+\delta+\gamma}z\right.$$
$$\left. -\frac{r+\delta+\lambda}{r+\delta+\gamma}\gamma\left(\frac{z}{r+\delta}+\frac{a\tilde{\theta}_0 c}{(1-a)(r+\delta)}\right)\right],$$

$$\implies \theta_0^a(r+\delta+\lambda) + \frac{r+\delta+\lambda}{r+\delta+\gamma}a\theta_0 = \frac{1-a}{c}[p-\kappa-z] + \frac{1-a}{c}\frac{\lambda a\theta_1 c}{(1-a)(r+\delta)}$$
$$-\frac{1-a}{c}\frac{(r+\delta+\lambda)\gamma}{r+\delta+\gamma}\frac{a\tilde{\theta}_0 c}{(1-a)(r+\delta)},$$

$$\implies \underset{\checkmark}{(r+\delta+\lambda)\theta_0^a} + \frac{r+\delta+\lambda}{r+\delta+\gamma}a\theta_0 - \underset{\checkmark}{\frac{\lambda a}{r+\delta}\theta_1} + \underset{\checkmark}{\frac{a\gamma(r+\delta+\lambda)}{(r+\delta)(r+\delta+\gamma)}\tilde{\theta}_0} = \underset{\checkmark}{\frac{(1-a)(p-\kappa-z)}{c}}.$$

In the last expression all the checked terms are identical to equation (48). To conclude the proof, it suffices to show that the multiplier of θ_0 in the (only) unchecked term above coincides with the multiplier of θ_0 in equation (48). That last multiplier is given by:

$$a+\frac{a\lambda}{r+\delta}-\frac{a\gamma(r+\delta+\lambda)}{(r+\delta)(r+\delta+\gamma)} = \frac{a}{r+\delta}\left(r+\delta+\lambda-\frac{\gamma(r+\delta+\lambda)}{r+\delta+\gamma}\right)$$
$$= \frac{(r+\delta+\lambda)a}{r+\delta}\left(1-\frac{\lambda}{r+\delta+\gamma}\right) = \frac{a(r+\delta+\lambda)}{r+\delta+\gamma},$$

which is indeed the multiplier of θ_0 in the unchecked term in the expression above. This concludes the proof.

D.2 Random search and bargaining with segmented markets

We also consider segmented markets for each type of workers where there is random matching within each market and wage is determined through Nash bargaining. As a result, the job finding rate of each type of worker is given by $f_i = b_i^{a-1}$, and the arrival rate of workers to the firms in each market is given by $q_i = b_i^a$, where the queue lengths are defined as $b_i = u_i/v_i$.

Next, we define the value functions of vacant firms in each submarket:

$$\begin{aligned} rV_0 &= -c + q_0(J_0 - V_0), \\ r\tilde{V}_0 &= -c + \tilde{q}_0(\tilde{J}_0 - \tilde{V}_0), \\ rV_1 &= -c + q_1(J_1 - V_1), \\ r\tilde{V}_1 &= -c + \tilde{q}_1(\tilde{J}_1 - \tilde{V}_1). \end{aligned}$$

Note that all the remaining value functions, namely employment and unemployment value functions for each type of worker and value functions for productive firms are identical to the value functions provided in the baseline model except now the arrival rates will be adjusted as discussed above.

For each market, we have the free entry condition. Namely, in equilibrium, we have $V_0 = \tilde{V}_0 = V_1 = \tilde{V}_1 = 0$. This implies

$$\begin{aligned} c &= q_0 \left(\frac{p - k - w_0}{r + \lambda + \delta} \right), \\ c &= \tilde{q}_0 \left(\frac{p - k - \tilde{k} - \tilde{w}_0}{r + \lambda + \delta} \right), \\ c &= q_1 \left(\frac{p - w_1}{r + \lambda + \delta} \right), \\ c &= \tilde{q}_1 \left(\frac{p - \tilde{k} - \tilde{w}_1}{r + \lambda + \delta} \right). \end{aligned}$$

The four free entry condition mentioned above coupled with eight Beveridge curves and the four wage curves identical to the baseline model (again only adjusted for arrival rates) determine the steady state equilibrium. The steady state equilibrium with random search and bargaining with segmented markets consists of unemployment levels for each type, $(u_0, \tilde{u}_0, u_1, \tilde{u}_0)$, employment levels for each type, $(e_0, \tilde{e}_0, e_1, \tilde{e}_0)$, wages in each market, $(w_0, \tilde{w}_0, w_1, \tilde{w}_0)$, and, vacancies in each market, $(v_0, \tilde{v}_0, v_1, \tilde{v}_0)$. The derivations of the wage curves and Beveridge curves are omitted as they are identical to the baseline model.

Under our parameterization, the model of "random search and segmented markets" predicts that the job finding rates of the four worker types will be as follows: $f_0 = 0.99$, $f_1 = 0.47$, $\tilde{f}_1 = 0.43$, and $\tilde{f}_0 = 0.39$. Let us remind the reader (see Table 6) that the analogous numbers for the competitive search specification are: $f_0 = 0.83$, $f_1 = 0.47$, $\tilde{f}_1 = 0.44$, and $\tilde{f}_0 = 0.42$. These numbers are similar, indicating that the random search and segmented markets specification is governed by the same economic forces as the competitive search version of the model. The splitting of the surplus may now be implemented through Nash bargaining, as opposed to wage posting, but, crucially, firms are still able to grasp a healthy share of the future surplus that is generated when they enter the type-0 submarket. Consequently, this specification too predicts that the job-finding rate of type-0 workers will be higher than that of types 1 and $\tilde{1}$, which is in contrast with the data presented in Figure 1.