

Learning Your Comparative Advantages

Theodore Papageorgiou*

Yale University, Department of Economics[†]

December 26, 2007

Abstract

We introduce an equilibrium labor market model, where workers gradually learn about their unobserved production abilities. While engaged in a productive activity (occupation), workers observe their output and extract information that allows them to make inferences about their unobserved aptitudes. Because workers are learning about themselves, their output informs them not only about their productivity in their current occupation, but about their likely productivity in other occupations as well. As workers acquire more information, they self-select into the occupations in which they expect to perform best, and their wages increase. Returns to experience here capture improved job selection by workers as they sort through occupations and learn about their productive abilities. Our setup can account for the offsetting worker flows across occupations, the within-occupation wage inequality, as well the decline in the probability of occupational switching as workers grow older, that we observe in the data. We use this framework to investigate whether the interaction of learning and search frictions can lead to further decreases in output. Indeed, an increase in the unemployment rate similar to the one experienced by many European countries in the early 1970s is found to reduce the flow of output per *employed* worker by 1% annually. If the increase in the unemployment rate disproportionately affects the young, then our estimate is a lower bound for the true cost.

*I thank Giuseppe Moscarini for his invaluable advice and guidance throughout this project, as well as Joe Altonji, Björn Brügemann and Bill Brainard for their suggestions and encouragement. This paper has benefited from discussions and comments of numerous colleagues and seminar participants, but I am particularly grateful to Costas Arkolakis, Eduardo Engel, Nicholas Kalouptsidis, Omar Licandro, Iourii Manovskii, Rafael Melo and Tony Smith.

[†]Cowles Foundation, PO Box 208281, New Haven, CT 06520-8281, USA. Phone: (203) 432-3717. E-mail: theodore.papageorgiou@yale.edu

1 Introduction

“Know thyself”: Learning about one’s self is a life-long process that involves introspection and realization of one’s talents, passions and limitations. In the context of the labor market it involves workers learning about their aptitudes and skills which are not always known to them immediately. While employed in a given productive activity (occupation), workers can observe their own output and make inferences about their unknown talents. Because workers are learning about themselves, the information they acquire allows to better assess their fit not only in their current occupation, but in other occupations as well. For instance, a computer programmer who found out he enjoyed team work and had good interpersonal skills may decide to switch to management. Another worker reported “My career as a small business consultant grew from my love of training employees in my chain of donut shops”¹.

In this paper we explore whether this new paradigm can explain a number of documented facts that are central to labor markets, such as the offsetting gross worker flows across occupations, the increase in wages over a worker’s lifetime, the decline in the occupational switching rate as workers get older, as well as the observed wage dispersion. We also use our setup to investigate the impact of high unemployment on productivity.

More specifically, we develop an equilibrium labor market model where each worker is assigned a set of skills, or type, at birth which remain unchanged throughout the course of his life. Different types have different productivities across occupations, but both the market and the worker are uncertain about the worker’s type. Output realizations reveal information about the worker’s underlying productivity and therefore his type (although stochastic retirement shocks prevent the worker from perfectly finding out his type), which is immediately reflected in his wage and occupational mobility decisions. Therefore, his output informs him not only about his productivity in his current occupation, but about his likely productivity in other occupations as well. The worker’s preferred occupation, however, may not be the one in which he is expected to be most productive, since he also takes into account differences in the speed of learning across occupations in his decision. We moreover incorporate search frictions in our setup. Firms and workers negotiate on splitting their match’s surplus. The worker’s wage is pinned down by the outcome of the negotiation, while free entry ensures zero firm profits (in expectation) in each occupation.

Our model differs substantially from the standard labor market learning model (Jovanovic, 1979, 1984), where a worker learns about the underlying quality of his occupational match rather than about his type. In that model, any information the worker receives about

¹This example is taken from “Making Career Sense of Labour Market Information,” a career guide for Canadian Career Counselors and is also used in Kambourov-Manovskii (2006).

his current firm/occupational match does not reveal anything about future matches: once a worker leaves his current match, he starts over. As a consequence, that model underlines the importance of tenure, rather than of labor market experience, in the wage formation. Indeed there is a large literature on the relative importance of tenure vs. experience on wage formation. Abraham and Farber (1987), Altonji and Shakotko (1987) and Altonji and Williams (2005) find small effects of firm tenure on the wage, whereas Topel (1990) reports larger estimates. Kambourov and Manovskii (2007a) present results that underline the importance of occupational, rather than firm or industry tenure on wages². Gathmann and Schönberg (2006) provide evidence that tenure in the previous occupation also affects the current wage.

Furthermore in the standard learning model, once the worker separates from his previous match he searches randomly since he does not have any additional information about future potential matches. On the contrary in our setup, the worker will direct his search efforts towards the occupation he believes he fits best.

The desire to learn about one's unobserved abilities can also provide an explanation for the decline in occupational mobility with experience, which has been documented in McCall (1990), Neal (1999) and others. Younger workers, who are more uncertain about their type, are more likely to switch occupations, while more experienced ones are more likely to be in an occupation that conforms best with their respective aptitudes.

Our model is partly motivated by the observation that *bidirectional* worker flows between occupations are very large; Kambourov and Manovskii (2006) report that approximately 13% of the labor force switches one-digit occupation every year. These represent transitions across one of nine broadly defined occupational groups, that most likely entail fundamental career changes. In addition, gross worker mobility largely exceeds net mobility: at any given time, almost as many workers enter an occupation as others are exiting it. Clearly an explanation which emphasizes only occupational demand shocks would fail to capture this aspect of the data³. On the contrary, learning about one's abilities provides an intuitive account of these offsetting flows: workers leave an occupation because they have discovered they can perform better elsewhere, while others enter it because they have come to believe it could provide a good fit for them.

The results of Farber and Gibbons (1996) and Altonji and Pierret (2001) who find that the market learns over time about the workers unobserved characteristics motivate our exercise. In particular, Farber and Gibbons (1996) show that the correlation of the worker's unobserved characteristics (such as test scores which might be a good proxy for his abilities)

²They also find positive returns to overall labor market experience.

³Occupation-specific productivity shocks may however explain the *increase* in occupational mobility observed over the past few decades. See Kambourov and Manovskii (2007b) for more details.

with his wage should increase over time, as learning accumulates. Altonji and Pierret (2001) extend their results and show that if the market uses observable characteristics (such as education) to make inferences about the worker's latent abilities, then the correlation of these observables with the wage should fall over time. They find that the data appear to support for these propositions.

MacDonald (1982) also constructs an equilibrium model of the labor market where workers learn about their unobserved abilities. He shows that this form of learning can reproduce the increase in wages over a worker's lifetime and also a positively skewed wage distribution, while all workers in a given occupation receive the same wage. Altonji (2005) has argued that if high-skilled jobs exhibit a higher speed of learning, then the market will be slow to learn the productivity of a worker who starts out in a low-skilled job.⁴ Miller (1984) also combines labor market learning and occupational choice, by extending the model of Jovanovic (1979) and assuming that jobs with similar expected productivity and rate of information acquisition belong to the same occupation. In that model, as in Jovanovic (1979), workers learn about the latent quality of their match, rather than their own unobserved skills. Similarly Antonovic and Golan (2007) examine the occupational choice of workers in the case of general human capital, when jobs differ at the rate of which they reveal information about the workers' skills. The importance of a worker's occupational/career choice has been stressed in several papers: McCall (1990) presents a model where the quality of a match consists of a job-specific and an occupation-specific component. In his setup, information about the quality of these matches arrives at random time intervals. His model predicts that increased tenure in the previous job lowers the likelihood of separation from the current one if both jobs are in the same occupation. Using data from the NLSY he finds evidence that supports his hypothesis. Along the same lines, Neal (1999) presents a model with no learning, where workers draw their job-specific and their career-specific matches. He argues that workers first search for a suitable career and then for a job within that career and provides evidence from the NLSY to confirm his prediction. Sicherman and Galor (1990) explore the issue of career mobility by focusing on intra-firm occupational upgrading and the transferability of human capital across occupations.

We extend the existing literature by showing that this form of learning is consistent with, not only the observed returns to experience, but also two other key labor market facts: the decline in the occupational mobility as workers age and the shape of the cross-sectional wage distribution, including its well-known fat right tail. Furthermore our setup incorporates many features that make it attractive for empirical implementation: we allow workers to

⁴His model assumes a hierarchical ability level, which is a special case of our framework. Our setup is similar to Roy (1951), in that it also allows each worker-type to be the most productive in his occupation.

have different initial priors about their abilities and therefore make different occupational choices; different occupations can have different speeds of learning which workers take into account in their decisions⁵; we include search frictions in our framework and finally we obtain closed-form expressions for steady state distribution of workers by expected productivity and the flows of workers, across occupations, to and from unemployment etc.

We use our framework to investigate the impact of increased unemployment on the output of *employed* workers. In the early 1970s, many European countries experienced large increases in their unemployment rate which persisted for many years, for some to this day. A decrease in the number of employed workers will lower overall output in the economy, simply because there are unused production resources. We demonstrate here that however there is an additional cost, resulting from the reduction of output of employed workers.

In particular, under these new labor market conditions, workers in this economy spend less time employed on average and therefore learn less about their talents. As a result, now workers are more unsure of their skills on average and therefore more likely, once employed to find themselves in an occupation that does not match their aptitudes. Using parameters recovered from the structural estimation of the model, we find that an increase in the unemployment rate similar to the one experienced by several European countries in the early 1970s, results in a reduction to the flow of output per employed worker by 1%. Put differently, employed workers in this economy are now producing 1% less on average every year than they would had they been employed longer and had more time to acquire information about their aptitudes. If the hypothesized increase in the unemployment rate disproportionately affects the young, then this estimate is a lower bound on the true cost.

The paper is organized as follows: Section 2 presents the economic environment and introduces learning and search frictions. Section 3 describes the behavior of agents in the model; we derive the value functions for the worker and the firms and solve out for the wage and optimal worker behavior. Section 4 discusses the model's implications for returns to experience and the occupational switching probability and Section 5 derives the steady state distribution of workers by expected productivity. Section 6 discusses the empirical relevance and predictions of the results, while Section 7 analyzes the effect of high unemployment on the flow of output of employed workers. Section 8 examines possible extensions and Section 9 concludes. The appendix contains detailed derivations and proofs.⁶

⁵As in Antonovic and Golan (2007).

⁶The interested reader can also consult the technical appendix to the present paper, which can be found at <http://pantheon.yale.edu/~tp222>

2 The Economy

Time is continuous. There is a population of risk-neutral workers of mass one. Workers are born and retire at a Poisson rate γ , so that the total population remains constant. Both firms and workers share a common discount rate equal to r . Each worker can be either employed or unemployed and let his flow value of leisure/unemployment benefit while unemployed be equal to b .

Each worker is endowed with a set of skills, which determine how productive he is in each occupation. We call each such set of skills a type. If, for example, there are two occupations in the economy, each type is a two-element vector, where each vector denotes the worker's mean productivity in each occupation. One can think of each of these elements as the productivity resulting from the combination of the different skills of that particular worker-type. For instance, it may represent how the worker's interpersonal, analytical and organizational skills translate into productivity in this particular occupation.⁷

Workers draw their type at birth and it remains unchanged throughout their lifetime: there is no human capital accumulation or learning-by-doing. We examine the case of two possible types, let's call them white (w) and blue (b).

There are two occupations, W and B . In Section 8 we discuss how the model is extended to more than two occupations. Within an occupation, all firms are identical⁸ and there is a large mass of ex ante homogeneous firms ensuring free entry in each occupation.

When the worker of type ν and a firm in occupation i form a match, they produce output that depends on the worker's productivity in that occupation and some i.i.d. shock. In particular, flow output is given by:

$$dY_t^{\nu i} = a_i^\nu dt + \sigma_i dZ_t$$

where dZ_t follows a normal with mean zero and variance dt . a_i^ν is the occupation- and type-specific mean, whereas σ_i is the occupation-specific output noise.

2.1 Learning Frictions

We introduce informational frictions into this economy by assuming that neither the worker, nor the market observe the former's type. The productivity of each type in each occupation,

⁷One can also interpret a worker's occupation-specific productivity as capturing both abilities and preferences. For instance a worker may have the ability to perform quantitative tasks, but not enjoy it and put in the effort, resulting in a low productivity.

⁸A worker must switch firms in order to switch occupations. Gathmann and Schönberg (2006), using German administrative data, report that only 10 percent of switches across 64 occupational groups occur within the same firm.

a_i^v , is common knowledge and so is output. As we will see in the next section, both parties will use output observations to make inferences about the worker's type.

Although they do not perfectly observe their underlying skills, workers are allowed to have different initial beliefs about their types. Let p denote the probability that a worker's type is white. Each worker draws his initial prior, p , from a beta distribution, $g(\cdot)$. This initial belief is also shared by the firms.

Their type will be determined after workers draw their initial belief, with probabilities given by that outcome: the Bernoulli distribution from which the worker's type is drawn, is completely characterized by the initial prior. For example, if a worker's initial prior is 0.3, then a separate draw will determine his type, where white will come up with probability 30% and blue with probability 70%. Thus his initial belief will be informative about his type.

2.2 Search Frictions

Since we wish to explore the impact increased unemployment has on learning, we also introduce search frictions into the model by assuming that it takes time for employed and unemployed workers to contact firms. While unemployed, each worker will choose to search in only one occupation⁹. Unemployed workers will be matched with firms at rate λ_i which we will take as exogenous for the time being, but which we endogenize in the appendix. Similarly employed workers in occupation i can also contact other firms, at rate $\eta_i \lambda_i$. Worker-firm matches can dissolve either endogenously, when the worker's posterior hits some trigger (to be specified later), or exogenously at rate δ_i , after which the worker becomes unemployed. These exogenous shocks capture separations that occur for reasons beyond the modeling assumptions of this paper.

Workers and firms split the surplus generated by search frictions, using Nash bargaining, with $\beta \in [0, 1]$ denoting the worker's bargaining power. If an employed worker contacts another firm, the firms engage in an ascending auction to lure the worker as in Moscarini (2005): firms make offers of a one time transfer plus the promise to bargain bilaterally after that. We analyze the behavior of the agents in this situation and the resulting equilibrium in Section 3.2.

2.3 Productivities: Absolute vs. Comparative Advantage

Assume throughout that a type white (w) worker is more productive in occupation W rather than in occupation B ($a_W^w > a_B^w$) and that type blue (b) worker is better in occupation B

⁹This will be optimal if there is a limited amount of time he can search and there are constant returns to searching.

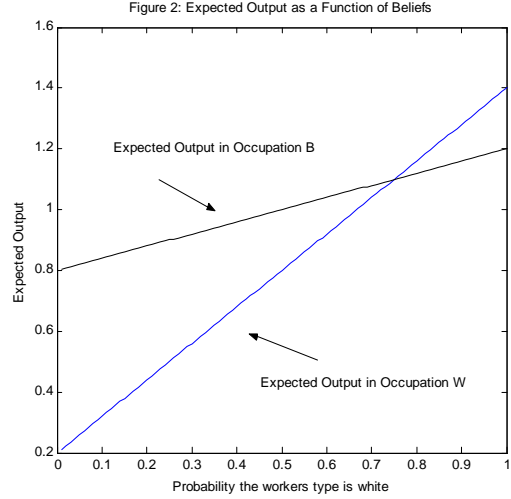
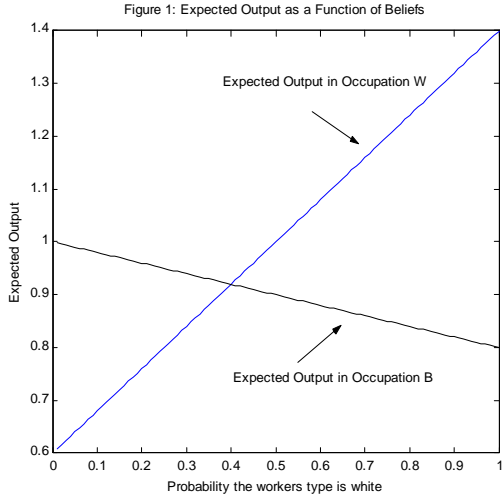
rather than in occupation W ($a_W^b < a_B^b$). Clearly if both types are more productive in the same occupation, the worker would always choose it.

If there were no learning frictions, workers would sort according to absolute advantages as described above. On the other hand, if there was only one occupation, there would only be learning about the worker's ability. In a world with two (or more) occupations and imperfect information, workers sort according to comparative rather than absolute advantage. To see this, notice that in the absence of search frictions and without taking into account differences in the speed of learning (discussed in the next section), the worker will choose to work in W if and only if his expected output in W is higher than in B :

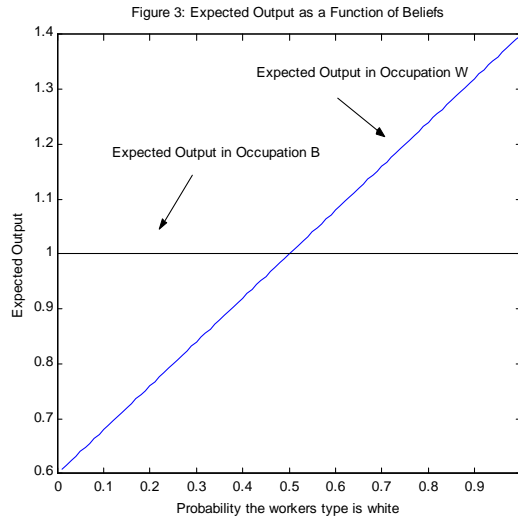
$$pa_W^w + (1-p)a_W^b > pa_B^w + (1-p)a_B^b \Leftrightarrow \frac{p}{1-p} > \frac{a_B^b - a_W^b}{a_W^w - a_B^w}$$

Therefore his occupational choice depends on relative productivities. Even if p is close to 1, the worker will still choose to work in occupation B if the relative productivity differences across occupations for type blue are large enough compared to those of type white. Essentially, the worker chooses to work in B , if and only if the relative gain from producing in that occupation, when his type is blue, is sufficiently larger than the relative gain from working in W , when his type is white. Sorting according to comparative advantage occurs in a similar fashion when there are three or more occupations.

There are three cases of interest that can arise in the two occupations, two types framework. The first is when each worker type is better than the other in the occupation in which he is more productive: as we assumed above a white type is better in W than in B ($a_W^w > a_B^w$), but we moreover assume that he is better than the blue type in W ($a_W^w > a_W^b$). Similarly a blue type, who is more productive in B rather than in W ($a_W^b < a_B^b$), is now assumed to be better than a white type in B ($a_B^w < a_B^b$). This is depicted in Figure 1 which shows the worker's expected output in each occupation as a function of his belief. Note that although expected outputs are equalized for $p = 0.4$, this does not imply that he will necessarily be indifferent between the two occupations at that point, since differences in the speed of learning are also important to him.



This becomes more apparent in Figure 2, which depicts the second case: now type white is not only better in W compared to blue ($a_W^w > a_W^b$), but he is also better in B ($a_B^w > a_B^b$). In other words the white type has an absolute advantage in both occupations¹⁰. In Figure 2, expected productivities are equalized for $p = 0.75$. As we will see in the next section however, at that value of the posterior the worker would still prefer to work in occupation W , because the spread between the output of the two types is larger and therefore output realizations now reveal more information (assuming that the output noise parameter σ_i is the same in both occupations).



¹⁰The case where the blue type has an absolute advantage in both occupations is equivalent.

Finally, Figure 3 depicts the case where there is no learning in one occupation. In this case occupation B is an absorbing state, since, once the worker enters it, no further information about his type is revealed (there is no way to distinguish between the two types) and his posterior no longer changes so he has no incentive to switch again. Notice that at $p = 0.5$ he will still choose occupation W because of the informational value of working there, even though his expected flow output is larger in occupation B .

In the remainder of the paper we will assume that a white type is more productive than a blue type in occupation W ($a_W^w > a_W^b$); whether or not he is also more productive than a blue type in occupation B will be important for our results.

3 Behavior

3.1 Belief Formation

If there was no output noise ($\sigma_i = 0$), worker and firms would be able to perfectly observe the worker's productivity in occupation i and learn his type, as long as $a_i^w \neq a_i^b$. If however $\sigma_i > 0$, they observe it only imperfectly. After noticing flow output each period, $dY_t^{\nu i}$, the market and the worker both update their belief using Bayes' rule. Since the market shares the same initial belief as the worker about the latter's type and output history is perfectly observable, at no point is there any divergence between the beliefs of the two. The belief, p , constitutes a sufficient statistic of both the worker's output history as well as his initial belief. As in Liptser and Shyryaev (1977), using Bayes' rule and Ito's lemma, the belief process is reduced to:

$$dp_t = p_t(1-p_t) \frac{a_i^w - a_i^b}{\sigma_i} \frac{dY_t^{\nu i} - (p_t a_i^w + (1-p_t) a_i^b) dt}{\sigma_i}$$

The change in beliefs will therefore depend on three components: the current variance in beliefs $p_t(1-p_t)$, the signal to noise ratio $\frac{a_i^w - a_i^b}{\sigma_i}$ and the normalized difference between realized and expected flow output. The higher the variance, the lower the precision of the current belief and the more likely it is to change. High signal to noise ratio (large difference in the means and/or low noise), signifies more informative signals. Finally, a large deviation between expected and actual output reveals more information, inducing a larger change in beliefs. Also note that this last term on the right hand side is a standard Wiener process with respect to the unconditional probability measure over types and signal realizations used by the agents of this economy.

3.2 Wage Determination

Since p summarizes the worker's and the firm's beliefs about the worker's type, it also captures their expectations regarding the future value of their match and will serve as a state variable for their values, as well as for the bargained wage. As mentioned above, the worker's wage is determined by (continuous) bilateral negotiation between the worker and the firm; as the beliefs change about the worker's type, so will the worker's wage, as both the worker and the firm revise their value of the match's surplus. The (cooperative) outcome of this negotiation is by assumption given by Nash bargaining, which dictates that a worker's wage in occupation $i \in \{W, B\}$, given beliefs p about his type, $w_i(p)$, is set according to:

$$w_i(p) = \arg \max_w [J_i(p)]^{1-\beta} [V_i(p) - U(p)]^\beta \quad (1)$$

where $J_i(p)$ denotes the asset value of the firm in occupation i , $V_i(p)$ denotes the value of a worker employed in occupation i and $U(p)$ denotes the value of an unemployed worker, given beliefs p about the worker's type.

As mentioned above, when an employed worker contacts another firm, then the firms engage in an ascending auction to lure the worker as in Moscarini (2005): firms make bids of a one time transfer plus the promise to bargain bilaterally after that. Let $S_i(p)$ denote the surplus of the match in occupation i ¹¹. Since the posterior belief about the worker is shared by all agents, the value of $S_i(p)$ will be common knowledge. Furthermore assume that worker search activity is not observable by their employer.

Let's first examine the case where the incumbent firm's occupation, i , is different from the occupation of the poaching firm, k ($i \neq k$). In this case, if $S_k(p) > S_i(p)$, the poaching firm, k , can always outbid the incumbent, i , in a bidding war: the highest possible bid the incumbent is willing to make is worth $S_i(p)$ ¹² which is strictly less than the maximum bid of the poaching firm, $S_k(p)$. The incumbent knows this so he does not even bother bidding and the worker switches to the poaching firm; the lump-sum transfer is zero and from then on, the worker will bargain bilaterally with the firm over the match's surplus and receive $V_i(p)$ ¹³. Alternatively, if $S_k(p) < S_i(p)$ the poaching firm knows it cannot outbid the incumbent so it never bothers to bid. If $S_k(p) = S_i(p)$ the worker stays with the incumbent. The above strategies constitute a subgame perfect Nash equilibrium.

¹¹Here $S_i(p) = W_i(p) - U(p) + J_i(p) - P_i$, where P_i is the value of a unfilled vacancy in occupation i . Free firm entry ensures that $P_i = 0$, $i \in \{W, B\}$.

¹²In that case the worker will receive $W_i(p)$, as prescribed by the Nash bargaining problem (eq. (1)) and $S_i(p) - W_i(p)$ as one-time lump sum transfer, effectively capturing the entire surplus of the match.

¹³Both firms bidding up to $S_i(p)$ and $S_k(p)$ respectively also forms a subgame Nash equilibrium of this game. However such an equilibrium is not robust to the perturbation of adding a small cost of bidding.

In the case where a worker receives an offer from a firm in the same occupation, since he is equally valuable to both of them, adding a small cost for the poaching firm to enter an auction will sustain a Nash equilibrium where the worker stays with incumbent with no retention bonus. Thus the worker will never have incentive to search for a job in the same occupation.¹⁴

In our setup, the solution to the Nash bargaining problem results in the linear sharing rule:

$$\beta J_i(p) = (1 - \beta) [V_i(p) - U(p)] \quad (2)$$

which provides the necessary condition to determine the worker's wage¹⁵.

3.3 Value Functions and Worker Behavior

In what follows we will assume for simplicity that the job finding rate is the same for both occupations $\lambda_W = \lambda_B = \lambda$. This assumption ensures that occupational choice and mobility, as well as wage formation, depend mainly on differences in productivity and speed of learning across occupations¹⁶. In Section 8 we discuss how the model is modified when we allow for different job finding rates across occupations.

The process that governs the change in beliefs is a diffusion without a drift, so using Ito's lemma we can write the flow value of an employed worker in occupation $i \in \{W, B\}$ as:

$$\begin{aligned} rV_i(p) = & w_i(p) + \frac{1}{2} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1-p)^2 V_i''(p) \\ & - \delta_i [V_i(p) - U(p)] - \gamma V_i(p) + \eta_i \lambda \max \{V_k(p) - V_i(p), 0\} \end{aligned} \quad (3)$$

where r denotes the worker's and the firm's discount rate, $V_i(\cdot)$ is the value of the worker in occupation i , $w_i(\cdot)$ is the occupation-specific wage function and $U(\cdot)$ the asset value of an unemployed worker.

¹⁴One can extend the present setup to allow for within-occupation job switches by assuming heterogeneity in firm productivities and/or differences in the non-pecuniary benefits to the worker.

¹⁵Shimer (2006) has shown that with on-the-job search this linear sharing rule may not always be bilaterally efficient. In particular he has argued that, if $S_k(p) > S_i(p)$ the incumbent employer might have incentive to pay the worker a higher wage in exchange for not searching on the job. However, as Moscarini (2005) notes, in a framework like the present where job search is costless (and not observed), such a strategy by the employer cannot work, because the worker would accept the higher wage, but continue to search on the job, knowing that any poaching firm can outbid the wage he is currently being paid. Thus in present setup, the Nash bargaining solution will be bilaterally efficient.

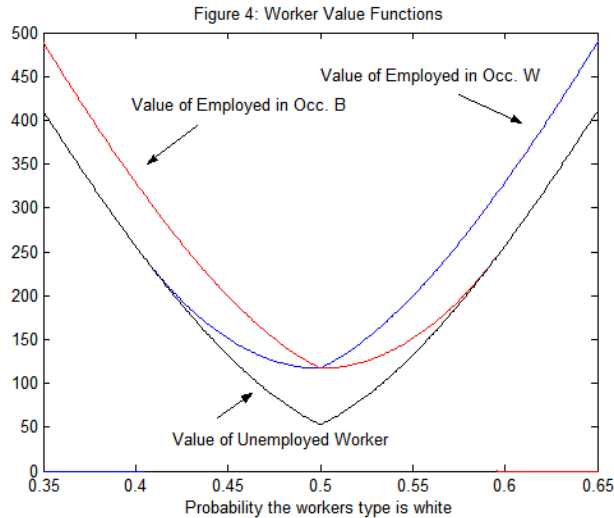
¹⁶We do allow occupations to have different exogenous separation rates and different job finding rates for employed workers, which are also taken into consideration in the worker's decisions.

The value of being employed is equal to the wage enjoyed by the worker plus a term that captures the gains from learning, minus the capital loss resulting from either exogenous match separation which is realized at rate δ_i , or retirement occurring at rate γ , plus any potential gains from on the job search (as we saw in the previous section, a worker has no incentive to search in the same occupation). Notice that the value of the learning component will depend on the signal to noise ratio and the current precision of the agent's beliefs. In our model, in contrast to the Jovanovic model, the value of learning extends beyond the duration of the current occupational match: what the worker (and the market) learns about himself in occupation i will also be useful if he's unemployed or employed in another occupation k .

Similarly the flow value to the firm of a filled vacancy in occupation i is given by:

$$rJ_i(p) = \bar{a}_i(p) - w_i(p) + \frac{1}{2} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1-p)^2 J_i''(p) - \delta_i J_i(p) - \gamma J_i(p) - \eta_i \lambda J_i(p) I \{V_i(p) < V_k(p)\} \quad (4)$$

where $J_i(\cdot)$ is the asset value of the firm, $\bar{a}_i(p) = pa_i^W + (1-p)a_i^B$ is the expected output given beliefs p about the worker, and $I\{\cdot\}$ is an indicator function which in this case marks whether the worker is searching on the job or not. Therefore the flow value of the firm is equal to expected output, minus the wage, plus a term that measures the value of learning to the firm, minus the potential capital loss resulting from an exogenous separation, worker retirement or worker transition to another job. For the firm, unlike the worker, the value of learning is limited only to the duration of the current match.



The worker and firm will mutually decide to separate if the value of the match's surplus

becomes zero. For the case of occupation W let \underline{p} denote the value of the posterior such that $V_W(\underline{p}) = U(\underline{p})$ (or equivalently $J_W(\underline{p}) = 0$); while $p > \underline{p}$ the worker prefers being employed in occupation W to being unemployed (but not necessarily to working in occupation B), but when p reaches \underline{p} the worker optimally quits to unemployment. Similarly if we let \bar{p} denote the value such that $V_B(\bar{p}) = U(\bar{p})$ (or equivalently $J_B(\bar{p}) = 0$), the worker optimally quits to unemployment when his p reaches \bar{p} ¹⁷. The optimality of the worker's decision entails a smooth pasting condition at \underline{p} and \bar{p} . These thresholds can be seen in Figure 4: \underline{p} is equal to approximately 0.4, while \bar{p} equals approximately 0.6. The worker's value of being employed in occupation W is always above his value of being unemployed, as long as $p > 0.4$. As p converges to 0.4 the difference between the two values shrinks, until they become tangent (smooth pasting), at which point the worker quits and enters unemployment to search for the job in occupation B .

The worker's flow value of being unemployed is:

$$rU(p) = b + \lambda \max_i [V_i(p) - U(p)] - \gamma U(p) \quad (5)$$

the flow benefit b , plus the excess value from being employed in occupation i times the job finding rate, λ , minus the capital loss in case of retirement. Again notice that since workers are learning about their general human capital, the value of being unemployed is a function of the worker's current belief about his type, unlike the Jovanovic model where beliefs are reset upon separation.

Let \hat{p} denote the value of the posterior at which the worker is indifferent between searching for a job in occupation W or B :

$$V_W(\hat{p}) = V_B(\hat{p})$$

For $p \geq \hat{p}$, the worker will search for employment in occupation W while for $p < \hat{p}$, he will search in occupation B . Note that \hat{p} also determines whether an employed worker searches on the job or not: as beliefs cross \hat{p} the relative value of being employed in one occupation compared to the other changes. In the symmetric case depicted in Figure 4, this is true for $\hat{p} = 0.5$. In the appendix we show that \hat{p} is unique.

3.4 Triggers and Wages

We will now describe how we solve for the wage and the worker's decision rules. For a complete derivation of all the steps the reader should refer to the appendix. Since the

¹⁷In the appendix we show that $V_W(\cdot) - U(\cdot)$ (as well as $J_W(\cdot)$) is increasing in p everywhere. Similarly $V_B(\cdot) - U(\cdot)$ (and $J_B(\cdot)$) will be decreasing in p everywhere.

worker's behavior changes depending on whether his posterior is to the left or right of \hat{p} we consider each case separately

In equations (2) through (5) we have four unknowns functions: the 3 value functions and the wage function. After some algebra (see appendix), we can solve for the wage as a function of the firm's value in the case where the worker's outside option is searching for another job in his current occupation:

$$w_i(p) = (1 - \beta)(b + \lambda[V_i(p) - U(p)]) + \beta\bar{a}_i(p) + \frac{1}{2} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 [\beta(J_i''(p) + V_i''(p)) - V_i''(p)] \quad (6)$$

The worker's wage weights his outside option, namely his value of being unemployed b , plus the option value of search while unemployed, and his inside option. The weights are given by his bargaining power coefficient, β . Here his inside option is his share of the output plus his share of match's total value of learning that is *in excess* of his own private value. For example if $\beta(J_i''(p) + V_i''(p)) - V_i''(p) < 0$, the worker will compensate his employer for the additional benefit he enjoys from learning. Remember that worker's gain from learning exceeds the life of the match, while the firm's benefit is limited only to the current match.

To understand better how the asset value of learning is split by the worker and the firm, notice that if we substitute in his wage, the worker's value while employed (eq. (3)) becomes:

$$(r + \gamma)V_i(p) = (1 - \beta)(b + \lambda[V_i(p) - U(p)]) + \beta \left(\bar{a}_i(p) + \frac{1}{2} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 (J_i''(p) + V_i''(p)) \right) - \delta_i[V_i(p) - U(p)]$$

where it is clear how the worker benefits only from his bargained share of the value of learning.

Similarly, in the case where the worker's outside option differs from his current occupation his wage is:

$$\begin{aligned}
w_i(p) &= (1 - \beta) (b + \lambda [V_k(p) - U(p)]) + \beta \bar{a}_i(p) \\
&+ \frac{1}{2} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 [\beta (J_i''(p) + V_i''(p)) - V_i''(p)] \\
&- \eta_i \lambda [(1 - \beta) (V_k(p) - V_i(p)) + \beta J_i(p)]
\end{aligned} \tag{7}$$

The interpretation is similar to equation (6) above with two differences: the worker's outside option is now different (if unemployed he will look for a job in another occupation) and his wage is reduced by an amount proportional to his search intensity. When the worker leaves his current firm for a firm in another occupation, the separation is no longer bilaterally efficient, because there are lost rents for the incumbent firm. Therefore when the worker searches on the job, the worker compensates his firm by an amount that is given by the weighted average of the worker's gains, $V_k(p) - V_i(p)$, and the firm's losses, $J_i(p)$.

Substituting in for the wage in the firm's value in each case results in a differential equation with respect to $J_i(\cdot)$ in the case where the worker's outside option is his current occupation and in the case where it is not:

$$(r + \gamma + \delta_i + \beta\lambda) J_i(p) = (1 - \beta) (\bar{a}_i(p) - b) + \frac{r + \gamma + \beta\lambda}{2(r + \gamma)} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 J_i''(p) \tag{8}$$

$$\begin{aligned}
(r + \gamma + \delta_i + \eta_i\lambda) J_i(p) &= (1 - \beta) (\bar{a}_i(p) - b) + \frac{\beta\lambda}{2(r + \gamma)} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 J_k''(p) \\
&+ \frac{1}{2} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 J_i''(p) - \beta\lambda (1 - \eta_i) J_k(p)
\end{aligned} \tag{9}$$

Using the boundary conditions $J_W(p), J_B(p) < \infty, \forall p$ we can solve the differential equation for the asset value of the firm in the case where $i = W$ and $p \geq \hat{p}$:

$$J_W(p) = \frac{(1 - \beta) (\bar{a}_W(p) - b)}{r + \gamma + \delta_W + \beta\lambda} + C_1^W p^{\frac{1}{2}} (1 - p)^{\frac{1}{2}} \left(\frac{1 - p}{p} \right)^{\frac{1}{2}} \sqrt{\frac{4 + h_W}{h_W}} \tag{10}$$

as well as for the case where $i = B$ and $p \leq \hat{p}$:

$$J_B(p) = \frac{(1 - \beta) (\bar{a}_B(p) - b)}{r + \gamma + \delta_B + \beta\lambda} + C_2^B p^{\frac{1}{2}} (1 - p)^{\frac{1}{2}} \left(\frac{p}{1 - p} \right)^{\frac{1}{2}} \sqrt{\frac{4 + h_B}{h_B}} \tag{11}$$

where $h_i = \frac{1}{2} \frac{r + \gamma + \beta\lambda}{(r + \gamma + \delta_i + \beta\lambda)(r + \gamma)} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2$ and C_1^W and C_2^B are undetermined coefficients.

Now we can substitute in for $J_k(p)$ in equation (9) and using the condition that $J_W(\underline{p}) = 0$ and $J'_W(\underline{p}) = 0$ for occupation W ($J_B(\bar{p}) = 0$ and $J'_B(\bar{p}) = 0$ for occupation B), we are able to solve the resulting differential equations (resulting expressions in the appendix).

To complete the solution of the value functions and the worker's decision rules we need to pin down the value of the 5 remaining unknowns: the 3 triggers, \underline{p} , \bar{p} and \hat{p} , as well as the 2 yet undetermined coefficients C_1^W and C_2^B . We need 5 conditions to do so. Optimality of searching behavior while unemployed, $V_W(\hat{p}) = V_B(\hat{p})$, provides the one of these conditions. The remaining 4 are given by continuity of the total value of the match ($V_i(\cdot) + J_i(\cdot)$) at \hat{p} , as well as continuity of its first derivative, for each occupation. The solution of the resulting non-linear system of 5 equations and 5 unknowns allows us to fully characterize both the wage for every occupation and value of the posterior, as well as the optimal behavior of the worker.

Proposition 1 *An unemployed worker searches for a job in occupation W if his posterior belief of being type w is greater than \hat{p} and in occupation B if his posterior is less. \hat{p} also determines the worker's search behavior on the job. He quits his job in occupation W when p reaches \underline{p} and in occupation B when p hits \bar{p} . The value of a firm employing a worker who is not searching on the job is given by (10) and (11) for occupations W and B respectively and by (17) and (18) (of the appendix) respectively when the worker is searching on the job. In the first case the worker's wage is given by (6) and in the second by (7). Finally the values of the 3 triggers \underline{p} , \bar{p} and \hat{p} , and the 2 undetermined coefficients C_1^W and C_2^B are determined by the solution of the system of equations (19)-(23), also in the appendix.*

4 Returns to Experience and Probability of Occupational Switching

Workers are expected to learn their type over time, since output signals are informative. Although a worker's posterior belief is a martingale, if we condition on his true type, his posterior is either strictly increasing or decreasing in expectation. For example, for the case of a type w worker in occupation i :

$$E(p_{t+\Delta t} | \text{true type is } w) = p_t + p_t(1-p_t)^2 \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 \Delta t > p_t = E(p_t)$$

In other words, if his true type is w , his posterior will converge in expectation to one almost surely. Similarly the belief of a blue worker will converge to zero.

Since market production contains information about the worker's type, this information is

incorporated in the priors and the true type is slowly revealed. For example if a white worker is employed in occupation B , as output realizations reveal information, the posterior about the worker's type is expected to rise and will eventually switch to working in occupation W , either by on-the-job search or quitting to unemployment. Similarly for a blue worker employed in occupation W .

Lemma 2 *Conditional on his true type, in expectation the worker's posterior is strictly increasing over time when his type is w and decreasing when it is b .*

Therefore workers are expected to learn their true type over time and self-select into the occupation they perform best.

In the case where no worker has an absolute advantage ($a_B^w < a_B^b$), which is depicted in Figure (1), the worker's expected output will increase over time and so will the value of the match's surplus. White workers will eventually self-select to occupation W , while blue workers will switch to occupation B . Their expected output will increase and in general so will their wage. To see this, notice that the value of the match's surplus is essentially equal to the value of output that it produces, plus the value of learning. As long as output increases in expectation, so will the surplus and therefore so will the worker's wage. We cannot prove however that the wage will be universally increasing in expectation. It is possible for example that for some parameter values, in some regions of p , as p approaches 0 or 1, the decline in the value of learning is larger than the increase in expected output (in expectation). In those cases, the worker's wage may decline, before it picks up again (it is not possible that the wage is universally declining). In the estimation of the model (Papageorgiou, 2007), we show that the model does indeed produce an upward sloping wage-age profile.

In the case however where a white worker type is better at both occupations ($a_B^w > a_B^b$), as in Figure 2, expected output will increase only for the white workers. In this case, we expect only white types to enjoy positive returns to experience, while blue types will see their wage decline over time, as their type is revealed. Now the wage decline masks what would be an even worse outcome should the blue worker stay in occupation W .

Furthermore, the occupational probability switching rate falls with experience: as long a worker in occupation W is above \hat{p} (below \hat{p} for occupation B), the instantaneous probability of switching occupations is zero; he would never accept an offer from another firm in the other occupation and if he were to become unemployed exogenously, he would search for a job in the same occupation. Once his p crosses \hat{p} , the instantaneous probability jumps, since now he is searching on the job and switches occupations at rate $\eta_i \lambda$ and if his match is destroyed due to a δ_i shock, he will search for a job in occupation B . Of course, when p hits \underline{p} (\bar{p} when employed in B), he quits and searches for a job in the other occupation. Therefore,

since beliefs are expected to converge to zero or one, occupational mobility declines with labor market experience. The greater a worker's employment history, the more likely it is that he will have self-selected into the occupation that he is most productive.

5 Aggregation

Summarizing worker behavior: if unemployed, a worker will look for a job in occupation W , if his posterior, p , is greater than \hat{p} and in occupation B otherwise. In both occupations he lands a job at rate λ . While employed in W , he will optimally quit if and when p hits \underline{p} , unless the match gets hit by an exogenous shock δ_W and is dissolved. If while employed, p falls below \hat{p} he starts searching on the job and is matched with a firm in occupation B at rate $\eta_W\lambda$. Similarly if employed in occupation B he will optimally quit if p reaches \bar{p} , conditional on not having been hit by a δ_B shock. Again if p exceeds \hat{p} he searches on the job for another match in occupation W , which materializes at rate $\eta_B\lambda$. Finally he exogenously retires at rate γ , regardless of whether he is employed or unemployed.

Since the process for the beliefs is Markovian, the stationary distribution will be ergodic. Let $F_i(p)$ denote the population of workers employed in occupation i whose posterior probability of being of a white type is less than p (thus $f_i(p)$ denotes the population density of employed workers in occupation i). Similarly $Z_i(p)$ is the population of those unemployed who are looking for a job in occupation i whose posterior probability of being a white type is less than p and $z_i(p)$ denotes the corresponding population density of unemployed workers in occupation i .

Following Karlin and Taylor (1981) (chapter 15), the Kolmogorov forward equation for occupation W for every $p \geq \underline{p}$ and $p \neq \bar{p}$ is given by:

$$0 = \frac{df_W(p)}{dt} = \frac{d^2}{dp^2} \left[\frac{1}{2} \left(\frac{a_W^w - a_W^b}{\sigma_W} \right)^2 p^2 (1-p)^2 f_W(p) \right] - \delta_W f_W(p) - \gamma f_W(p) + \lambda z_W(p) - \eta_W \lambda f_W(p) I\{p < \hat{p}\} + \eta_B \lambda f_B(p) I\{\hat{p} \leq p \leq \bar{p}\} \quad (12)$$

This equation ensures that flows in and out of every $p \neq \bar{p}$ in the distribution of employed workers in occupation W are equal. The first term captures the net change in p caused by workers moving into p from the right and left of that point, as well as those workers moving away from p . The second and third term measure the outflow from p resulting exogenous job destruction and worker retirement shocks respectively, whereas the fourth term captures the inflow of new workers from unemployment at p . Finally the last two terms reflect the outflow

from p of workers to other jobs in occupation B and the worker inflow from occupation B to newly created matches in occupation W at p .

Similarly we can write the Kolmogorov forward equation for occupation B . We would like to use these two equations and solve out the resulting system for $f_W(\cdot)$ and $f_B(\cdot)$. However we first need to solve out for the population density of unemployed workers in each occupation i , $z_i(\cdot)$. To do so we make use of the condition that in the steady state, flows in and out of every p in the distributions of unemployed workers must equal. Therefore those searching for a job in occupation W , for every $p \geq \hat{p}$ and $p \neq \bar{p}$:

$$\delta_W f_W(p) + \delta_B f_B(p) + \gamma g(p) = \lambda z_W(p) + \gamma z_W(p) \quad (13)$$

The first two terms on the left hand side represent the inflow to unemployment due to exogenous match destruction shocks from occupations W and B respectively, whereas the third represents for the inflow of newly born workers at p . The two terms on the right hand side account for the exit of workers from p because they either find a job or they retire. Using this condition, we can easily substitute out for $z_W(p)$. Similarly we can write the corresponding condition for those unemployed in occupation B .

After solving out for $z_W(\cdot)$ and $z_B(\cdot)$ we can substitute them into the two forward equations and derive a system of second order differential equations, with respect to $f_W(\cdot)$ and $f_B(\cdot)$. The resulting solution contains 12 undetermined coefficients and we make use of 12 conditions to pin them down. For more details on the solution, the interested reader should refer to the appendix in the present paper, as well as the technical appendix found at the author's website.¹⁸

In the appendix, we also show that the distribution of workers' posteriors in occupation W (B) will feature a fat right (left) Pareto tail, if $\gamma > \left(\frac{a_W^w - a_W^b}{\sigma_W}\right)^2 \frac{\lambda + \gamma}{\delta_W + \lambda + \gamma}$ or $\gamma > \left(\frac{a_W^w - a_W^b}{\sigma_W}\right)^2$ ($\gamma > \left(\frac{a_B^w - a_B^b}{\sigma_B}\right)^2$) approximately. Otherwise, it will be increasing as p goes to 0 or 1. These conditions have an intuitive economic explanation: since beliefs are reset only upon retirement, if the speed of learning in the occupation is slower than the retirement rate, worker's beliefs are less likely to reach the extremes of the support before being forced to reset. Furthermore we derive conditions under which the within occupations cross-sectional distribution of wages features a fat, Pareto-type tail.

¹⁸Currently at <http://pantheon.yale.edu/~tp222>

6 Discussion

Before we proceed to investigate the impact of increased unemployment on output per worker we discuss some of the model's implications and assumptions. In our economy, returns to experience and wage dispersion are not the result of heterogeneity across workers in their observed ability levels (human capital accumulation). Although there are search frictions in this model, they neither are the driving force behind generated wage dynamics and dispersion as is the case in e.g. the Burdett and Mortensen (1998) model. On the contrary, here returns to experience and wage dispersion capture both learning, as well as the selection of workers into activities they fit best. Even though a worker performs differently across occupations, previous work experience is an important determinant of the current wage. Even in the case of an exogenous match destruction, the wage in the new job will continue to reflect all the accumulated knowledge about the worker's type.

It is important to notice that if we assume a one-dimensional ability level, returns to labor market experience may be negative for the truly low ability workers, whereas in the alternative case, where some workers perform better in one occupation and others at another, *all* workers enjoy positive returns to experience in expectation. This is a potentially testable implication of the model that allows one to distinguish between the two cases.¹⁹

As noted in the Introduction, the model predicts that at any given time there are simultaneous and offsetting flows to and from each occupation: at any one time workers exit a given occupation because they realize they are more productive elsewhere, while simultaneously, other workers enter it because they have come to believe it may be a good fit for them. Given the growing empirical evidence that points to the excess of gross worker flows over the net²⁰, this prediction contrasts sharply with models that assume that workers switch occupations because of occupational productivity shocks that affect all workers in the same way.

Furthermore, we are able to obtain closed-form expressions for the cross-sectional distribution of worker posteriors in each occupations and show that under certain conditions that they feature Pareto-type tails. Moreover we show that under certain conditions, the wage distribution of both occupations will feature a fat right tail. Therefore our setup is potentially capable of capturing the well-documented shape of the empirical wage distribution and account for the observed wage inequality.

Search in our framework, although undirected within an occupation, is directed between occupations: workers make their choice, based on their expected productivity in each occupation and the speed of learning. Therefore occupational choice is not random, as would

¹⁹The results of the model's structural estimation (Papageorgiou, 2007) favor the case where each worker-type is the most productive in his or her occupation.

²⁰Murphy and Topel (1987), Jovanovic and Moffitt (1990), Kambourov and Manovskii (2006).

be the case in a model in which occupational match is an experience good and a worker's productivity in a given occupation is independent of his productivity in another occupation. Such a model would suggest that in the occupational transition matrix created from the data, all the off-diagonal elements would be approximately equal, whereas the model described in the present paper implies that from a given occupation, transitions to certain occupations are more likely than to others. In the occupational transition matrices in McCall (1990) and Kambourov and Manovskii (2006), there are noticeable divergences among the off-diagonal elements, indicating that switches to some occupations are more likely than to others, from a given occupation. Gathmann and Schönberg (2006) also find that the current occupation affects future occupational choices.

Some assumptions of the model require some consideration. First of all, there are no information asymmetries in our setup: both the worker and the market have the same amount of information regarding the former's type. In reality however, one would expect a degree of informational asymmetry between both the incumbent firm and other firms, as well as between the worker and the market. Let us examine each case separately.

We do not view the assumption of no informational discrepancy between the incumbent and the other firms as particularly restrictive; in most occupations a worker's CV contains sufficient information to largely bridge the informational gap between the incumbent and a poaching firm. For example an architect's CV contains all the relevant information about his previous work, in order to allow a third party to make an informed estimate of his ability. Furthermore, the incumbent firm's actions also reveal information regarding the worker's abilities (e.g. promotions, bonuses, firing etc). Furthermore, recent results by Schönberg (2007) who tests for asymmetric employer learning, show that this assumption is not particularly restrictive.²¹ Given the above we conclude that the additional level of realism asymmetric information would lend to the model does not warrant the additional technical complexity that would be required.²²

A potentially more severe limitation of the model is that it assumes that the worker and the market have the same information. For example, Holmstrom (1999) has argued that if a worker has superior information about his type, he might take actions that influence the market's learning process. Alternatively one might think that the market can better assess the fit of a young, inexperienced worker in each occupation. The extension of the present model to include information asymmetries is an interesting, but challenging question that is left for future research. Equally important would be an empirical investigation on the extent

²¹Pinkston (2006) finds evidence of public learning as experience increases and asymmetric learning as the length of the employment spell increases.

²²For an example of a 3-period labor market model with firm asymmetric information the interested reader should refer to Eeckhout (2006).

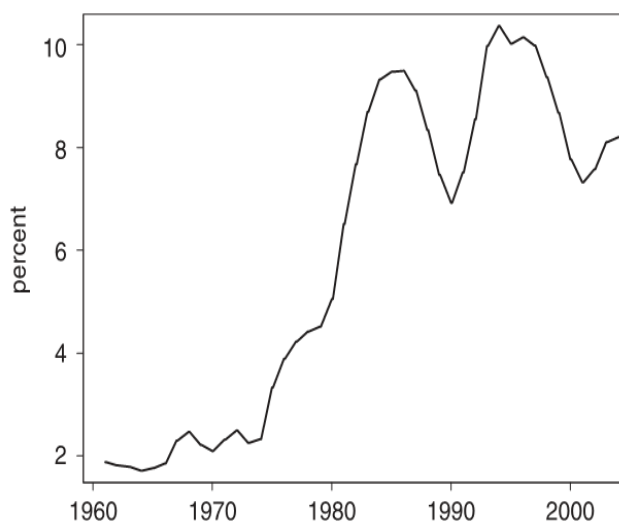
of asymmetric information between the worker and the market about the former's aptitudes, which to this day remains an open question.

Another potential limitation of the model is that it does not generate job-to-job transitions within the same occupation. Since the worker is equally valuable to all firms in the same occupation, there is no incentive for him to switch. However this is not an inherent shortcoming of the model and can be amended by extending the model to allow for heterogeneity in labor demand within each occupation.

7 The Productivity Cost of High Unemployment

Our model provides an appropriate framework to investigate the impact increased unemployment may have on labor allocation. In the early 1970s several European countries experienced large increases in their unemployment rate that persisted for many years. This is depicted in Figure 5, which shows the average European unemployment rate since 1960. Although much research has focused on identifying the causes²³, little attention has been paid to the costs of high unemployment on the allocation of labor resources. Clearly an increase in the percentage of workers not employed will decrease output produced. Nevertheless, we will argue here that there is an additional cost for workers who *are* employed.

Figure 5, EU15 Unemployment Rate, since 1960



Source, OECD database (Blanchard, 2006)

Using our framework, one can observe that an by reducing the time each worker spends

²³For an overview of the literature and evolution of ideas on the sources of the increased, the interested reader may refer to Blanchard (2006).

employed, there is less overall learning and sorting in economy. Workers on average are now more unsure about their abilities and therefore more likely to be employed in an occupation that does not match their talents.

To explore this issue further and assess the magnitude of the costs involved, we will use the parameter values generated by the model's structural estimation in Papageorgiou (2007). In that paper we use the 1996 panel of Survey of Income and Program Participation and partition the occupations into white collar jobs (W) and blue collar jobs (B). We only use data on workers up to 45 years of age, who have a high school degree, but not a bachelor's. In our estimation we use information from the within-occupation wage distributions, as well as worker flows to and from unemployment and across occupations.

The thought experiment we perform in this exercise is to simulate an increase in the US unemployment rate in 1996, similar to the one that occurred in Europe in the early 1970s. If one believes that the fundamental structure of the two economies is not too different, then our estimated cost will provide a good approximation for the true cost suffered by Europe during those decades.

Table 1			
	$\lambda = 0.36$	$\lambda = 0.12$	$\lambda = 0.36$
	$\delta_W = 0.008, \delta_B = 0.01$	$\delta_W = 0.008, \delta_B = 0.01$	$\delta_W = 0.036, \delta_B = 0.036$
Unemployment Rate	3.6%	10%	10%
Output	10.89	10.06	10.08
Output per worker	11.29	11.17	11.19
Percent Mismatched	31.77%	32.12%	32.03%

Table 1 presents the results of our exercise. Our baseline estimates, in column 1, imply an unemployment rate of approximately 3.6%. Hourly output in this economy is \$10.89 (total population has mass 1), whereas hourly output per *employed* worker is \$11.29. The percentage of workers whose "true", but unobserved, talents do not match perfectly the requirements of their current occupation (mismatched), is 31.77%.

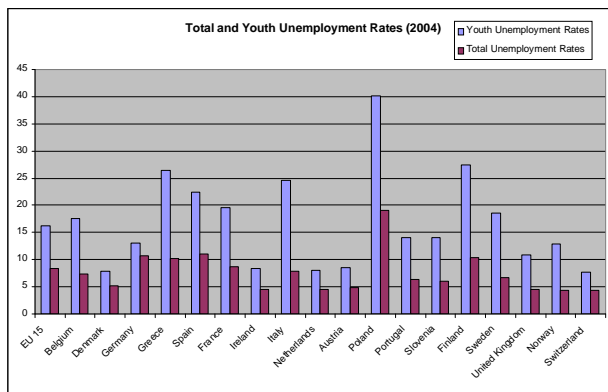
Since, the evidence from European countries shows that most of the increase in unemployment is due to a decrease in the job-finding rate, rather than an increase in the separations rates, we appropriately reduce the unemployed workers' estimated job finding rate, to generate an unemployment level of 10%. Not surprisingly, total output falls substantially, to \$10.06 per hour, mainly because fewer workers are employed and producing output. However, output per employed worker also drops to \$11.17 per hour, approximately 1% lower than before. As noted earlier, because workers now spend more time unemployed, there is

less learning and therefore workers are mismatched more often. Indeed the percentage of mismatched workers has grown to 32.12%, confirming our intuition.

We should point out that this decrease in worker productivity by 1% annually, may imply a sizeable welfare loss. If the increased unemployment rates persist for many years, as they have for many countries, the total amount of lost output can be quite substantial. This cost is in addition to the decline in total output caused by lower utilization of the productive resources.

Some of the increase in the mismatch rate is also the result of the change in workers' behavior. Employed workers who believe they are in an occupation that does not match their talents (i.e. those close to the triggers, \underline{p} and \bar{p}), will now be less willing to quit to unemployment and look for a job in the other occupation, since they are likely to be unemployed for a longer time. To investigate the importance of this mechanism, instead of lowering the job-finding rate, λ , we increase the rates of exogenous separations, δ_W and δ_B in order to generate similar levels of high unemployment. The results, which are presented in the third column, indicate that the magnitude of the loss in output per worker is quantitatively similar. Therefore we conclude that this mechanism, although present, is not as quantitatively important as the decline in learning.

Figure 6



Source, Eurostat

We should note that in this exercise, we have assumed that the rate at which unemployed workers find jobs does not vary across workers. In reality there is certainly heterogeneity in the time it takes workers to locate employment. Perhaps the most striking difference is observed across age groups. As shown in Figure 6, young workers have much larger unemployment rates than older ones. However they are also the ones who benefit from learning the most: since they are the more unsure about their skills compared to more experienced workers, they are more likely to find themselves in the "wrong" occupation. If

the increase in the unemployment rate affects this group harder, then the misallocation of the economy's productive resources will be more severe. We should therefore consider our estimate above as a lower bound on the true productivity cost of high unemployment.

Our results also contribute to the debate on the impact of increased unemployment insurance on resource allocation. Both Acemoglu and Shimer (1999) and Marimon and Zilibotti (1999) have illustrated mechanisms under which an increase in unemployment insurance can lead to improved resource allocation²⁴. If however the increase in unemployment insurance leads to a higher unemployment rate, then the above exercise has shown that it can also lead to a deterioration in the allocation of workers across productive activities. Our results may partly explain why most empirical studies²⁵ have had difficulty detecting the positive impact of unemployment insurance on the match quality that had been previously presumed.

8 Extensions

It is relatively straightforward to extend the model to include three or more occupations. In the three occupation case for example, each worker type is characterized by a three element vector denoting his productivity in each occupation respectively²⁶. All of the main results will go through in this case as well: although each occupation values his skills differently, overall work experience will be an important determinant of the worker's wage. If different worker types have an absolute advantage in different occupations, then all workers will enjoy positive returns to labor market experience in expectation. On the other hand, if one type of worker performs better in all occupations, then again some workers will face negative returns to tenure as their type is revealed to be the low ability one. Of course, search between occupations will still be directed and will depend on the worker's previous employment history. Moreover, occupational mobility will be negatively correlated with employment experience, since workers will eventually self-select to the occupation at which they perform best.

We can also extend the model to include different job finding rates across occupations. This would however add an additional level of complexity to analysis: with different job finding rates, it may be optimal for an unemployed worker to obtain a job in the occupation with a high job finding rate and immediately start looking for a job, while employed, in

²⁴In Acemoglu and Shimer (1999) increased unemployment insurance encourages risk averse workers to apply for higher quality jobs which are more difficult to obtain. Marimon and Zilibotti (1999) show how unemployment insurance can act as a search subsidy that allows workers to obtain more suitable jobs.

²⁵See for example van Ours and Vodopivec (2006) and references therein.

²⁶One must assume that no occupation is dominated by the other two for at least one value of the posterior, so that any point in time all occupations have a positive mass of workers.

another occupation. In that case \hat{p} would still determine when the worker decides to search on the job, but there would be a different threshold that characterizes his search decision if unemployed. If the difference in the job finding rates is small however this behavior will not be optimal, because the cost of searching on the job (decreased search efficiency) will outweigh the benefit and the worker will prefer to look for a job directly in the occupation he is interested in.

The setup introduced here can also be extended to allow for general human capital accumulation. For instance, one can assume that the worker’s productivity in each occupation, as well as his flow value of unemployment, b , increases at some constant rate, as in Bagger et al. (2007). In this case the increase in wages would reflect learning-by-doing, as well as learning about unobserved talents and improved occupational choice.²⁷

9 Conclusion

We have introduced a labor market model where workers learn about themselves and eventually self-select into occupations they perform best. We show that this type of learning has implications on several aspects of labor markets. For instance, it offers an economic interpretation for the decline in the occupational switching probability as workers grow older. Furthermore we propose an explanation for the large bidirectional flows across occupations observed in the data. Our framework provides insights regarding the wage inequality and can also replicate the observed shape of the wage distribution, including its fat right tail. We also show that learning about unobserved abilities can provide an explanation for the increase in the workers’ wages over their lifetime.

We use our setup to investigate whether the interaction of search and learning frictions can lead to further decreases in output. An increase in the unemployment rate by approximately 6.5% is found to reduce output per *employed* worker by approximately 1%, mostly due to the increased misallocation of labor resources across productive activities. These results have important implications regarding policy analysis. For instance, any policy prescription that can alleviate search frictions will also have a positive impact on worker productivity. Furthermore, our results have raised questions regarding the effectiveness of unemployment insurance in improving labor resource allocation. On the contrary, temporary jobs and employment subsidies could provide valuable labor market experience, especially to young workers, thus allowing them to learn their abilities faster and self-select in the occupations they fit best.

Aggregation in our setup is analytically tractable and we are able to provide closed-form

²⁷Occupational mobility, however, would be driven exclusively by the latter type of learning.

expressions for the distribution of workers by expected productivity, as well as worker flows across occupations and to and from unemployment. This greatly facilitates an empirical implementation of the model which can provide important insights regarding the quantitative importance of the model's implications. This task is taken up in a companion paper (Papageorgiou, 2007), where we show that our framework can *quantitatively* account for the within-occupation wage distributions, the relative occupational shares in employment, the decline with age in the probability of occupational switching, as well as a sizeable part of the observed returns to experience.

Appendix

A Wage and Trigger Derivation

As is standard in learning problems, the value functions of the worker and the firm are weakly convex in beliefs p ; for the case of an employed worker for example it is true that $E(V_i(p)) \geq V_i(p) = V_i(E(p))$. The equality holds because from the worker's point of view, beliefs are a martingale (if they were expected to go up or down, the worker would not have fully incorporated all available information when updating his prior); the inequality is true because additional information allows the worker to make decisions that improve his situation. For example, if tomorrow's prior decreases, the worker can start searching on the job and achieve higher utility (revealed preference). It therefore follows from Jensen's inequality that $V(\cdot)$ is a convex function of beliefs. The same argument ensures that $J_i(\cdot)$ and $U(\cdot)$ are also convex.

Furthermore, $J_W(\cdot)$ will be globally increasing in p . To see this, remember that $J_W(p) \geq 0 \forall p$, $J_W(\underline{p}) = 0$ and $J'_W(\underline{p}) = 0$. Since $J_W(\cdot)$ is convex, then it must be increasing everywhere. Intuitively, for \tilde{p} slightly larger than \underline{p} , $J_W(\tilde{p}) > 0$, and thus it must be increasing in that region. Convexity ensures that it is increasing everywhere. Similarly we can show that $J_B(\cdot)$ is decreasing in p .

To see that the threshold \hat{p} is unique, remember that it is defined from $V_W(\hat{p}) - U(\hat{p}) = V_B(\hat{p}) - U(\hat{p})$. However from the Nash bargaining solution we know that $\beta J_i(p) = (1 - \beta)[V_i(p) - U(p)] \Leftrightarrow V_i(p) - U(p) = \frac{\beta}{1-\beta} J_i(p)$. Monotonicity and convexity of $J_i(\cdot)$ and therefore of $V_i(p) - U(p)$, implies that $V_W(p) - U(p)$ and $V_B(p) - U(p)$ cross at most once. The assumption that $\lim_{p \rightarrow 0} V_B(p) > \lim_{p \rightarrow 0} V_W(p)$ and that $\lim_{p \rightarrow 1} V_W(p) > \lim_{p \rightarrow 1} V_B(p)$ (a type worker-type is better in a white-collar job, whereas a blue worker-type is more productive in a blue-collar job), ensures that they cross exactly once.

We will now derive the solution to the value functions and the triggers.

Let's start with the case where the worker's outside option is the occupation he is currently employed in, hence $i = k$.

Now equations (3) through (5) become:

$$\begin{aligned}
 rV_i(p) &= w_i(p) + \frac{1}{2} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1-p)^2 V_i''(p) \\
 &\quad - \delta_i [V_i(p) - U(p)] - \gamma V_i(p)
 \end{aligned} \tag{3'}$$

$$rJ_i(p) = \bar{a}_i(p) - w_i(p) + \frac{1}{2} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1-p)^2 J_i''(p) - \delta_i J_i(p) - \gamma J_i(p) \quad (4')$$

$$rU(p) = b + \lambda [V_i(p) - U(p)] - \gamma U(p) \quad (5')$$

We subtract the worker's value of being unemployed (eq. (5')) from his value of being employed (eq. (3')) and multiply through by $(1 - \beta)$:

$$(1 - \beta) r (V_i(p) - U(p)) = (1 - \beta) (w_i(p) - b) + \frac{1}{2} (1 - \beta) \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 V_i''(p) - \delta_i (1 - \beta) [V_i(p) - U(p)] - \gamma (1 - \beta) V_i(p) - \lambda (1 - \beta) [V_i(p) - U(p)] + \gamma (1 - \beta) U(p)$$

We similarly multiply the asset value of a filled vacancy (eq. (4')) by β :

$$\beta r J_i(p) = \beta \bar{a}_i(p) - \beta w_i(p) + \frac{1}{2} \beta \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 J_i''(p) - \beta \delta_i J_i(p) - \beta \gamma J_i(p)$$

We then subtract the above two equations and using the surplus sharing condition (eq. (2)) we obtain:

$$\begin{aligned} w_i(p) - (1 - \beta) b + \frac{1}{2} (1 - \beta) \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 V_i''(p) \\ - \lambda (1 - \beta) [V_i(p) - U(p)] - \beta \bar{a}_i(p) - \frac{1}{2} \beta \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 J_i''(p) = 0 \Leftrightarrow \\ w_i(p) = \beta \bar{a}_i(p) + (1 - \beta) b + \beta \lambda J_i(p) - \frac{1}{2} (1 - \beta) \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 V_i''(p) \\ + \frac{1}{2} \beta \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 J_i''(p) \end{aligned} \quad (14)$$

Using the surplus sharing condition once again:

$$\begin{aligned}
\beta J_i''(p) &= (1 - \beta)(V_i''(p) - U''(p)) \\
V_i''(p) &= \frac{\beta}{1 - \beta} J_i''(p) + U''(p)
\end{aligned} \tag{15}$$

However from the value of the unemployed worker (eq. (5')), we have:

$$\begin{aligned}
U(p) &= \frac{b}{r + \gamma + \lambda} + \frac{\lambda}{r + \gamma + \lambda} V_i(p) \Rightarrow \\
U''(p) &= \frac{\lambda}{r + \gamma + \lambda} V_i''(p)
\end{aligned}$$

Substituting out for $U''(p)$ in equation (15) results in:

$$\begin{aligned}
V_i''(p) &= \frac{\beta}{1 - \beta} J_i''(p) + \frac{\lambda}{r + \lambda + \gamma} V_i''(p) \Rightarrow \\
V_i''(p) &= \frac{\beta}{1 - \beta} \frac{r + \gamma + \lambda}{r + \gamma} J_i''(p)
\end{aligned} \tag{16}$$

We can now substitute out for $V_i''(p)$ in equation (14) and obtain the wage as a function of $J_i(\cdot)$ only:

$$\begin{aligned}
w_i(p) &= \beta \bar{a}_i(p) + (1 - \beta)b + \beta \lambda J_i(p) - \frac{1}{2} \beta \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 \frac{r + \gamma + \lambda}{r + \gamma} J_i''(p) \\
&\quad + \frac{1}{2} \beta \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 J_i''(p)
\end{aligned}$$

$$\Leftrightarrow w_i(p) = \beta \bar{a}_i(p) + (1 - \beta)b + \beta \lambda J_i(p) - \frac{\beta \lambda}{2(r + \gamma)} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 J_i''(p)$$

The alternative expression for the wage (eq. (6)) can be derived by using equation (14) above and the surplus sharing condition (eq. (2)).

Replacing the wage into the value of the firm function (eq. (4')) produces a differential equation with respect to $J_i(\cdot)$ (eq. (8) in the main text):

$$(r + \delta_i + \gamma + \beta \lambda) J_i(p) = (1 - \beta)(\bar{a}_i(p) - b) + \frac{r + \gamma + \beta \lambda}{2(r + \gamma)} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2 p^2 (1 - p)^2 J_i''(p)$$

The general solution to the above differential equation is:

$$J_i(p) = \frac{(1-\beta)(\bar{a}_i(p) - b)}{r + \gamma + \delta_i + \beta\lambda} + C_1^i p^{\frac{1}{2} - \frac{1}{2}\sqrt{\frac{4+h_i}{h_i}}} (1-p)^{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{4+h_i}{h_i}}} \\ + C_2^i p^{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{4+h_i}{h_i}}} (1-p)^{\frac{1}{2} - \frac{1}{2}\sqrt{\frac{4+h_i}{h_i}}}$$

where $h_i = \frac{1}{2} \frac{r+\gamma+\beta\lambda}{(r+\gamma+\delta_i+\beta\lambda)(r+\gamma)} \left(\frac{a_i^w - a_i^b}{\sigma_i} \right)^2$ and C_1^i and C_2^i are undetermined coefficients. For the case of $i = W$, when $p \rightarrow 1$, $\lim_{p \rightarrow 1} C_2^i p^{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{4+h_i}{h_i}}} (1-p)^{\frac{1}{2} - \frac{1}{2}\sqrt{\frac{4+h_i}{h_i}}} = C_2^i \cdot 1 \cdot \lim_{p \rightarrow 1} (1-p)^{\frac{1}{2} - \frac{1}{2}\sqrt{\frac{4+h_i}{h_i}}} = \infty$ because since $h_i > 0$, then $\sqrt{\frac{4+h_i}{h_i}} > 1$, so $\frac{1}{2} \left(1 - \sqrt{\frac{4+h_i}{h_i}} \right) < 0$. However since the profits of the firm are bounded from above by the total value of the surplus when the worker is known to be type white which is finite, it must be the case that $C_2^W = 0$. A similar argument for $i = B$ and $p \rightarrow 0$ leads to $C_1^B = 0$.

Therefore for $p \geq \hat{p}$ for the case of occupation W , we obtain equation (10) in the main text and for $p \leq \hat{p}$ and occupation B we are left with equation (11).

The case where the worker's occupation of choice if unemployed differs from his current one ($i \neq k$) is a somewhat more involved and we analyze it the technical appendix to this paper. The resulting expressions for the asset value of the firm are the following.

For occupation W and $p < \hat{p}$:

$$J_W(p) = \frac{r + \gamma + \delta_W + \eta_W \lambda}{\left(\frac{a_W^w - a_W^b}{\sigma_W} \right)^2} p^{\frac{1}{2}} (1-p)^{\frac{1}{2}} \cdot \int_p^1 \left[\theta_{WB} \tau + \kappa_{WB} C_2^B \tau^{\xi_{WB}} (1-\tau)^{1-\xi_{WB}} + c_{WB} \right] (\tau(1-\tau))^{-\frac{3}{2}} \\ \cdot \left(\left(\frac{1-\tau}{\tau} \frac{p}{1-p} \right)^{l_{WB}} - \left(\frac{\tau}{1-\tau} \frac{1-p}{p} \right)^{l_{WB}} \right) d\tau \quad (17)$$

$$\text{where } l_{WB} = \sqrt{\frac{1}{4} + \frac{2(r+\gamma+\delta_W+\eta_W\lambda)}{\left(\frac{a_W^w - a_W^b}{\sigma_W} \right)^2}}, \quad c_{WB} = \frac{1-\beta}{r+\gamma+\delta_W+\eta_W\lambda} \left(\beta\lambda(1-\eta_W) \frac{(a_B^b - b)}{r+\gamma+\delta_B+\beta\lambda} - (a_W^b - b) \right),$$

$$\theta_{WB} = \frac{1-\beta}{r+\gamma+\delta_W+\eta_W\lambda} \left(\beta\lambda(1-\eta_W) \frac{a_B^w - a_B^b}{r+\gamma+\delta_B+\beta\lambda} - (a_W^w - a_W^b) \right),$$

$$\kappa_{WB} = -\frac{\beta\lambda}{r+\gamma+\delta_W+\eta_W\lambda} \left(\frac{r+\gamma+\delta_B+\beta\lambda}{r+\gamma+\beta\lambda} \left(\frac{a_B^w - a_B^b}{a_B^w - a_B^b} \frac{\sigma_B}{\sigma_W} \right)^2 - (1-\eta_W) \right) \text{ and } \xi_{WB} = \frac{1}{2} + \frac{1}{2}\sqrt{\frac{4+h_B}{h_B}}.$$

Similarly, for the case of occupation B and $p > \hat{p}$:

$$J_B(p) = \frac{r + \gamma + \delta_B + \eta_B \lambda}{\left(\frac{a_B^w - a_B^b}{\sigma_B}\right)^2} p^{\frac{1}{2}} (1-p)^{\frac{1}{2}} \cdot \int_{\bar{p}}^p \left[\theta_{BW} \tau + \kappa_{BW} C_1^W \tau^{\xi_{BW}} (1-\tau)^{1-\xi_{BW}} + c_{BW} \right] (\tau(1-\tau))^{-\frac{3}{2}} \cdot \left(\left(\frac{1-\tau}{\tau} \frac{p}{1-p} \right)^{l_{BW}} - \left(\frac{\tau}{1-\tau} \frac{1-p}{p} \right)^{l_{BW}} \right) d\tau \quad (18)$$

where $l_{BW} = \sqrt{\frac{1}{4} + \frac{2(r+\gamma+\delta_B+\eta_B\lambda)}{\left(\frac{a_B^w - a_B^b}{\sigma_B}\right)^2}}$ and $c_{BW} = \frac{1-\beta}{r+\gamma+\delta_B+\eta_B\lambda} \left(\beta\lambda(1-\eta_B) \frac{(a_W^b - b)}{r+\gamma+\delta_W+\beta\lambda} - (a_B^b - b) \right)$,
 $\theta_{BW} = \frac{1-\beta}{r+\gamma+\delta_B+\eta_B\lambda} \left(\beta\lambda(1-\eta_B) \frac{a_W^w - a_B^b}{r+\gamma+\delta_W+\beta\lambda} - (a_B^w - a_B^b) \right)$,
 $\kappa_{BW} = -\frac{\beta\lambda}{r+\gamma+\delta_B+\eta_B\lambda} \left(\frac{r+\gamma+\delta_W+\beta\lambda}{r+\gamma+\beta\lambda} \left(\frac{a_B^w - a_B^b}{a_W^w - a_W^b} \frac{\sigma_W}{\sigma_B} \right)^2 - (1-\eta_B) \right)$ and $\xi_{BW} = \frac{1}{2} - \frac{1}{2} \sqrt{\frac{4+h_W}{h_W}}$.

Finally, to pin down the values of \underline{p} , \bar{p} , \hat{p} , C_1^W and C_2^B we will use the 5 conditions mentioned in the main text. Straightforward derivations imply that they can be rewritten as follows:

$$\lim_{p \rightarrow \hat{p}^-} J_W(p) = \lim_{p \rightarrow \hat{p}^+} J_W(p) \quad (19)$$

$$\lim_{p \rightarrow \hat{p}^-} J_B(p) = \lim_{p \rightarrow \hat{p}^+} J_B(p) \quad (20)$$

$$J_W(\hat{p}) = J_B(\hat{p}) \quad (21)$$

$$\lim_{p \rightarrow \hat{p}^-} (V'_W(p) + J'_W(p)) = \lim_{p \rightarrow \hat{p}^+} (V'_W(p) + J'_W(p)) \quad (22)$$

$$\lim_{p \rightarrow \hat{p}^-} (V'_B(p) + J'_B(p)) = \lim_{p \rightarrow \hat{p}^+} (V'_B(p) + J'_B(p)) \quad (23)$$

B Endogenizing the Job Finding Rate

In order to close the model we need to endogenize the worker's job finding rate λ_i . We assume that in each occupation i , matches are randomly formed according to an increasing, concave and homogenous of degree one matching function $m(p_i, v_i)$, where p_i is the mass of workers (employed or unemployed) petitioning for a job in occupation i and v_i is the mass of occupation i vacancies. Unemployed workers will be matched with firms at rate $\lambda_i = \frac{m(p_i, v_i)}{p_i}$. Each firm can post a vacancy at flow cost cv_i , which earns no revenue while empty.

Since the job searching is less efficient, the effective mass of workers searching for a job in each occupation is equal to:

$$\begin{aligned}
p_W &= \int_{\hat{p}}^1 z_W(p) dp + \eta_B \int_{\hat{p}}^{\bar{p}} f_B(p) dp \\
p_B &= \int_0^{\hat{p}} z_B(p) dp + \eta_W \int_{\underline{p}}^{\hat{p}} f_W(p) dp
\end{aligned}$$

A firm in occupation W contacts a worker at rate $\frac{m(p_W, v_W)}{v_W}$, so the value of an unfilled vacancy in occupation W , P_W will equal:

$$P_W = -cv_W + \frac{m(p_W, v_W)}{v_W} \int_{\hat{p}}^1 J_{WW}(p) z_W(p) dp + \frac{m(p_W, v_W)}{v_W} \eta_B \int_{\hat{p}}^{\bar{p}} J_{WW}(p) f_B(p) dp$$

Therefore the asset value of a vacancy is equal to its flow cost plus the expected value of filled posting, when it contacts a worker. Note that because on the job search has a different effectiveness, the effective mass of potential workers is η_i times the mass of the employed searchers.

Similarly for occupation B :

$$P_B = -cv_B + \frac{m(p_B, v_B)}{v_B} \int_0^{\hat{p}} J_{BB}(p) z_B(p) dp + \frac{m(p_B, v_B)}{v_B} \eta_W \int_{\underline{p}}^{\hat{p}} J_{BB}(p) f_W(p) dp$$

Free firm entry and exit ensures that $P_i = 0$ and therefore:

$$cv_W = \frac{m(p_W, v_W)}{v_W} \int_{\hat{p}}^1 J_{WW}(p) z_W(p) dp + \frac{m(p_W, v_W)}{v_W} \eta_B \int_{\hat{p}}^{\bar{p}} J_{WW}(p) f_B(p) dp$$

and:

$$cv_B = \frac{m(p_B, v_B)}{v_B} \int_0^{\hat{p}} J_{BB}(p) z_B(p) dp + \frac{m(p_B, v_B)}{v_B} \eta_W \int_{\underline{p}}^{\hat{p}} J_{BB}(p) f_W(p) dp$$

The above two conditions pin down the equilibrium mass of unfilled vacancies in each occupation and therefore the workers' job finding rate, thus closing the model.

Note that in the derivations of the preceding sections we assumed that $\lambda_W = \lambda_B = \lambda$ to simplify the steady state analysis; however this assumption should not be very restrictive, since worker behavior does not diverge much for small differences in the job finding rates. A possible extension of the model with different job finding rates across occupations is discussed

in Section 8.

C Aggregation

The Kolmogorov forward in occupation B , for every $p \leq \bar{p}$ and $p \neq \underline{p}$ is given by:

$$0 = \frac{df_B(p)}{dt} = \frac{d^2}{dp^2} \left[\frac{1}{2} \left(\frac{a_B^w - a_B^b}{\sigma_B} \right)^2 p^2 (1-p)^2 f_B(p) \right] - \delta_B f_B(p) - \gamma f_B(p) + \lambda z_B(p) - \eta_B \lambda f_B(p) I \{ \hat{p} \leq p \} + \eta_W \lambda f_W(p) I \{ \underline{p} \leq p < \hat{p} \} \quad (24)$$

Furthermore, the corresponding condition (13) for unemployed workers in occupation B , is that for every $p < \hat{p}$ and $p \neq \underline{p}$:

$$\delta_W f_W(p) + \delta_B f_B(p) + \gamma g(p) = \lambda z_B(p) + \gamma z_B(p) \quad (25)$$

We then use eq.(13) and (25) to solve out for $z_W(p)$ and $z_B(p)$ respectively and after substituting them into the two forward equations (12) and (24), we derive the following system of differential equations:

$$\begin{aligned} & \frac{d^2}{dp^2} \left[\frac{1}{2} \left(\frac{a_W^w - a_W^b}{\sigma_W} \right)^2 p^2 (1-p)^2 f_W(p) \right] - \gamma \frac{\delta_W + \lambda + \gamma}{\lambda + \gamma} f_W(p) + \frac{\lambda \delta_B}{\lambda + \gamma} f_B(p) \\ & + \frac{\gamma \lambda}{\lambda + \gamma} g(p) - \eta_W \lambda f_W(p) I \{ p < \hat{p} \} + \eta_B \lambda f_B(p) I \{ \hat{p} \leq p \leq \bar{p} \} = 0 \end{aligned}$$

$$\begin{aligned} & \frac{d^2}{dp^2} \left[\frac{1}{2} \left(\frac{a_B^w - a_B^b}{\sigma_B} \right)^2 p^2 (1-p)^2 f_B(p) \right] - \gamma \frac{\delta_B + \lambda + \gamma}{\lambda + \gamma} f_B(p) + \frac{\lambda \delta_W}{\lambda + \gamma} f_W(p) \\ & + \frac{\gamma \lambda}{\lambda + \gamma} g(p) - \eta_B \lambda f_B(p) I \{ \hat{p} \leq p \} + \eta_W \lambda f_W(p) I \{ \underline{p} \leq p < \hat{p} \} = 0 \end{aligned}$$

The solution to the differential equations contains 12 undetermined coefficients, so we need 12 conditions to pin them down. We go over those conditions in the technical appendix to the present paper²⁸.

The resulting expression for workers employed in occupation B is the following:

²⁸Located at <http://pantheon.yale.edu/~tp222>

In the case where $p \in [0, \underline{p}]$:

$$\begin{aligned}
f_B(p) &= C_2^B p^{-1-q_B} (1-p)^{q_B-2} \\
&\quad - \frac{d_B}{\sqrt{c_B(4+c_B)}} p^{q_B-2} (1-p)^{-1-q_B} \int_0^p \tau^{a-q_B} (1-\tau)^{q_B+\psi-1} d\tau \\
&\quad + \frac{d_B}{\sqrt{c_B(4+c_B)}} p^{-1-q_B} (1-p)^{q_B-2} \int_0^p \tau^{q_B+a-1} (1-\tau)^{\psi-q_B} d\tau
\end{aligned} \tag{26}$$

for $p \in (\underline{p}, \widehat{p})$:

$$\begin{aligned}
f_B(p) &= C_3^B p^{q_B-2} (1-p)^{-1-q_B} + C_4^B p^{-1-q_B} (1-p)^{q_B-2} \\
&\quad + \frac{1}{\sqrt{c_B(4+c_B)}} p^{q_B-2} (1-p)^{-1-q_B} [C_6^W m_B \int_p^{\widehat{p}} \tau^{-\frac{1}{2}-\kappa_B-q_B} (1-\tau)^{-\frac{3}{2}+\kappa_B+q_B} d\tau \\
&\quad - C_6^W n_B \int_p^{\widehat{p}} \tau^{-\frac{1}{2}+\kappa_B-q_B} (1-\tau)^{-\frac{3}{2}-\kappa_B+q_B} d\tau + d_B \int_p^{\widehat{p}} \tau^{a-q_B} (1-\tau)^{\psi+q_B-1} d\tau] \\
&\quad - \frac{1}{\sqrt{c_B(4+c_B)}} p^{-1-q_B} (1-p)^{q_B-2} [C_6^W m_B \int_p^{\widehat{p}} \tau^{-\frac{3}{2}-\kappa_B+q_B} (1-\tau)^{-\frac{1}{2}+\kappa_B-q_B} d\tau \\
&\quad - C_6^W n_B \int_p^{\widehat{p}} \tau^{-\frac{3}{2}+\kappa_B+q_B} (1-\tau)^{-\frac{1}{2}-\kappa_B-q_B} d\tau + d_B \int_p^{\widehat{p}} \tau^{a+q_B-1} (1-\tau)^{\psi-q_B} d\tau]
\end{aligned}$$

and for $p \in [\widehat{p}, \bar{p}]$:

$$f_B(p) = C_5^B (p(1-p))^{-\frac{3}{2}} \left(\left(\frac{1-p}{p} \right)^{\kappa_W} - \left(\frac{1-\bar{p}}{\bar{p}} \right)^{2\kappa_W} \left(\frac{p}{1-p} \right)^{\kappa_W} \right)$$

where $C_2^B, C_3^B, C_4^B, C_5^B$ and C_6^W are undetermined coefficients and $c_B = \frac{\left(\frac{a_B^w - a_B^b}{\sigma_B}\right)^2 (\lambda + \gamma)}{2\gamma(\delta_B + \lambda + \gamma)}$,
 $q_B = \frac{1}{2} - \frac{1}{2} \sqrt{\frac{4+c_B}{c_B}}$, $d_B = -\frac{\lambda}{(\delta_B + \lambda + \gamma)B(a, \psi)}$, $\kappa_W = \sqrt{\frac{1}{4} + \frac{2(\delta_B + \gamma + \eta_B \lambda)}{\left(\frac{a_B^w - a_B^b}{\sigma_B}\right)^2}}$, $\kappa_B = \sqrt{\frac{1}{4} + \frac{2(\delta_W + \gamma + \eta_W \lambda)}{\left(\frac{a_W^w - a_W^b}{\sigma_W}\right)^2}}$,
 $n_W = -\lambda \frac{\delta_B + \eta_B \lambda + \eta_B \gamma}{\gamma(\delta_W + \lambda + \gamma)}$, $n_B = -\lambda \frac{\delta_W + \eta_W \lambda + \eta_W \gamma}{\gamma(\delta_B + \lambda + \gamma)}$, $m_W = -n_W \left(\frac{1-\bar{p}}{\bar{p}}\right)^{2\kappa_W}$, $m_B = -n_B \left(\frac{p}{1-\bar{p}}\right)^{2\kappa_B}$.

A similar expression give us the mass of workers at every value of the posterior in occupation B . In our solution we eventually end up with 8 undetermined coefficients, which are given by the solution of a linear system of 8 equations and 8 unknowns. See technical appendix for details.

From the first integral in right hand side of equation (26), a necessary condition for $f_B(p)$ not to go to infinity, as p goes to zero is that $q_B + a - 1 > 0$.

We are interested in the behavior of $f_B(p)$ as p approaches 0.

We first show that as p approaches 0, the density of workers employed in occupation B is captured by the first term in equation (26), while the other two become quantitatively insignificant.

Let's transform the integrals in equation (26). Let $\tau = ps$. Then we can rewrite the first integral in the equation (26) above as:

$$\begin{aligned} \int_0^p \tau^{a-q_B} (1-\tau)^{q_B+\psi-1} d\tau &= \int_0^1 (ps)^{a-q_B} (1-ps)^{q_B+\psi-1} p ds \\ &= p^{1+a-q_B} \int_0^1 s^{a-q_B} (1-ps)^{q_B+\psi-1} ds \end{aligned}$$

Similarly the second integral can be rewritten as:

$$\int_0^p \tau^{q_B+a-1} (1-\tau)^{\psi-q_B} d\tau = p^{q_B+a} \int_0^1 s^{q_B+a-1} (1-ps)^{\psi-q_B} ds$$

Thus we can now rewrite equation (26) as:

$$\begin{aligned} f_B(p) &= C_2^B p^{-1-q_B} (1-p)^{q_B-2} \\ &\quad - \frac{d}{\sqrt{c_B(4+c_B)}} p^{a-1} (1-p)^{-1-q_B} \int_0^1 s^{a-q_B} (1-ps)^{q_B+\psi-1} ds \\ &\quad + \frac{d}{\sqrt{c_B(4+c_B)}} p^{a-1} (1-p)^{q_B-2} \int_0^1 s^{q_B+a-1} (1-ps)^{\psi-q_B} ds \end{aligned}$$

Clearly, as p approaches zero, all three terms will eventually converge to zero, since they all have a term of the form p^θ , for some $\theta > 0$. The question is which one converges the slowest, p^{-1-q_B} or p^{a-1} . However we already have a restriction that $q_B + a - 1 > 0 \Leftrightarrow a - 1 > -q_B$. Thus it must be true that $a - 1 > -1 - q_B$ and therefore only the first term will be quantitatively important as p goes to zero.

Moreover we have that:

$$\frac{d(C_2^B p^{-1-q_B} (1-p)^{q_B-2})}{dp} = C_2^B p^{-2-q_B} (1-p)^{q_B-3} (-1 - q_B + 3p)$$

Since $f_B(p) \geq 0, \forall p$, then $C_2^B > 0$. Furthermore, if $\gamma > \left(\frac{a_B^w - a_B^b}{\sigma_B}\right)^2 \frac{\lambda + \gamma}{\delta_B + \lambda + \gamma}$, $f_B(p)$ will be decreasing as p goes to zero. In that case, the distribution of posterior will feature a fat Pareto type tail²⁹. Otherwise, it will be increasing as p goes to zero.

²⁹This is true because only the first term is quantitatively important as p goes to zero.

A similar derivation shows that $f_W(p)$ will feature a fat right tail, as long $\gamma > \left(\frac{a_W^w - a_W^b}{\sigma_W}\right)^2 \frac{\lambda + \gamma}{\delta_W + \lambda + \gamma}$. Furthermore the wage function is given by:

$$w_B(p) = \beta \bar{a}_B(p) + (1 - \beta) b - \frac{1}{2} \left(\frac{a_B^w - a_B^b}{\sigma_B}\right)^2 p^2 (1 - p)^2 (1 - \beta) W_B''(p) \\ + \frac{1}{2} \left(\frac{a_B^w - a_B^b}{\sigma_B}\right)^2 p^2 (1 - p)^2 \beta J_B''(p) + \beta \lambda J_B(p)$$

where:

$$J_{BB}(p) = \frac{(1 - \beta)(\bar{a}_B(p) - b)}{r + \gamma + \delta_B + \beta \lambda} + C_2^{BB} p^{\frac{1}{2}} \left(1 + \sqrt{\frac{4 + q_{BB}}{q_{BB}}}\right) (1 - p)^{\frac{1}{2}} \left(1 - \sqrt{\frac{4 + q_{BB}}{q_{BB}}}\right)$$

As p approaches zero, we drop all higher terms and we are left with

$$w_B(p) = p \left[\beta (a_B^w - a_B^b) + \beta \lambda (1 - \beta) \frac{a_B^w + a_B^b}{r + \gamma + \delta_B + \beta \lambda} \right] + \beta \lambda (1 - \beta) \frac{a_B^b - b}{r + \gamma + \delta_B + \beta \lambda} + \beta a_B^b + (1 - \beta) b$$

For the case where $\beta (a_B^w - a_B^b) + \beta \lambda (1 - \beta) \frac{a_B^w + a_B^b}{r + \gamma + \delta_B + \beta \lambda} = 1$, we can write the wage function as $w_B(p) = p + c_0$ where $c_0 = \beta \lambda (1 - \beta) \frac{a_B^b - b}{r + \gamma + \delta_B + \beta \lambda} + \beta a_B^b + (1 - \beta) b$. In this case the density of wages $f^w(w) = f(w - c_0)$, will also feature a fat *left* tail of the Pareto type. Alternatively in the case where $\beta (a_B^w - a_B^b) + \beta \lambda (1 - \beta) \frac{a_B^w + a_B^b}{r + \gamma + \delta_B + \beta \lambda} = -1$ the wage distribution in occupation B , will feature a fat *right* tail of the Pareto type. Notice that this has an intuitive economic explanation. In both cases few workers are expected to reach the vicinity of zero (as long as $\gamma > \left(\frac{a_B^w - a_B^b}{\sigma_B}\right)^2 \frac{\lambda + \gamma}{\delta_B + \lambda + \gamma}$). The first case is the high ability, low ability workers case ($a_B^w - a_B^b = \frac{r + \gamma + \delta_B + \beta \lambda}{\beta(r + \gamma + \delta_B + \beta \lambda + \lambda(1 - \beta))} > 0 \Rightarrow a_B^w > a_B^b$). Now these workers will receive the lowest wages in occupation B , since they are most likely to be the low-ability ones. Hence the within occupation wage distribution will feature a fat left tail!

In the alternative case ($a_B^w - a_B^b = -\frac{r + \gamma + \delta_B + \beta \lambda}{\beta(r + \gamma + \delta_B + \beta \lambda + \lambda(1 - \beta))} < 0 \Rightarrow a_B^w < a_B^b$), these workers will receive the highest wages in the occupation as they are recognized to be the best. Now the within occupation wage distribution will feature a fat right tail, consistent with empirical observations.

Similarly the wage distribution in occupation W will feature a fat right tail if $a_W^w - a_W^b = \frac{r + \gamma + \delta_W + \beta \lambda}{\beta(r + \gamma + \delta_W + \beta \lambda + \lambda(1 - \beta))}$.

References

- [1] Acemoglu, D. and R. Shimer (1999): “Efficient Unemployment Insurance,” *Journal of Political Economy*, 107(5), 893-928.
- [2] Abraham, K. and H. S. Farber (1987): “Job Duration, Seniority, and Earnings,” *American Economic Review*, 77(3), 278-297.
- [3] Altonji, J. G. (2005): “Employer Learning, Statistical Discrimination and Occupational Attainment,” *American Economic Review*, 95(2), 112-117.
- [4] Altonji, J. G. and C. Pierret (2001): “Employer Learning and Statistical Discrimination,” *Quarterly Journal of Economics*, 116(1), 313-350.
- [5] Altonji J. G. and R. J. Shaker (1987): “Do Wages Rise with Job Seniority?,” *The Review of Economic Studies*, 54(3), 437-459.
- [6] Altonji J. G. and N. Williams (2005): “Do Wages Rise with Job Seniority?: A Reassessment,” *Industrial and Labor Relations Review*, 58(3), 370-397.
- [7] Antonovics K. and L. Golan (2007): “Experimentation and Job Choice,” Mimeo, Carnegie Mellon University, Tepper School of Business.
- [8] Bagger, J., F. Fontaine, F. Postel-Vinay and J.-M. Robin (2007): “A Tractable Equilibrium Search Model of Individual Wage Dynamics with Experience Accumulation,” Mimeo, University of Bristol, Department of Economics.
- [9] Blanchard, O. (2006): “European Unemployment: The Evolution of Facts and Ideas,” *Economic Policy*, 21(45), 5-59.
- [10] Burdett K. and D. T. Mortensen (1998): “Wage Differentials, Employer Size and Unemployment,” *International Economic Review*, 39(2), 257-273.
- [11] Eeckhout, J. (2006): “Employer Learning and General Human Capital,” Mimeo, University of Pennsylvania, Department of Economics.
- [12] Farber, H. S. and R. Gibbons (1996): “Learning and Wage Dynamics,” *Quarterly Journal of Economics*, 111(4), 1007-1047.
- [13] Gathmann, C. and U. Schönberg (2006): “How General Is Specific Human Capital,” IZA Discussion Paper No. 2485, Bonn, Germany.

- [14] Holmstrom B. (1999): “Managerial Incentive Problems: A Dynamic Perspective,” *The Review of Economic Studies*, 66(1), 169-182.
- [15] Jovanovic, B. (1979): “Job Matching and the Theory of Turnover,” *Journal of Political Economy*, 87(5) Part 1, 972-990.
- [16] Jovanovic, B. (1984): “Matching, Turnover and Unemployment,” *Journal of Political Economy*, 92(1), 108-122.
- [17] Jovanovic, B. and R. Moffitt (1990): “An Estimate of a Sectoral Model of Labor Mobility,” *Journal of Political Economy*, 98(4), 827-852.
- [18] Kambourov, G. and I. Manovskii (2006): “Rising Occupational and Industry Mobility in the United States: 1968-1997,” *International Economic Review*, forthcoming.
- [19] Kambourov, G. and I. Manovskii (2007a): “Occupational Specificity of Human Capital,” *International Economic Review*, forthcoming.
- [20] Kambourov, G. and I. Manovskii (2007b): “Occupational Mobility and Wage Inequality,” Mimeo, University of Pennsylvania, Department of Economics.
- [21] Karlin, S. and H. M. Taylor (1981): *A Second Course in Stochastic Processes* (New York: Academic Press).
- [22] Liptser, R. and A. Shyryaev (1977): *Statistics of Random Processes*, Vol. 2 Berlin: Springer-Verlag.
- [23] MacDonald, G. M. (1982): “A Market Equilibrium Theory of Job Assignment and Sequential Accumulation of Information,” *The American Economic Review*, 72(5), 1038-1055.
- [24] Marimon R. and F. Zilibotti (1999): “Unemployment vs. Mismatch of Talents: Reconsidering Unemployment Benefits,” *The Economic Journal*, 109(455), 266-291.
- [25] McCall, B. (1990): “Occupational Matching: A Test of Sorts,” *Journal of Political Economy*, 98(1), 45-69.
- [26] Miller, R. A. (1984): “Job Matching and Occupational Choice,” *Journal of Political Economy*, 92(6), 1086-1120.
- [27] Moscarini, G. (2005): “Job Matching and the Wage Distribution,” *Econometrica*, 73(2), 481-516.

- [28] Murphy, K. and R. Topel (1987): “The Evolution of Unemployment in the United States: 1968-1985,” in S. Fischer (ed.), *NBER Macroeconomic Annual 1987* (Cambridge: MIT Press).
- [29] Neal, D. (1999): “The Complexity of Job Mobility among Young Men,” *Journal of Labor Economics*, 17(2), 237-261.
- [30] Papageorgiou, T. (2007): “Reinterpreting the Returns to Labor Market Experience,” Mimeo, Yale University, Department of Economics.
- [31] Pinkston, J. C. (2006): “A Model of Asymmetric Employer Learning With Testable Implications,” Bureau of Labor Statistics Working Paper 390, Washington, D.C.
- [32] Roy, A. D. (1951): “Some Thought on the Distribution of Earnings,” *Oxford Economics Papers*, 3, 135-146.
- [33] Schönberg, U. (2007): “Testing for Asymmetric Employer Learning,” *Journal of Labor Economics*, forthcoming.
- [34] Shimer, R. (2006): “On-the-Job Search and Strategic Bargaining,” *European Economic Review*, 50(4), 811-830.
- [35] Sicherman, N. and O. Galor (1990): “A Theory of Career Mobility,” *Journal of Political Economy*, 98(1), 169-192.
- [36] Topel, R. (1990): “Specific Capital, Mobility and Wages: Wages Rise with Job Seniority,” *Journal of Political Economy*, 99(1), 145-176.
- [37] van Ours, J. C. and M. Vodopivec (2006): “Duration of Unemployment Benefits and Quality of Post-Unemployment Jobs: Evidence from a Natural Experiment,” World Bank Policy Research Working Paper 4031, The World Bank, Washington, DC.