Learning Your Comparative Advantage

Theodore Papageorgiou*

Yale University, Department of Economics†

June 2007 [in progress, do not quote]

Abstract

This paper introduces a labor market model where workers learn about their general human capital, rather than the quality of their current match. Different worker types perform differently across occupations, but neither the worker, nor the market, have full information about the worker’s type. Output realizations reveal information about the worker’s productivity and therefore her type. Wage and occupational mobility decisions depend on both the worker’s comparative advantage and differences in the speed of learning across occupations. The model has different implications than the standard labor market learning model (Jovanovic (1979)): human capital is occupation-specific, since each occupation puts a different price on the worker’s skills, but tenure in other occupations will also be important in wage determination. Returns to overall employment experience capture the selection of workers sorting through occupations and learning about their general human capital. Moreover occupational choice is no longer random, since previous work experience reveals information about the worker’s fit in other occupations as well. In this framework, an increase in unemployment benefits may result in matches of lower quality.

1 Introduction

This paper introduces a labor market model with learning and search frictions, where workers learn about their type, rather than the quality of their match as they sort through different

---

*I thank Giuseppe Moscarini for his invaluable advice and guidance throughout this project, as well as Joe Altonji and Bjorn Brügemann for their suggestions and help. In addition, I am grateful to Daron Acemoglu, Bill Brainard, Ken Burdett, Ricardo Caballero, Eduardo Engel, Fatih Guvenen, Nicholas Kalouptsidis, Per Krusell, Fabian Lange, Iourii Manovskii, Rafael Melo, Fabien Postel-Vinay, Aleh Tsyvinski and Randy Wright for useful comments and discussions.

†PO Box 208268, New Haven, CT 06520-8268, USA. Phone: (203) 432-3595. E-mail: theodore.papageorgiou@yale.edu
occupations. This assumption turns out to have different implications about the importance of labor market experience and occupational choice compared to the standard labor market learning model (Jovanovic (1979 and 1984)). In particular, overall employment experience, will be an important determinant of wages, capturing the self-selection of workers into occupations in which they are better matched, while occupational choice is no longer random but based on workers’ previous labor market performance. Furthermore, the present framework provides new insights on the impact of unemployment insurance on mismatch.

One can distinguish two types of human capital in the literature: specific and more general human capital; the former can be firm, occupation or industry specific, whereas the latter transcends these boundaries and stays with the worker as he switches over to new firms/occupations/industries. Specific human capital implies returns to (occupational, firm or industry) tenure, whereas general human capital underlines the importance of returns to labor market experience. Both forms of human capital can be explained by either learning by doing or by selection\(^1\). Regarding occupation specific capital, Kambourov and Manovskii present both theoretical (2004) and empirical evidence (2005) in favor of the hypothesis of learning by doing, whereas the case for selection has been famously been made by Jovanovic (1979) and more recently by Moscarini (2005). Furthermore, the results of Altonji and Pierret (2001), who find that firms learn about their workers’ unobserved characteristics over time, provides evidence on the importance of selection. General human capital investment through learning by doing has been advocated since Becker’s (1964) influential book on the matter. The present paper makes the case for selection in a setting with general human capital, by introducing a model that can explain returns to experience with no learning by doing. Unlike the Jovanovic model, the worker doesn’t disregard any knowledge he has acquired upon switching occupations or becoming unemployed. Each worker enters the labor market with a belief regarding his type, which will evolve throughout his working life until his retirement, despite firm or occupational switches. The model captures returns to experience, while accounting for other facts that have been documented in the labor market, such as the declining occupational switching hazard, the shape of the wage distribution and offsetting flows between occupations.

More specifically, in this economy, every worker is assigned a type at birth which remains 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 latter’s type. Output realizations reveal information about the worker’s underlying productivity

---

\(^1\)Search models also provide an explanation for returns to tenure and (potentially) returns to experience. For example Burdett-Coles (2003) present a model where the firm offers the worker a wage-tenure contract, guaranteeing him increasing wage over time.
and therefore his type, which is immediately reflected in his wage and occupational mobility decisions. Therefore, while employed, workers learn, not only about their fit in their current occupation, but in other occupations as well. Furthermore, imperfect information about the worker’s type implies that workers sort across occupations according to comparative rather than absolute advantages: the worker’s expected productivity in one occupation compared to others will largely determine the worker’s search decision. The worker however also takes into account differences in the speed of learning across occupations in his decision; in fact, his preferred occupation may not be the one in which he is expected to be most productive. Firms and workers negotiate on splitting the 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. Stochastic retirement shocks prevent the worker from perfectly finding out his type.

In our setup, returns to human capital are occupation-specific, since different occupations value the worker’s skills differently, but his tenure in other occupations is also an important determinant of his wage and the hazard rate of occupational switching. Unlike other models where previous experience in other occupations is not important (e.g. Kambourov and Manovskii (2004)), in the proposed environment, overall labor market experience matters: learning about general human capital occurs in all occupations (perhaps at different rates), but there is an occupation-specific price for the worker’s skills. As we will show, in the case where no worker type performs better in both occupations, workers with more employment experience receive higher wages in expectation, and are less likely to switch occupations either by on the job search or by quitting to unemployment. Returns to overall labor market experience here capture the selection effect of workers sorting through occupations and learning about their general human capital.

Furthermore, unlike in the standard learning model (Jovanovic (1979)), search here is directed between occupations. Since learning involves a worker’s type rather than his match, the information the worker acquires while working in one occupation is relevant for other occupations as well. Her current performance shapes expectations about his productivity in other occupations, so if he decides to switch he will direct his search efforts to the one that fits him best. For example, it is reasonable to expect that a worker who was previously employed as a sales representative and decided to switch because of poor communication skills, probably would not choose (or be hired in) a public relations position. Alternatively if he discovered that he excelled in getting the message across, he might chose to pursue a career in advertising.

The notion of occupational specific human capital conforms with our intuition: there is no reason for a lawyer to be less effective in his practice if he switches law firms or for a
chef to be less creative if he works in a different restaurant. Several authors have presented
evidence of the occupational specificity of human capital: McCall (1990) shows that tenure
in a worker’s first job has a negative effect on the hazard rate of separation of his second job
if no occupational switch has taken place. More recently Kambourov and Manovskii (2005)
document returns to occupational tenure.

The idea of learning about the worker’s general human capital rather than match quality
is not new in the literature. Altonji (2005) presents a model in which the market is slow to
learn about the worker’s productivity if he starts out in low-skill level jobs. In his model
higher skilled jobs also exhibit a higher speed of learning for the worker. Along the same
lines, Gibbons et al. (2005) present a model with learning frictions where different sectors
value differently workers’ ability levels. They estimate their model and find evidence of
sorting across sectors: high-wage sectors employ high-skill workers and offer high returns to
skill. Eeckhout (2006) also introduces a model of learning about the worker’s ability, where
firms have asymmetric information and engage in a second price auction that determines the
worker’s wage.

All the above papers assume a hierarchical model of ability: more able workers perform
better in all occupations. The present model incorporates the unimodal ability level as a
special case and we’ll be able to draw conclusions regarding its empirical relevance. Moreover
in none of the above models does the worker take into account differences in the speed of
learning in his occupational choice. In the proposed model differences in the speed of learning
will be important in both the worker’s choice of occupation and his wage. Our proposed
model combines learning with search frictions and unemployment which the above authors
abstract from. Furthermore the current framework allows us to evaluate the importance of
overall employment experience on the worker’s wage and also the impact of an unemployment
insurance on mismatch.

The combination of search and learning frictions lends additional realism to the model,
while allowing the evaluation of actual policies and counterfactual policy experiments, such
as the effect increased unemployment benefits have on the quality of matches formed. As
is standard in most labor market models, an increase in unemployment insurance results
in an elevated unemployment rate. In the current framework however, an increased unem-
ployment rate results in lower worker learning and therefore less sorting, leading to lower
quality worker-occupation matches. This underlying mechanism contrasts with the results of
the earlier theoretical literature (Acemoglu and Shimer (1999)) that suggest that increased
unemployment insurance results in less mismatch, by allowing workers to search for better
matches.

Our model is analytically tractable and we are able to derive the steady state distribu-
tion of workers and worker flows in closed form, therefore facilitating a potential empirical estimation of the model's underlining parameters.

The paper is organized as follows: Section 2 presents the economic environment and introduces learning and search frictions. Section 3 solves the model: we first derive the value functions for the worker and the firm and solve for the wage and optimal worker behavior. We then discuss the relationships between wages, labor market experience and occupational mobility. After solving out for the steady state distribution of workers, we close the model by deriving the optimal firm behavior. Section 4 discusses the empirical relevance and predictions of the results, while Section 5 analyzes the effect of unemployment insurance on worker-occupation match quality. Section 6 examines possible extensions and Section 7 concludes. The appendix contains detailed derivations and proofs.

2 The Economy

We will examine the case of two occupations and two types. In Section 6 we discuss how the model is extended to more than two occupations.

Let the two possible types of workers be Green (G) and Red (R), and the two possible occupations be Cook (C) and Truck driver (T). Each type performs differently in each occupation, so there are four different type-occupation productivity combinations. Moreover, assume a population of risk neutral workers whose mass is normalized to 1. Each worker can be either employed or unemployed and let his flow value of leisure/unemployment benefit be equal to $b$. Workers are born and retire at rate $\gamma$, so that the total population remains constant. Moreover assume there is a large mass of ex ante homogeneous firms ensuring free entry in each occupation.

2.1 Learning Frictions and Output Determination

Each worker has a unique type that is drawn at birth and doesn’t change throughout the course of his life. We assume that both he and the market observe an initial signal about his type and form the same initial belief. When the worker of type $\nu$ 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. Output is perfectly and publicly observable by both the market and the worker, whereas productivity is not. Cumulative output is given by:

$$Y_t^{\nu i} = a_t^{\nu i} + \sigma_i Z_t$$
where $Z_t$ is a Weiner process and $\sigma_i \geq 0$ is an occupation specific constant that captures the output variance, while $a_i^\nu$ is the type $\nu$, occupation $i$ specific productivity\(^2\). If $\sigma_i = 0$, workers would be able to perfectly observe $a_i^\nu$. However if $\sigma_i > 0$, then they observe it only imperfectly. After noticing flow output each period, $dY_t^{\nu i}$, the market and the worker both update their belief in a Bayesian fashion. There is no learning while the worker is unemployed.

Since the market shares the same initial belief with the worker about his type and output history is perfectly observable, at no point is there any divergence between the beliefs of the two. In reality there is certainly a degree of informational asymmetry between the incumbent firm and other firms regarding the worker’s ability, however we do not view this assumption 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 any third party to make an informed estimate of his ability. 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\(^3\).

Recent results by Schönberg (2007) show that this assumption is not particularly restrictive. Similarly, tractability issues force us to abstract from signaling issues that may arise when the worker has superior information about his type.

Let $p \in [0,1]$ denote the posterior belief that the worker’s type is $G$. As in Moscarini (2005), since there are only two types, the agent’s beliefs follow a Bernoulli distribution and $p$ is a sufficient statistic of the worker’s output history and initial belief. After observing output, beliefs are updated in a Bayesian fashion; after applying Bayes’ rule and Ito’s lemma, the belief process is reduced to:

$$dp_t = \frac{1}{\sigma_i} \left( \frac{p_t a_i^G - a_i^R}{\sigma_i} dY_t^{\nu i} - (p_t a_i^G + (1 - p_t) a_i^R) dt \right)$$

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 $\frac{a_i^G - a_i^R}{\sigma_i}$ 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 represents a standard Wiener

---

\(^2\)To avoid trivialities, we assume that $a_i^\nu > b$ for at least one combination of $i$ and $\nu$.

\(^3\)For an example of a 3-period labor market model with firm asymmetric information the interested reader should refer to Eeckhout (2006).
2.2 Absolute vs. Comparative Advantage

We assume throughout, without loss of generality, that a type Green worker is more productive in being an cook rather than a truck driver \((a_G^C > a_G^T)\) and that type Red is better as a truck driver than a cook \((a_R^T < a_R^C)\). Clearly if both types are more productive in the same occupation, the worker would always choose it, as long as the differences in the speed of learning (signal to noise ratio) across occupations are not too large.

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, the worker will choose to work as a cook if and only if his expected output as a cook is higher than as a truck driver:

\[
p a_G^C + (1 - p) a_R^C > p a_G^T + (1 - p) a_R^T \Leftrightarrow
\]

\[
\frac{p}{1 - p} > \frac{a_R^T - a_R^C}{a_G^C - a_G^T}
\]

Therefore his occupational choice depends on relative productivities. Even if \(p\) is close to 1, the worker will still choose to work as a truck driver if the relative productivity differences across occupations for type Red are large enough compared to those of type Green. Essentially, the worker chooses to work as truck driver, if and only if the relative gain from producing in that occupation, when his type is Red, is sufficiently larger than the relative gain from working as a cook, when his type is Green. 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 Green type is better as a cook than a truck driver \((a_G^C > a_G^T)\), but we moreover assume that he is better than the Red type at being a cook \((a_G^C > a_R^C)\). Similarly a Red type, who is more productive as a truck driver rather than a cook \((a_R^C < a_R^T)\), is now assumed to be better than Green in being a truck driver \((a_T^R < a_T^R)\). This is depicted in Figure 1 which shows the expected output of the worker as function of his posterior in each occupation. 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 her.

This becomes more apparent in Figure 2, which depicts the second case: now type $G$ is not only better as a cook than type $R$ ($a^G_C > a^R_C$), but he is also better as a truck driver ($a^G_T > a^R_T$). In other words type Green has an absolute advantage in both occupations\footnote{The case where type $R$ has an absolute advantage in both occupations is equivalent.}. In Figure 2, expected productivities are equalized for $p = 0.75$. At that value of the posterior however, the worker would still prefer to work in occupation $C$, because the spread between the output of the two types is larger and therefore output realizations now reveal more information (assuming that the noise parameter $\sigma_i$ is the same in both occupations).

Finally, Figure 3 depicts the case where there is no learning in one occupation. In this case occupation $T$ is an absorbing state, since once the worker enters it, his posterior no longer changes and he has no incentive to switch again. Notice that at $p = 0.5$ he will still choose occupation $C$ because of the informational value of working there, even though his expected flow output is larger in occupation $T$.

In the remainder of the paper we will assume that type $G$ is more productive than type $R$ as a cook ($a^G_C > a^R_C$); whether he is also more productive than $R$ as a truck driver will be important for our results.

### 2.3 Search Frictions

While unemployed, each worker will choose to search in only one occupation\footnote{This will be optimal if there is a limited amount of time she can search and there are constant returns to searching.}. Unemployed workers will be matched with firms at rate $\lambda_i$ which we will take as exogenous for the time being. Worker-firm matches can dissolve either endogenously, or exogenously at rate $\delta$, after which the worker becomes unemployed. These exogenous shocks capture separations that occur for reasons beyond the modeling assumptions of this paper. Finally the worker’s wage is determined in such a way so as to split the match’s surplus according to Nash bargaining where $\beta$ is the worker’s bargaining power.

We will also allow workers to search while employed. Search activity is not observable by their employer and they are less effective at contacting new firms however, so that they are matched with a firm when searching in occupation $i$ at rate $\eta \lambda_i$, where $\eta \in [0, 1)$. When $\eta = 0$, there is no on the job search. If an employed worker contacts another firm, he accepts the new position, if and only if his value from working at the new firm strictly exceeds his current value.

If a worker employed in occupation $i$ contacts a firm in another occupation $k$, 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. Let $S_i(p)$ denote the surplus of the match in occupation $i$, which is assumed to be common knowledge. As we will see, when bargaining takes place bilaterally between the worker and the firm, the (cooperative) Nash bargaining solution prescribes that the firm receives $(1 - \beta)$ share of the surplus, where $\beta$ is the worker’s bargaining coefficient. Therefore when two firms compete for the same worker, both the incumbent and the poacher are willing to make a one time lump-sum transfer of up to $(1 - \beta)$ of their respective surplus to the worker. If $S_k(p) > S_i(p)$, the poaching firm can always outbid the incumbent in a bidding war; the incumbent knows this so it doesn’t 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. 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. It is also true that, both firms offering up to $(1 - \beta)S_i(p)$ and $(1 - \beta)S_k(p)$ respectively as a one time transfer also forms a subgame perfect Nash equilibrium of this game. However such an equilibrium is not robust to the perturbation of adding a small cost to bidding.

If the 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.

3 Model Solution

3.1 Value Functions

In what follows we will assume for simplicity that the job finding rate is the same for both occupations $\lambda_C = \lambda_T = \lambda^7$. We also maintain the assumption that a type Green worker is more productive than type Red as a cook in occupation $C$ ($a_C^G > a_C^R$). Since $p$ is a sufficient statistic of the worker’s output history and initial belief, it will also serve as a state variable

---

6 Shimer (2006) notes that, if $S_j(p) > S_i(p)$ 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, 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 she is currently being paid. Thus in this setup, the Nash bargaining solution will be bilaterally efficient.

7 This assumption ensures that occupational choice and mobility, as well as wage formation, depend only on differences in productivity and speed of learning across occupations. In Section 6 we discuss how the model is modified when we allow for different job finding rates across occupations.
for the value of an employed or unemployed worker, as well as of his employer-firm.

The process that governs the change in beliefs is a Brownian motion without a drift, so using Ito's lemma we can write the flow value of an employed worker in occupation \( i \) as:

\[
\begin{align*}
    \omega_W^i(p) &= w_i(p) + \frac{1}{2} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 W''_i(p) \\
    &- \delta [W_i(p) - U(p)] - \gamma W_i(p) + \eta \lambda \max \{ W_k(p) - W_i(p), 0 \}
\end{align*}
\]

where \( \omega \) denotes the worker's and the firm's discount rate, \( W_i(\cdot) \) is the value of the worker in occupation \( i \in \{C, T\} \), \( w_i(\cdot) \) is the occupation specific wage function and \( U(\cdot) \) the asset value of an unemployed worker.

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 gain loss resulting from either exogenous match separation which is realized at rate \( \delta \), or retirement occurring at rate \( \gamma \), plus any potential gains from on the job search (as we saw in Section 2.3.1, the 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 this case, 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 herself in occupation \( i \) will also be useful if she’s employed in another occupation \( k \).

The worker will quit his current job and switch occupations if he turns out to be sufficiently unproductive in his current occupation. For the case of occupation \( C \) let \( \bar{p}^W \) denote the value of the posterior such that \( W_C(\bar{p}^W) = U(\bar{p}^W) \); while \( p > \bar{p}^W \) the worker prefers being employed in occupation \( C \) to being unemployed (but not necessarily to working in occupation \( T \) ), but when \( p \) reaches \( \bar{p}^W \) the worker optimally quits to unemployment. Similarly if we let \( \tilde{p}^W \) denote the value such that \( W_T(\tilde{p}^W) = U(\tilde{p}^W) \), the worker optimally quits to unemployment when his \( p \) reaches \( \tilde{p}^W \). The optimality of the worker’s decision entails a smooth pasting condition at \( \bar{p}^W \) and \( \tilde{p}^W \). These thresholds can be seen in Figure 4: \( \bar{p}^W \) is equal to approximately 0.3, while \( \tilde{p}^W \) equals approximately 0.7. The worker’s value of being employed in occupation \( C \) is always above his value is being unemployed, as long as \( p > 0.3 \). As \( p \) converges to 0.3 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 \( T \).

\footnote{It can be shown that \( W_A(p) - U(p) \) (as well as \( J_A(p) \)) is increasing in \( p \) in the \([\bar{p}^W, \tilde{p}] \) interval. Similarly \( W_B(p) - U(p) \) (and \( J_B(p) \)) will be decreasing in \( p \) when \( p \in [\hat{p}, \bar{p}W] \), where \( \hat{p}, J_A(\cdot) \) and \( J_B(\cdot) \) are defined further down in the main text.}
Similarly the flow value to the firm of a filled vacancy in occupation $i$ is given by:

$$
 rJ_i (p) = \pi_i (p) - w_i (p) + \frac{1}{2} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 J_i'' (p) \\
- \delta J_i (p) - \gamma J_i (p) - \eta \lambda J_i (p) I \{ W_i (p) < W_k (p) \} 
$$

(2)

where $J_i (\cdot)$ is the asset value of the firm, $\pi_i (p) = p a_i^G + (1 - p) a_i^R$ 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 by 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.

Again for the case of occupation $C$, there exists $p^T_C$ such that $J_C (p^T_C) = 0$ and $J'_C (p^T_C) = 0$ at which point the firm fires the worker; similarly there exists a corresponding trigger $\bar{p}^T$ for occupation $T$.

The worker’s flow value of being unemployed is:

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

(3)

The flow value of being unemployed is equal to the flow benefit $b$, plus the excess value from being employed in occupation $i$ times the job finding rate, $\lambda$, 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 match destruction.

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

$$
 W_C (\hat{p}) = W_T (\hat{p}) 
$$

For $p \geq \hat{p}$, the worker will search for employment in occupation $C$ while for $p < \hat{p}$, he will search in occupation $T$. 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, as depicted in Figure 4, for $\hat{p} = 0.5$.

Finally, the worker’s wage is determined by Nash Bargaining, where $\beta \in [0, 1]$ is the worker’s share of the surplus:
\[ w_i(p) = \arg \max_w \left[ J_i(p) \right]^{1-\beta} [W_i(p) - U(p)]^\beta \]

which yields:

\[ \beta J_i(p) = (1 - \beta) [W_i(p) - U(p)] \] (4)

Since the match surplus is shared, separations are bilaterally efficient and therefore \( p_W = p^T \) and \( p^{W'} = p^{T'} \).

### 3.2 Worker Behavior and Wage Determination

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 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 (1) through (4) 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 his current occupation:

\[ w_i(p) = \beta \tilde{a}_i(p) + (1 - \beta) b + \beta \lambda J_i(p) - \frac{\beta \lambda}{2 (r + \gamma)} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 J''_i(p) \] (5)

It is easier to understand the wage equation however if we write it as follows:

\[
\begin{align*}
    w_i(p) &= (1 - \beta) (b + \lambda [W_i(p) - U(p)]) + \beta \tilde{a}_i(p) \\
    &\quad + \frac{1}{2} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 [\beta (J''_i(p) + W''_i(p)) - W''_i(p)] \\
\end{align*}
\] (5’)

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, \( \beta \). 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 \( \beta (J''_i(p) + W''_i(p)) - W''_i(p) < 0 \), the worker will compensate his employer for the additional benefit he enjoys from learning. Note 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. (1)) becomes:

$$(r + \gamma) W_i(p) = (1 - \beta)(b + \lambda [W_i(p) - U(p)])$$

$$+ \beta \left( \frac{\sigma_i(p)}{2} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 (J_i''(p) + W_i''(p)) \right)$$

$$- \delta [W_i(p) - U(p)]$$

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:

$$w_i(p) = (1 - \beta)(b + \lambda [W_k(p) - U(p)]) + \beta \sigma_i(p)$$

$$+ \frac{1}{2} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 \left[ \beta (J_i''(p) + W_i''(p)) - W_k''(p) \right]$$

$$- \eta \lambda [(1 - \beta)(W_k(p) - W_i(p)) + \beta J_i(p)]$$

The interpretation is similar to equation (5’) above with two differences: the worker’s outside option is now different 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 is 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, $W_k(p) - W_i(p)$, 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 isn’t:
\[(r + \gamma + \delta + \beta \lambda) J_i (p) = (1 - \beta) (\pi_i (p) - b) + \frac{r + \gamma + \beta \lambda}{2 (r + \gamma)} \left( \frac{a^G_i - a^R_i}{\sigma_i} \right)^2 p^2 \left(1 - p\right)^2 J_i'' (p) \]  

(7)

\[(r + \gamma + \delta + \eta \lambda) J_i (p) = (1 - \beta) (\pi_i (p) - b) + \frac{\beta \lambda}{2 (r + \gamma)} \left( \frac{a^G_i - a^R_i}{\sigma_i} \right)^2 p^2 \left(1 - p\right)^2 J_i'' (p) + \frac{1}{2} \left( \frac{a^G_i - a^R_i}{\sigma_i} \right)^2 p^2 \left(1 - p\right)^2 J_i'' (p) - \beta \lambda (1 - \eta) J_k (p) \]  

(8)

Using the boundary conditions \(J_C (p), J_T (p) < \infty, \forall p\) we can solve the differential equation for the asset value of the firm in the case where \(i = C\) and \(p > \hat{p}\):

\[J_C (p) = \frac{(1 - \beta) (\bar{\pi}_C (p) - b)}{r + \gamma + \delta + \beta \lambda} + C^C_1 p^{\frac{1}{2}} \left(1 - p\right)^{\frac{1}{2}} \left( \frac{1}{p} \right)^{\frac{1}{2}} \sqrt{\frac{4 + h_C}{\kappa_C}} \]  

(9)

as well as for the case where \(i = T\) and \(p < \hat{p}\):

\[J_T (p) = \frac{(1 - \beta) (\bar{\pi}_T (p) - b)}{r + \gamma + \delta + \beta \lambda} + C^T_2 p^{\frac{1}{2}} \left(1 - p\right)^{\frac{1}{2}} \left( \frac{p}{1 - p} \right)^{\frac{1}{2}} \sqrt{\frac{4 + h_T}{\kappa_T}} \]  

(10)

where \(\hat{p} = \frac{1}{2} \frac{r + \gamma + \delta + \beta \lambda}{(r + \gamma + \delta + \beta \lambda)(r + \gamma)} \left( \frac{a^G_i - a^R_i}{\sigma_i} \right)^2\) and \(C^C_1\) and \(C^T_2\) are undetermined coefficients.

Now we can substitute in for \(J_k (p)\) in equation (8) and using the condition that \(J_C (\hat{p}) = 0\) and \(J'_C (\hat{p}) = 0\) for occupation \(C\) (\(J_T (\hat{p}) = 0\) and \(J'_T (\hat{p}) = 0\) for occupation \(T\)), we are able to solve the resulting differential equations.

Indeed, for occupation \(C\) and \(p < \hat{p}\):

\[J_C (p) = \frac{r + \gamma + \delta + \eta \lambda}{(\frac{a^G_i - a^R_i}{\sigma_C})^2} p^{\frac{1}{2}} \left(1 - p\right)^{\frac{1}{2}} \cdot \int_{\hat{p}}^p \left[ \theta_{CT} \tau + \kappa_{CT} C^T_{CT} \tau^\xi_{CT} (1 - \tau)^{\xi_{CT} - 1} + c_{CT} \right] (\tau (1 - \tau))^{-\frac{3}{2}} \\
\cdot \left( \left( \frac{1 - \tau}{\tau} \right)^{l_{CT}} - \left( \frac{\tau - 1 - p}{\tau - 1} \right)^{l_{CT}} \right) d\tau \]  

(11)

where \(l_{CT} = \frac{1}{4} + \frac{2 (r + \gamma + \delta + \eta \lambda)}{(\frac{a^G_i - a^R_i}{\sigma_C})^2}\) and \(c_{CT}, \theta_{CT}, \kappa_{CT}\) and \(\xi_{CT}\) are defined in the appendix. For the case of occupation \(T\) and \(p > \hat{p}\):
\[ J_T(p) = \frac{r + \gamma + \delta + \eta \lambda}{\left(\frac{a_I^G - a_R^G}{\sigma_T}\right)^2} l_{TC} \cdot \int_{\tau} \left[ \theta_{TC} \tau + \kappa_{TC} C_1^C \tau \xi_{TC} (1 - \tau)^{1 - \xi_{TC}} + c_{TC} \right] \left( \tau (1 - \tau)^{-\frac{3}{2}} \right) \cdot \left( \left( \frac{1 - \tau}{1 - p} \right)^{l_{TC}} - \left( \frac{1 - \tau}{1 - p} \right)^{l_{TC}} \right) d\tau \]

where \( l_{TC} = \sqrt{\frac{1}{2} + \frac{2(r + \gamma + \delta + \eta \lambda)}{\left(\frac{a_I^C - a_R^C}{\sigma_T}\right)^2}} \) and \( c_{TC}, \theta_{TC}, \kappa_{TC} \) and \( \xi_{TC} \) are defined 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, \( p_l, p_R \) and \( \hat{p} \), as well as the 2 yet undetermined coefficients \( C_1^C \) and \( C_2^T \). We need 5 conditions to do so. Optimal stopping of on the job search at \( \hat{p} \) implies a value matching and a smooth pasting condition at that point, as shown in Figure 4. For occupation \( C \) this suggests \( \lim_{\varepsilon \to 0} W_C(\hat{p} - \varepsilon) = W_C(\hat{p}) \) and \( \lim_{\varepsilon \to 0} W'_C(\hat{p} - \varepsilon) = W'_C(\hat{p}) \), while for occupation \( T \) we have \( \lim_{\varepsilon \to 0} W_T(\hat{p} - \varepsilon) = W_T(\hat{p}) \) and \( \lim_{\varepsilon \to 0} W'_T(\hat{p} - \varepsilon) = W'_T(\hat{p}) \). Finally optimality of searching behavior while unemployed provides the last condition which was already mentioned above, \( W_C(\hat{p}) = W_T(\hat{p}) \). 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.

Equation (7) implies that \( J_C(\cdot) \) will be an increasing function of \( p \) for \( p > \hat{p}^0 \). Similarly, if \( a_I^G < a_R^R \), (i.e. type \( R \) is more productive as a truck driver than type \( G \)), \( J_T(\cdot) \) will be a decreasing function of \( p \) for \( p < \hat{p} \). In other words, the value of the worker to the firm is maximized as the worker’s type is revealed. This will not be the case when \( a_I^G > a_R^R \), i.e. type \( G \) has an absolute advantage in both occupations: now the worker’s expected productivity in occupation \( T \) is increasing in his posterior and \( J_T(\cdot) \) will generally be increasing in \( p \), for \( p < \hat{p} \) (as will \( J_C(\cdot) \) for \( p > \hat{p} \)); as the worker is revealed to be the least able of the two types, the firm’s value declines. In all cases, \( J_C(\cdot) \) and \( J_T(\cdot) \) are convex functions of the posterior.

We can now characterize the behavior of the wage function. From equation (5) the wage function \( w_C(\cdot) \) will generally be increasing in \( p \), when \( p > \hat{p} \), whereas \( w_T(\cdot) \) will be decreasing in \( p \) when \( p < \hat{p} \) and \( a_I^G < a_R^R \). On the other hand when type \( G \) performs better in both occupations \( (a_I^G > a_R^R) \), \( w_T(\cdot) \) will be increasing in \( p \) for \( p < \hat{p} \). The behavior of the wage function when the worker is searching on the job is less clear and will depend on the specific parameter values.

\(^9\)With the exception of the extreme cases where for some values of the posterior, the change in the option value of learning is sufficiently large so as to overshadow the change in expected productivity.
3.3 Returns to Experience and Hazard Rate of Occupational Switching

Workers are expected to learn their type over time, since output signals are informative. Although the unconditional expectation of 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 $G$ worker in occupation $i$:

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

In other words, if his true type is $G$, his posterior will converge in expectation to one almost surely.

Similarly for type $R$ worker employed in occupation $i$:

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

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 type $G$ worker is employed as a truck driver, as output realizations reveal information, the posterior about the worker’s type is expected to rise and will eventually switch to working for an accounting firm, either by on-the-job search or quitting to unemployment. Similarly for a type $R$ worker employed as a cook.

Moreover, in the case where type $G$ worker is more productive than type $R$ as a cook ($a_G^C > a_R^C$), and type $R$ worker is more productive than $G$ as a truck driver ($a_T^G < a_T^R$), the worker’s wage will increase with labor market experience, as his type is revealed and he ends up in the occupation that suits him best. In the case where type $G$ has an absolute advantage in both occupations ($a_T^G > a_T^R$), wages will be increasing with labor market experience for type $G$ workers, but they will be decreasing for type $R$ workers.

Furthermore, the occupational hazard switching rate falls with experience: as long a worker in occupation $C$ is above $\tilde{p}$ (below $\tilde{p}$ for occupation $T$), 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 $\tilde{p}$, the instantaneous probability jumps, since now he is searching on the job and switches occupations at rate $\eta \lambda$ and if his match is destroyed due to a $\delta$ shock, he will search for a job in occupation $T$. Of course, when $p$ hits $\tilde{p}$ ($\tilde{p}$ when employed in $T$), he quits and searches for a job in the other occupation. Therefore, since beliefs are expected to converge to zero or one with labor market experience,
in expectation occupation mobility will decline 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.

In the appendix we also show that the model also exhibits returns to occupational tenure as well. Moreover we demonstrate that in the case where no worker has absolute advantage in both occupations, expected future occupational tenure is increasing in the worker’s wage.

Summarizing we see that in the case where no worker type has an absolute advantage in both occupations, the wage is increasing in labor market experience as workers sort through occupations and their type is revealed, while occupational mobility falls. In the case where a worker type is more productive in both occupations, returns to labor market experience will be negative for the low ability workers.

3.4 The Ergodic Distribution

Summarizing so far: if unemployed, a worker will look for a job in occupation \( C \), if his posterior, \( p \), is greater than \( \hat{p} \) and in occupation \( T \) otherwise. In both occupations he lands a job at rate \( \lambda \). While employed in \( C \), he will optimally quit if and when \( p \) hits \( \hat{p} \), unless the match gets hit by an exogenous shock \( \delta \) 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 \( T \) at rate \( \eta \lambda \). Similarly if employed in occupation \( T \) he will optimally quit if \( p \) reaches \( \bar{p} \), conditional on not having been hit by a \( \delta \) shock. Again if \( p \) exceeds \( \hat{p} \) he searches on the job for another match in occupation \( C \), which materializes at rate \( \eta \lambda \). Finally he exogenously retires at rate \( \gamma \), regardless of whether he is employed or unemployed.

Since the process for the beliefs is Markovian, the stationary distribution will be ergodic. To calculate the ergodic steady state distribution of workers in this economy, we also need to specify the "birth" distribution. New workers also arrive unemployed at rate \( \gamma \), so that the total population mass remains constant and equal to unity. At birth, their initial belief and their true type is determined: they draw their initial prior from a probability density function \( g(\cdot) \) with support \([0, 1]\). Based on that realization, their type is drawn with probabilities determined by the draw of their posterior. For example, if a worker’s initial posterior is equal to 0.3, the realization of his type will result from a random draw where the outcome is \( G \) with probability 30\% and \( R \) with probability 70\%. Therefore his initial belief will be informative about his type. In what follows we will assume that \( g(\cdot) \) follows a beta\((a, \psi)\), which is a fairly general distribution family, so that \( g(p) = \frac{1}{B(a, \psi)}p^{a-1}(1-p)^{\psi-1} \) where \( B(a, \psi) = \int_0^1 x^{a-1}(1-x)^{\psi-1} \) denotes the beta function\(^{10}\). The mean value of the distribution also

\(^{10}\)The population density from which the posterior is drawn is equal to \( g(p) M \), where \( M \) is the total mass
pins down the true percentage of type Green workers in the economy.

Let \( T_i(p) \) denote the population of workers employed in occupation \( i \) whose posterior probability of being of type \( G \) 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 of type \( G \) 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 \( C \) for every \( p \geq \frac{1}{2} \) and \( p \neq \frac{1}{2} \) is given by:

\[
0 = \frac{df_C(p)}{dt} = \frac{d^2}{dp^2} \left[ \frac{1}{2} \left( \frac{a_C^G - a_C^R}{\sigma} \right)^2 p^2 (1 - p)^2 f_C(p) \right] - \delta f_C(p) - \gamma f_C(p) + \lambda z_C(p) - \eta \lambda f_C(p) I \{ p < \frac{1}{2} \} + \eta \lambda f_T(p) I \{ \frac{1}{2} \leq p \leq \frac{1}{2} \} \tag{13}
\]

This equation ensures that flows in and out of every \( p \neq \frac{1}{2} \) in the distribution of employed workers in occupation \( C \) 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 \( T \) and the worker inflow from occupation \( T \) to newly created matches in occupation \( C \) at \( p \).

The Kolmogorov forward for occupation \( T \), for every \( \frac{1}{2} \leq p \leq \frac{1}{2} \) is similar:

\[
0 = \frac{df_T(p)}{dt} = \frac{d^2}{dp^2} \left[ \frac{1}{2} \left( \frac{a_T^G - a_T^R}{\sigma} \right)^2 p^2 (1 - p)^2 f_T(p) \right] - \delta f_T(p) - \gamma f_T(p) + \lambda z_T(p) - \eta \lambda f_T(p) I \{ \frac{1}{2} \leq p \leq \frac{1}{2} \} + \eta \lambda f_C(p) I \{ \frac{1}{2} \leq p < \frac{1}{2} \} \tag{14}
\]

Equivalently, flows in and out of every \( p \) in the distributions of unemployed workers must equal. Therefore those searching for job in occupation \( C \), for every \( \frac{1}{2} \leq p \) and \( p \neq \frac{1}{2} \):

\[
\delta f_C(p) + \delta f_T(p) + \gamma g(p) = \lambda z_C(p) + \gamma z_C(p) \tag{15}
\]

The first two terms on the left hand side represent the inflow to unemployment due to
exogenous match destruction shocks from occupations $C$ and $T$ 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. The corresponding condition for those unemployed in occupation $T$, indicates that for every $p < \tilde{p}$ and $p \neq \tilde{p}$:

$$\delta f_C (p) + \delta f_T (p) + \gamma g (p) = \lambda z_T (p) + \gamma z_T (p)$$

(16)

We then use the two conditions above (eq.(15) and (16)) to solve out for $z_C (p)$ and $z_T (p)$ respectively and after substituting them into the two forward equations (13) and (14), we derive the following system of differential equations:

$$\frac{d^2}{dp^2} \left[ \frac{1}{2} \left( \frac{a^G_C - a^R_C}{\sigma_C} \right)^2 p^2 (1 - p)^2 f_C (p) \right] - \gamma \frac{\delta + \lambda + \gamma}{\lambda + \gamma} f_C (p) + \frac{\lambda \delta}{\lambda + \gamma} f_T (p) + \frac{\gamma \lambda}{\lambda + \gamma} g (p) - \eta \lambda f_C (p) I \{p < \tilde{p}\} + \eta \lambda f_T (p) I \{\tilde{p} \leq p \leq \bar{p}\} = 0$$

(17)

$$\frac{d^2}{dp^2} \left[ \frac{1}{2} \left( \frac{a^G_T - a^R_T}{\sigma_T} \right)^2 p^2 (1 - p)^2 f_T (p) \right] - \gamma \frac{\delta + \lambda + \gamma}{\lambda + \gamma} f_T (p) + \frac{\lambda \delta}{\lambda + \gamma} f_C (p) + \frac{\gamma \lambda}{\lambda + \gamma} g (p) - \eta \lambda f_T (p) I \{\tilde{p} \leq p\} + \eta \lambda f_C (p) I \{\tilde{p} \leq p < \bar{p}\} = 0$$

(18)

By taking cases the above system becomes more manageable. The appendix contains a detailed description of the solution, as well as the resulting expression for the steady state distribution of workers in the two occupations.

The solution to the differential equations contains 12 undetermined coefficients, so we need 12 conditions to pin them down. As described in the appendix, in the solution derivation we have already used the boundary conditions $f_C (\bar{p}) = 0$ and $f_T (\bar{p}) = 0$: once a worker hits the separation trigger in either occupation, he immediately quits to unemployment and therefore there are no employed worker at $\bar{p}$ and $\bar{p}$. In the solution derivation we furthermore use the condition that the mass of employed workers in each occupation is bounded from below by zero and from above by the total mass of workers, which in this case has been normalized to one ($\int_0^1 f_T (p) dp \geq 0 > -\infty$ and $\int_0^1 f_C (p) dp \leq 1 < \infty$); this condition also requires that $a, \psi > 1$ (i.e. the birth distribution is unimodal rather than U-shaped).

For the remaining 8 undetermined coefficients we make use of the following 8 conditions.

The endogenous exit of workers from occupation $T$, at $p = \bar{p}$ generates an atom of
unemployed workers at that point equal to\(^{11}\):

\[
 z_C(\bar{p}) = -\frac{1}{2} \left( a_G^C - a_R^C \right) \left( \frac{a_G^C - a_R^C}{\sigma_T} \right)^2 \bar{p}^2 (1 - \bar{p})^2 f'_T(\bar{p}^-)
\]

The inflow of unemployed workers at \( p = \bar{p} \) into employment in occupation \( C \), \( \lambda z_C(\bar{p}) \), will create a kink rather than a discontinuity in the distribution of employed workers: although there is an increased inflow of workers at \( p = \bar{p} \), compared to the every other point of the distribution, as soon as new workers find a job and produce output, they start updating their beliefs, moving to points to the right and left of \( \bar{p} \) and therefore preserving continuity:

\[
 \lim_{\varepsilon \to 0} f_C(\bar{p} + \varepsilon) = f_C(\bar{p}) \quad (19)
\]

There will be a disruption however in the first derivative of \( f_C(\cdot) \): the difference in the slope of \( f_C \) to the left and right of \( \bar{p} \) will equal the increased inflow from unemployment at that point\(^{12}\):

\[
 \frac{1}{2} \left( \frac{a_G^C - a_R^C}{\sigma_C} \right)^2 \bar{p}^2 (1 - \bar{p})^2 (f'_C(\bar{p}^-) - f'_C(\bar{p}^+)) = -\frac{\lambda}{2(\lambda + \gamma)} \left( \frac{a_G^C - a_R^C}{\sigma_T} \right)^2 \bar{p}^2 (1 - \bar{p})^2 f'_T(\bar{p}^-) \quad (20)
\]

Finally there is no reason for there to be a discontinuity in the level or the first derivative of \( f_C \) at \( \bar{p} \), so\(^ {13}\):

\[
 \lim_{\varepsilon \to 0} f_C(\bar{p} + \varepsilon) = f_C(\bar{p}) \quad (21)
\]

\[
 \lim_{\varepsilon \to 0} f'_C(\bar{p} + \varepsilon) = f'_C(\bar{p}) \quad (22)
\]

The above 4 conditions (eq(19) through(22)), along with the corresponding 4 conditions for occupation \( T \), constitute a linear system of 8 equations and 8 unknowns that complete solution to the steady distribution. See the appendix for more details.

The distribution of the posteriors for both occupations will have a fat right tail (fat left tail for the case of occupation \( T \)), as long as \( \gamma > \left( \frac{a_G^C - a_R^C}{\sigma_i} \right)^2 \) for \( i \in \{C, T\} \). 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

\(^{11}\)This condition, along with equation (15) ensure that the total flows in and out of unemployment in occupation \( A \) are equal.

\(^{12}\)It is straightforward to show that equation (18) along with the Kolmogorov forward equation for occupation \( A \) (eq. (13)) ensure that total flows in and out of occupation \( A \) employment are equal.

\(^{13}\)Unemployed workers no longer search in occupation \( A \) when the posterior crosses \( \bar{p} \) from above, but this creates a discontinuity in the second derivative of \( f_A(\cdot) \) rather than in the first, as is the case at \( \bar{p} \).
less likely to reach the extremes of the support before being forced to reset.

Notice also that a low signal-to-noise ratio also increases the endogenous exit flow from each occupation, equal to \( \frac{1}{2} \left( \frac{a_C^2 - a_T^2}{\sigma_C} \right)^2 \tilde{p}^2 (1 - \tilde{p})^2 f_C'(\tilde{p}+) \) for occupation \( C \) and \(-\frac{1}{2} \left( \frac{a_T^2 - a_C^2}{\sigma_T} \right)^2 \tilde{p}^2 (1 - \tilde{p})^2 f_T'(\tilde{p}-) \) for occupation \( T \), since speed of learning is now higher and workers sort through the occupations faster.

Given the ergodic distribution of posterior beliefs, we can also calculate the ergodic distribution of wages. Unfortunately the wage function is such that it is not possible to invert it; nevertheless it is relatively straightforward to solve for the ergodic distribution of wages numerically for a given set of parameter values, now that we have solved out for the ergodic distribution of beliefs.

In the case where no worker has absolute advantage in both occupations, the wage distribution has an empirically accurate shape, featuring a fat right tail in both occupations, similar to the one documented in the data. When however one worker type performs better in both occupations, the wage distribution will not obtain its familiar shape for both occupations, due to the form of the wage function in this case (see Section 3.2 for more details).

### 3.5 Firm Behavior

In order to close the model we need to endogenize the worker’s job finding rate \( \lambda_i \). 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 \( c v_i \), which earns no revenue while empty.

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

\[
\begin{align*}
    p_C &= \int_0^1 z_C(p) \, dp + \eta \int_{\tilde{p}}^\hat{p} f_T(p) \, dp \\
    p_T &= \int_0^\hat{p} z_T(p) \, dp + \eta \int_{\hat{p}}^\infty f_C(p) \, dp
\end{align*}
\]

A firm in occupation \( C \) contacts a worker at rate \( \frac{m(p_C, v_C)}{v_C} \), so the value of an unfilled vacancy in occupation \( C \), \( V_C \) will equal:
\[ V_C = -c v_C + \frac{m(p_C, v_C)}{v_C} \int_0^1 J_{CC} (p) \, dz_C (p) + \frac{m(p_C, v_C)}{v_C} \eta \int_p^\beta J_{CC} (p) \, df_T (p) \]

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 is less effective, the effective mass of potential workers is \( \eta \) times the mass of the employed searchers.

Similarly for occupation \( T \):

\[ V_T = -c v_T + \frac{m(p_T, v_T)}{v_T} \int_0^\beta J_{TT} (p) \, dz_T (p) + \frac{m(p_T, v_T)}{v_T} \eta \int_\beta^p J_{TT} (p) \, df_C (p) \]

Free firm entry and exit ensures that \( V_i = 0 \) and therefore:

\[ cv_C = \frac{m(p_C, v_C)}{v_C} \int_0^1 J_{CC} (p) \, dz_C (p) + \frac{m(p_C, v_C)}{v_C} \eta \int_\beta^p J_{CC} (p) \, df_T (p) \]

and:

\[ cv_T = \frac{m(p_T, v_T)}{v_T} \int_0^\beta J_{TT} (p) \, dz_T (p) + \frac{m(p_T, v_T)}{v_T} \eta \int_\beta^p J_{TT} (p) \, df_C (p) \]

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 \( \lambda_C = \lambda_T = \lambda \) 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. We discuss a possible extension of the model with different job finding rates across occupations in Section 6.

4 Discussion

Unlike the standard learning model, learning about the worker’s general human capital implies that overall labor market experience is an important determinant in the worker’s wage. Although each type performs differently across occupations, previous work experience is an important determinant of the wage as well, since it is informative about the worker’s type. Over time, workers sort into the occupations they perform best and their wage increases (at least in the case where no worker has an absolute advantage case in both occupations). Note that, even in the case of an exogenous match destruction, his wage in his new job will
continue to reflect all the accumulated knowledge about his type. Returns to experience capture the selection of workers into activities they are better suited.

In the case where a worker type has an absolute advantage in both occupations, returns to labor market experience will be negative for the low ability worker: as he is revealed to be the low ability type, his expected productivity falls and so does his wage. This can be a potential drawback for models that assume a one-dimensional ability level, where the more able worker can perform better in all occupations\textsuperscript{14}.

Furthermore as we saw in Section 3.3 the model predicts that occupational mobility is decreasing in labor market experience: increased work experience results in better occupational matches and therefore a lower occupational switching hazard. This is consistent with results in the empirical literature that document a negative correlation between worker experience and occupational mobility (McCall (1990) and Neal (1999)).

Apart from the individual wage tenure profile, the model also has the ability to capture the cross-sectional distribution of wages within an occupation. As described in Section 3.3, in the case where no one worker type is better in all occupations, the shape of the wage distribution is empirically accurate and features a fat right tail, as documented in the data. On the contrary the one-dimensional ability model cannot always replicate the shape of the wage distribution for both occupations, casting further doubt about its use.

Search in this model, although undirected within an occupation, is directed between occupations: workers choose which occupation to search in, based on their expected productivity in each occupation given their labor market experience so far and the speed of learning. Therefore occupational choice is not random, as would 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\textsuperscript{15}.

It also worth noting that the model predicts that at any given time there are simultaneous and offsetting flows to and from each occupation: workers sort through occupations learning about their human capital resulting in large offsetting flows between them. This is a well established empirical regularity that has been documented by many authors including

\textsuperscript{14}It is however consistent with Willis and Rosen (1979) whose results argue against the strict hierarchical model where one type of agent performs better at all tasks.

\textsuperscript{15}Table 2 in McCall (1990) features such a table using NLSY 79 data for workers switching from their first to their second job. In that table there are noticeable divergences among the off-diagonal elements, indicating that transitions to some occupations are more likely than to others, from a given occupation.
Jovanovic and Moffitt (1990), Murphy and Topel (1987), as well as others who point to the excess of gross worker flows over the net. 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.

One potential limitation of the model is that it doesn’t 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 to the model to allow for heterogeneity in labor demand within each occupation.

## 5 Unemployment Insurance and Mismatch

The model also provides an appropriate framework for the investigation of worker mismatch. Different types of workers sorting between occupations allows for a natural interpretation of mismatch and the present model would be suitable for the evaluation of policy initiatives designed to reduce worker mismatch. We define a worker-occupation pair mismatched if it would have not been created in a perfect information environment with no search frictions. For example, a worker of type $G$ would choose to work as a cook, if he knew his type, so if we observed him working as a truck driver we would consider him mismatched. The mismatch rate in this case can be defined as the number of mismatched workers over the total number of employed workers in the economy.

Previous authors have already pointed out how unemployment insurance reduces the mismatch rate: Acemoglu and Shimer (1999), as well as Marimon and Zilibotti (1999), argue that an increase in unemployment benefits induces workers to be more selective in their search process, thus resulting in better matches. In the current framework this effect is not present, since occupational choice of unemployed workers is hardly affected by the level of unemployment insurance. There is however a second mechanism that also leads to a reduction in mismatched workers from an increase in $b$, unemployment benefits: as the cost of being unemployed is alleviated, workers who believe that they are in bad matches, i.e. are near the quit triggers, find it less painful to leave their current job and search for a better one while unemployed. The law of large numbers implies that these matches on average will represent true mismatches, so an increase in the unemployment benefits does appear to result in an economy with better matched workers.

There is a however a more subtle mechanism that works the other way. An increase in unemployment benefits will result an increase in unemployment: the entry rate into unemployment increases because of the mechanism described above, but also less vacancies will
be created in equilibrium: the worker’s outside option is now higher, so his wage increases. Firms earn less profits from each match and therefore they lower incentives to create new vacancies. It should be noted that the because the worker’s job finding rate falls, this somewhat mitigates the increase in his outside option, which partially offsets this effect. Since workers learn about their type while employed, an increase in the unemployment rate results in lower worker sorting through different occupations and therefore more mismatch: workers spend less time employed and therefore have lower opportunities to learn about their type and are mismatched more often as a result. This result relies on the assumption that workers learn about their general human capital over their lifetime. In a model with infinitely lived workers who only care about their match quality (see for example Jovanovic (1979) or Moscarini (2005)), an increase in unemployment benefits only censors out subjectively bad matches, but does not affect sorting, since beliefs are reset among separation. In those models, in the beginning of the match all workers are identical, so previous unemployment spells do not affect their current posterior. In the framework introduced here where learning about general human capital is important, workers who have lower previous work experience are worst off and more likely to find themselves in the wrong occupation. Jovanovic (1984) extends the standard learning model to allow for an inspection component of a job, an initial signal about how the quality of the match, in addition to the existing experience component. In that case, an increase in the unemployment benefits, may have the standard effect of making workers more selective, by increasing the acceptable signal cutoff. In general, it would be interesting to study which of the two effects dominates: willingness to wait for a more promising match, or reduced learning about general human capital.

Note that the same result would also hold if search were costly and we allowed workers to choose their search effort. In that case, an increase in $b$, would lead workers to search less and therefore the unemployment rate would increase, leading to less learning and worst matches on average.

6 Extensions

It is relatively straightforward to extend the model to include three or more occupations and two types. In the three occupation case for example, each worker type is characterized by a three element vector denoting his productivity in each occupation respectively\footnote{One 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.}. 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 display positive returns to labor market experience. 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. Finally, an increase in unemployment insurance could well result in worst match formation, since the resulting higher unemployment rate will result in lower learning and less worker sorting.

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 another occupation. In that case \( \hat{p} \) would still determine when the worker decides to search on the job, but there would an 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 outweight the benefit and the worker will prefer to look for a job directly in the occupation he is interested in.

7 Conclusion

In their investigation of wage dynamics, three main alternatives have been put forth by economists\cite{17}: human capital investment, search and learning. This paper focuses on the last of these explanations by modifying the standard learning model to allow for learning about the worker’s general human capital. Indeed its predictions are quite different, underling the importance of overall employment experience in wage determination and occupational mobility. Returns to labor market experience here represent the selection of workers into occupations that are better suited for their skills, as they learn about their human capital. Furthermore, consistent with the empirical literature, occupational mobility is decreasing in employment tenure, while occupational choice is no longer random, but depends on the worker’s employment history.

The theoretical implications of the model raise questions about the assumption of the hierarchical model of ability often used in the literature where one worker type is more adept

\cite{17}Rubinstein and Weiss (2005)
in all occupations; such a framework appears to have difficulty in capturing positive returns to employment tenure and the shape of the wage distribution.

Furthermore the theory described above can be readily applied to the data: the steady state distribution of workers has been calculated, facilitating a potential structural estimation of the model. Job flows between occupations, as well as flows from and to unemployment (both exogenous as well as voluntary) also have closed form expressions and can be used in the likelihood estimation. Such an empirical estimation of the model could, for example, provide useful insights regarding the quantitative impact unemployment insurance has on worker sorting and mismatch. This mechanism could potentially be quantitatively significant, in line with most empirical studies\(^{18}\) that have difficulties detecting a positive impact of unemployment insurance on the match quality that had been presumed. Furthermore, an empirical estimation of the model would also shed light on another important policy question regarding the effectiveness of job subsidies.

\(^{18}\)See for example van Ours and Vodopivec (2006) and references therein.
Figures

Figure 1: Expected Output as a Function of Beliefs

Expected Output in Occupation C

Expected Output in Occupation T

Figure 2: Expected Output as a Function of Beliefs

Expected Output in Occupation T

Expected Output in Occupation C
Figure 3: Expected Output as a Function of Beliefs

Expected Output in Occupation C

Expected Output in Occupation T

Figure 4: Worker Value Functions

Value of Employed in Occ. C

Value of Employed in Occ. T

Value of Unemployed Worker
Appendix

A Wage and Trigger Derivation

For clarity, to distinguish between the case where the worker’s outside option is his current occupation or not (and consequently whether he is searching on the job), we introduce another subscript $k$ for the three value functions and the wage to denote the workers outside option. Formally we have:

$$k = \text{arg max}_i [W_i (p) - U_i (p)]$$

$J_{CT} (\cdot)$, for example, denotes the asset value of a filled vacancy in occupation $C$ with a worker whose outside option is occupation $T$, in other words when $p \in (\bar{p}, \bar{p})$. Similarly, $w_{TT} (p)$, denotes the wage of a worker who does not find it optimal to switch occupations by on the job search, thus $p \in [0, \bar{p}]$, while $U_C (\cdot)$ denotes the value of an unemployed worker searching for a job in occupation $C$, i.e. $p \in [\bar{p}, 1]$.

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 (1) through (3) become:

$$rW_{ii} (p) = w_{ii} (p) + \frac{1}{2} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 W''_{ii} (p)$$  \hspace{1cm} (1')

$$- \delta [W_{ii} (p) - U_i (p)] - \gamma W_{ii} (p)$$

$$rJ_{ii} (p) = \bar{a}_i (p) - w_{ii} (p) + \frac{1}{2} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 J''_{ii} (p)$$  \hspace{1cm} (2')

$$- \delta J_{ii} (p) - \gamma J_{ii} (p)$$

$$rU_i (p) = b + \lambda [W_{ii} (p) - U_i (p)] - \gamma U_i (p)$$  \hspace{1cm} (3')

We subtract the worker’s value of being unemployed (eq. (3')) from his value of being employed (eq. (1')) and multiply through by $(1 - \beta)$:
\[(1 - \beta) r (W_{ii} (p) - U_i (p)) = (1 - \beta) (w_{ii} (p) - b) + \frac{1}{2} (1 - \beta) \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 W''_{ii} (p) \]
\[-\delta (1 - \beta) [W_{ii} (p) - U_i (p)] - \gamma (1 - \beta) W''_{ii} (p) \]
\[-\lambda (1 - \beta) [W_{ii} (p) - U_i (p)] + \gamma (1 - \beta) U_i (p) \]

We similarly multiply the asset value of a filled vacancy (eq. (2')) by \(\beta\):

\[\beta r J_i (p) = \beta \pi_i (p) - \beta w_{ii} (p) + \frac{1}{2} \beta \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 J''_{ii} (p) \]
\[-\beta \delta J_{ii} (p) - \beta \gamma J_{ii} (p) \]

We then subtract the above two equations and using the surplus sharing condition (eq. (4)) we obtain:

\[w_{ii} (p) - (1 - \beta) b + \frac{1}{2} (1 - \beta) \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 W''_{ii} (p) \]
\[-\lambda (1 - \beta) [W_{ii} (p) - U_i (p)] - \beta \pi_i (p) - \frac{1}{2} \beta \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 J''_{ii} (p) = 0 \Rightarrow \]
\[w_{ii} (p) = \beta \pi_i (p) + (1 - \beta) b + \beta \lambda J_{ii} (p) - \frac{1}{2} (1 - \beta) \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 W''_{ii} (p) \]
\[+ \frac{1}{2} \beta \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 J''_{ii} (p) \]  \hspace{1cm} (23)

Using the surplus sharing condition once again:

\[\beta J''_{ii} (p) = (1 - \beta) (W''_{ii} (p) - U''_i (p)) \]
\[W''_{ii} (p) = \frac{\beta}{1 - \beta} J''_{ii} (p) + U''_i (p) \]  \hspace{1cm} (24)

However from the value of the unemployed worker (eq. (3')), we have:

\[U_i (p) = \frac{b}{r + \gamma + \lambda} + \frac{\lambda}{r + \gamma + \lambda} W_{ii} (p) \Rightarrow \]
\[U''_i (p) = \frac{\lambda}{r + \gamma + \lambda} W''_{ii} (p) \]

31
Substituting out for $U_0^i(p)$ in equation (24) results in:

$$W_0^{ii}(p) = \frac{\beta}{1-\beta} J_0^{ii}(p) + \frac{\lambda}{r + \lambda + \gamma} W_0^{ii}(p) \Rightarrow$$

$$W_0^{ii}(p) = \frac{\beta}{1-\beta} \frac{r + \gamma + \lambda}{r + \gamma} J_0^{ii}(p) \quad (25)$$

We can now substitute out for $W_0^{ii}(p)$ in equation (23) and obtain the wage as a function of $J_0^{ii}(\cdot)$ only (eq. (5) in the main text):

$$w_0^{ii}(p) = \beta \pi_i(p) + (1 - \beta) b + \beta \lambda J_0^{ii}(p) - \frac{1}{2} \beta \left( \frac{a_i^g - a_i^R}{\sigma_i} \right)^2 p^2 (1-p)^2 \frac{r + \gamma + \lambda}{r + \gamma} J_0^{ii}(p)$$

$$+ \frac{1}{2} \beta \left( \frac{a_i^g - a_i^R}{\sigma_i} \right)^2 p^2 (1-p)^2 J_0^{ii}(p)$$

$$\Leftrightarrow w_0^{ii}(p) = \beta \pi_i(p) + (1 - \beta) b + \beta \lambda J_0^{ii}(p) - \frac{1}{2} \beta \frac{\lambda}{2 (r + \gamma)} \left( \frac{a_i^g - a_i^R}{\sigma_i} \right)^2 p^2 (1-p)^2 J_0^{ii}(p)$$

The alternative expression for the wage (eq. (5')) can be derived by using equation (23) above and the surplus sharing condition (eq. (4)).

Replacing the wage into the value of the… rm function (eq. (2')) produces a differential equation with respect to $J_0^{ii}(\cdot)$ (eq. (7) in the main text):

$$(r + \delta + \gamma + \lambda \beta) J_0^{ii}(p) = (1 - \beta) (\pi_i(p) - b) + \frac{r + \gamma + \beta \lambda}{2 (r + \gamma)} \left( \frac{a_i^g - a_i^R}{\sigma_i} \right)^2 p^2 (1-p)^2 J_0^{ii}(p)$$

The general solution to the above differential equation is:

$$J_0^{ii}(p) = \frac{(1 - \beta) (\pi_i(p) - b)}{r + \gamma + \delta + \beta \lambda} + C_1^i p^{\frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + h_i}{h_i}}} (1-p)^{\frac{1}{2} + \frac{1}{2} \sqrt{\frac{4 + h_i}{h_i}}}$$

$$+ C_2^i p^{\frac{1}{2} + \frac{1}{2} \sqrt{\frac{4 + h_i}{h_i}}} (1-p)^{\frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + h_i}{h_i}}}$$

where $h_i = \frac{1}{2} \frac{r + \gamma + \beta \lambda}{(r + \gamma + \delta + \beta \lambda)(r + \gamma)} \left( \frac{a_i^g - a_i^R}{\sigma_i} \right)^2$ and $C_1^i$ and $C_2^i$ are undetermined coefficients. For the case of $i = C$, when $p \to 1$, $\lim_{p \to 1} C_2^i p^{\frac{1}{2} + \frac{1}{2} \sqrt{\frac{4 + h_i}{h_i}}} (1-p)^{\frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + h_i}{h_i}}} = C_2^i \cdot 1 \cdot \lim_{p \to 1} (1-p)^{\frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + h_i}{h_i}}} = 32$
\[ \because \text{since } h_i > 0, \text{ then } \sqrt{\frac{4 + 4h_i}{h_i}} > 1, \text{ so } \frac{1}{2} \left( 1 - \sqrt{\frac{4 + 4h_i}{h_i}} \right) < 0. \] 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 Green which is finite, it must be the case that \( C^C_2 = 0 \). A similar argument for \( i = T \) and \( p \to 1 \) leads to \( C^T_1 = 0 \).

Therefore for \( p \geq \hat{p} \) for the case of occupation \( C \), we obtain equation (9) in the main text and for \( p \leq \hat{p} \) and occupation \( T \) we are left with equation (10).

For the case where the worker’s occupation of choice if unemployed differs from his current one \( (i \neq k) \), equations (1) through (3) become:

\[
\begin{align*}
 rW_{ik} (p) &= w_{ik} (p) + \frac{1}{2} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 W''_{ik} (p) \\
 &\quad - \delta [W_{ik} (p) - U_k (p)] - \gamma W_{ik} (p) + \eta \lambda [W_{kk} (p) - W_{ik} (p)] \\
 &\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad\]
\[ W''_{ik} (p) = \frac{\beta}{1 - \beta} J''_{ik} (p) + \frac{\lambda}{r + \gamma + \lambda} W''_{kk} (p) \]

which we use to substitute out for \( W''_{ik} (p) \) in the wage equation (eq. (26)):

\[
\begin{align*}
    \omega_{ik} (p) &= (1 - \beta) b + \beta \bar{\alpha}_i (p) \\
    &- \frac{1}{2} (1 - \beta) \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 \left( \frac{\beta}{1 - \beta} J''_{ik} (p) + \frac{\lambda}{r + \gamma + \lambda} W''_{kk} (p) \right) \\
    &- (1 - \beta) \eta \lambda (W_{kk} (p) - W_{ik} (p)) + \beta \lambda J_{kk} (p) \\
    &+ \frac{1}{2} \beta \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 J''_{ik} (p) - \beta \eta \lambda J_{ik} (p) \\
    &\Rightarrow \\
    \omega_{ik} (p) &= (1 - \beta) b + \beta \bar{\alpha}_i (p) - \frac{\lambda (1 - \beta)}{2 (r + \gamma + \lambda)} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 W''_{kk} (p) \\
    &- (1 - \beta) \eta \lambda (W_{kk} (p) - W_{ik} (p)) + \beta \lambda J_{kk} (p) - \beta \eta \lambda J_{ik} (p)
\end{align*}
\]

(27)

Using the same transformation as in the case of equation (5'), we can rewrite equation (27) above, as equation (6) in the main text.

Once again, using the surplus sharing condition we have:

\[
W_{kk} (p) = \frac{\beta}{1 - \beta} J_{kk} (p) + U_k (p)
\]

and:

\[
W_{ik} (p) = \frac{\beta}{1 - \beta} J_{ik} (p) + U_k (p)
\]

Subtracting one from the other:

\[
W_{kk} (p) - W_{ik} (p) = \frac{\beta}{1 - \beta} (J_{kk} (p) - J_{ik} (p))
\]

Using the above result as well as equation (25), allows us to write the wage equation (eq. (27)) as function of \( J_{kk} (\cdot) \) and \( J_{ik} (\cdot) \):

\[
\omega_{ik} (p) = (1 - \beta) b + \beta \bar{\alpha}_i (p) - \frac{\beta \lambda}{2 (r + \gamma)} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 J''_{kk} (p) + \beta \lambda (1 - \eta) J_{kk} (p)
\]

(28)

Substituting in for the wage (eq. (28)) in the asset value of the firm condition gives us the following differential equation (eq. (8) in the main text):
\[(r + \delta + \gamma + \eta \lambda) J_{ik} (p) = \pi_i (p) - (1 - \beta) b - \beta \pi_i (p)\]
\[+ \frac{\beta \lambda}{2 (r + \gamma)} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 J''_{kk} (p) - \beta \lambda (1 - \eta) J_{kk} (p)\]
\[+ \frac{1}{2} \left( \frac{a_i^G - a_i^R}{\sigma_i} \right)^2 p^2 (1 - p)^2 J''_{ik} (p)\]

We can now use the expression we derived above and replace \(J_{kk} (\cdot)\) and \(J_{kk}'' (\cdot)\) in the above expression and obtain a differential equation in \(J_{ik} (p)\). More specifically, for the case of \(i = C\) and \(k = T\), we have:

\[\phi_{CT} p^2 (1 - p)^2 J''_{CT} (p) = J_{CT} (p) + c_{CT} + \theta_{CT} p + \kappa_{CT} C_T^2 p^{\xi_{CT}} (1 - p)^{\xi_{CT}}\]

where:

\[\phi_{CT} = \frac{1}{2 (r + \gamma + \delta + \eta \lambda)} \left( \frac{a_i^G - a_i^R}{\sigma_C} \right)^2\]
\[c_{CT} = \frac{1 - \beta}{r + \gamma + \delta + \eta \lambda} \left( \beta \lambda (1 - \eta) \frac{\left( a_T^R - b \right)}{r + \gamma + \delta + \beta \lambda} - \left( a_C^R - b \right) \right)\]
\[\theta_{CT} = \frac{1 - \beta}{r + \gamma + \delta + \eta \lambda} \left( \beta \lambda (1 - \eta) \frac{a_T^G - a_T^R}{r + \gamma + \delta + \beta \lambda} - \left( a_C^G - a_C^R \right) \right)\]
\[\kappa_{CT} = - \frac{\beta \lambda}{r + \gamma + \delta + \eta \lambda} \left( \frac{r + \gamma + \delta + \beta \lambda}{r + \gamma + \beta \lambda} \left( \frac{a_T^G - a_T^R}{\sigma_C} \sigma_T \right)^2 - (1 - \eta) \right)\]
\[\xi_{CT} = \frac{1}{2} + \frac{1}{2} \sqrt{\frac{4 + \eta_T}{\eta_T}}\]

Similarly for the case of \(i = T\) and \(k = C\):

\[\phi_{TC} p^2 (1 - p)^2 J''_{TC} (p) = J_{TC} (p) + c_{TC} + \theta_{TC} p + \kappa_{TC} C_T^2 p^{\xi_{TC}} (1 - p)^{\xi_{TC}}\]

where again:
\[ \phi_{TC} = \frac{1}{2(r + \gamma + \delta + \eta \lambda)} \left( \frac{a_G^T - a_R^T}{\sigma_T} \right)^2 \]

\[ c_{TC} = \frac{1 - \beta}{r + \gamma + \delta + \eta \lambda} \left( \beta \lambda (1 - \eta) \frac{(a_G^T - b)}{r + \gamma + \delta + \beta \lambda} - (a_R^T - b) \right) \]

\[ \theta_{TC} = \frac{1 - \beta}{r + \gamma + \delta + \eta \lambda} \left( \beta \lambda (1 - \eta) \frac{a_G^C - a_R^C}{r + \gamma + \delta + \beta \lambda} - (a_G^T - a_R^T) \right) \]

\[ \kappa_{TC} = -\frac{\beta \lambda}{r + \gamma + \delta + \eta \lambda} \left( \frac{r + \gamma + \delta + \beta \lambda}{r + \gamma + \beta \lambda} \left( \frac{a_G^C - a_R^C}{a_G^C - a_R^C} \right)^2 - (1 - \eta) \right) \]

\[ \xi_{TC} = \frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + h_C}{h_C}} \]

Using the value matching and smooth pasting conditions \( J_{CT} (\bar{p}) = 0 \) and \( J_{CT}' (\bar{p}) = 0 \) for the case of \( i = C \) and \( J_{TC} (\bar{p}) = 0 \) and \( J_{TC}' (\bar{p}) = 0 \) for the case of \( i = T \) we solve the above differential equations to obtain equations (11) and (12) in the main text.

Finally, to pin down the values of \( \bar{p}, \bar{p}, C_1^C \) and \( C_2^T \) we will use the 5 conditions mentioned in the main text. Straightforward derivations imply that they can be rewritten in terms of the value of the firm rather than that of the worker as follows:

\[ J_{CT} (\bar{p}) = J_{CC} (\bar{p}) \]
\[ J_{TC} (\bar{p}) = J_{TT} (\bar{p}) \]
\[ J_{CC} (\bar{p}) = J_{TT} (\bar{p}) \]
\[ \frac{r + \gamma + \lambda}{r + \gamma} J_{CC}' (\bar{p}) = J_{CT}' (\bar{p}) + \frac{\lambda}{r + \gamma} J_{TT}' (\bar{p}) \]
\[ \frac{r + \gamma + \lambda}{r + \gamma} J_{TT}' (\bar{p}) = J_{TC}' (\bar{p}) + \frac{\lambda}{r + \gamma} J_{CC}' (\bar{p}) \]

**B Returns to Occupational Tenure**

In this section we investigate how occupational tenure is related to the worker’s wage.

For the case of a worker employed in occupation \( C \):

\[ E (p_{t + \Delta t} | \text{no occupational switching until } t + \Delta t, p_t) = E (p_t | p_{t + \Delta t} > p_t, p_t) > E (p_{t + \Delta t} | p_t) = p_t \]

Her expectation of his future posterior conditional on no occupational switching until \( t + \Delta t \)
is strictly higher than the unconditional one, since occupational switching is less likely to occur for higher beliefs.

Similarly for a worker employed in occupation $T$:

$$E(p_{t+\Delta t} | \text{no occupational switching until } t + \Delta t, p_t) = E(p_t | p_{t+\Delta t} < \bar{p}, p_t) < E(p_{t+\Delta t} | p_t) = p_t$$

Therefore beliefs are a strict submartingale. Since sorting occurs over time, the longer a worker has remained in an occupation, the more likely it is that he has found to be better suited there and that his posterior will be converging to 0 or 1.

In the case where a type $G$ worker has an absolute advantage in both occupations ($a_G^C > a_R^C$), wages will be increasing on average with occupational tenure in occupation $C$, but decreasing in occupation $T$.

We will now explicitly derive the expression for expected future occupational tenure in both occupations and show that expected occupational tenure is increasing in the value of the posterior in occupation $C$ and decreasing in the value of the posterior in occupation $T$. Therefore, at least in the case where no worker has absolute advantage in both occupations, expected future occupational tenure is higher when the worker’s wage is higher.

Following Karlin and Taylor (1981), page 192 and Moscarini (2005), we compute the expected future occupational tenure $\xi_i(p)$ of worker with posterior $p$ employed in occupation $i$, as the solution of the following differential equations (one for each occupation):

$$\frac{1}{2} p^2 (1-p)^2 \left( \frac{a_G^C - a_R^C}{\sigma_C} \right)^2 \xi_C''(p) - \delta \xi_C(p) - \gamma \xi_C(p) - \eta \lambda \xi_C(p) I(p < \bar{p}) = -1$$

$$\frac{1}{2} p^2 (1-p)^2 \left( \frac{a_G^T - a_R^T}{\sigma_T} \right)^2 \xi_T''(p) - \delta \xi_T(p) - \gamma \xi_T(p) - \eta \lambda \xi_T(p) I(p > \bar{p}) = -1$$

We also have the following 8 conditions that will help us pin down the undetermined coefficients in each case: $\xi_C(1) = \frac{1}{\delta + \gamma}, \xi_T(0) = \frac{1}{\delta + \gamma}, \xi_C(\bar{p}) = 0, \xi_T(\bar{p}) = 0, \xi_C(\bar{p}^-) = \xi_C(\bar{p}^+), \xi_C(\bar{p}^-) = \xi_C(\bar{p}^+), \xi_T(\bar{p}^-) = \xi_T(\bar{p}^+)$ and $\xi_T(\bar{p}^-) = \xi_T(\bar{p}^+)$.

Using the fact that $\xi_C(1) = \frac{1}{\delta + \gamma}$, we can write up the solution to the above differential equation for the case of occupation $C$ when $p > \bar{p}$ as:

$$\xi_C(p) = \frac{1}{\delta + \gamma} + C_1 p^{\frac{1}{2} - \frac{1}{2 \sqrt{\frac{\delta + \gamma}{\sigma_C}}} (1 - p)^{\frac{1}{2} + \frac{1}{2 \sqrt{\frac{\delta + \gamma}{\sigma_C}}}}$$

where $\nu_1 = \frac{1}{2(\delta + \gamma)} \left( \frac{a_G^C - a_R^C}{\sigma_C} \right)^2$ and when $p < \bar{p}$:
\[ \xi_C(p) = \frac{1}{\delta + \gamma + \eta \lambda} + C_3^\xi p^{\frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}}} (1 - p)^{-\frac{1}{2} + \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}}} + C_4^\xi p^{\frac{1}{2} + \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}}} (1 - p)^{\frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}}} \]

where \( \nu_2 = \frac{1}{2(\delta + \gamma + \eta \lambda)} \left( \frac{\sigma_C^2 - a_C^2}{\sigma_C^2} \right)^2 \).

Note that \( C_3^\xi < 0 \), since \( \xi_C(1) = \frac{1}{\delta + \gamma} \) is the maximum value that \( \xi_C(\cdot) \) can attain. The 3 remaining undetermined coefficients are pinned down by \( \xi_C(p) = 0 \), \( \xi_C(\hat{p}^-) = \xi_C(\hat{p}^+) \) and \( \xi'_C(\hat{p}^-) = \xi'_C(\hat{p}^+) \).

We will now prove that \( \xi_C(\cdot) \) is increasing in \( p \).

For \( p > \hat{p} \) we have that:

\[ \xi_C(p) = C_3^\xi p^{-\frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}}} (1 - p)^{-\frac{1}{2} + \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}}} \left( \frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + \nu_1}{\nu_1}} - p \right) > 0 \]

Since \( \nu_1 > 0 \), \( \sqrt{\frac{4 + \nu_1}{\nu_1}} > 1 \) and thus \( \frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + \nu_1}{\nu_1}} - p < 0 \), but given the fact that \( C_3^\xi < 0 \), \( \xi_C(p) > 0 \).

Moreover for the case of \( p < \hat{p} \) we have:

\[ \xi'_C(p) = \left( \frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}} - p \right) C_3^\xi p^{-\frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}}} (1 - p)^{-\frac{1}{2} + \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}}} + \left( \frac{1}{2} + \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}} - p \right) C_4^\xi p^{-\frac{1}{2} + \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}}} (1 - p)^{-\frac{1}{2} - \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}}} \]

We know that \( \xi'_C(\hat{p}^-) = \xi'_C(\hat{p}^+) > 0 \) and since \( \xi_C(p) = 0 \), then it must be the case that \( \xi'_C(p) \geq 0 \) since expected tenure is bounded from below by zero. Given that \( \xi_C(\cdot) \) is continuous and differentiable everywhere in the interval \((\hat{p}_-, \hat{p})\), by the mean value theorem, it suffices to show that there is at most only one \( p \) between \( \hat{p} \) and \( \hat{p} \) where the slope is zero in order to prove that it is increasing in \( p \) in this interval as well.

From above we can see that \( \xi'_C(p) = 0 \) is equivalent to:

\[ C_3^\xi = -C_4^\xi \frac{1}{2} + \frac{1}{2} \sqrt{\frac{4 + \nu_2}{\nu_2}} - p \left( \frac{p}{1 - p} \right) \sqrt{\frac{4 + \nu_2}{\nu_2}} \]  \hspace{1cm} (29)

It can be shown that the right hand side of (29) is either globally increasing or globally decreasing depending on the sign of \( C_4^\xi \). Thus there can be at most only one \( p \) such that \( \xi'_C(p) = 0 \) and given the discussion above we conclude that \( \xi_C(\cdot) \) is globally increasing in \( p \).

A similar derivation proves that \( \xi_T(\cdot) \) is globally decreasing in \( p \).
C   Ergodic Distribution Derivation

Now, in order to solve the system of differential equations (17) and (18), we will have to take cases.

For \( p \in [0, \bar{p}] \), we know that \( f_C (p) = 0 \) (workers in occupation \( C \) quit once they hit \( p \)), so equation (18) reduces to:

\[
\frac{d^2}{dp^2} \left[ \frac{1}{2} \left( \frac{a_T^C - a_T^R}{\sigma_T} \right)^2 p^2 (1 - p)^2 f_T (p) \right] - \gamma \frac{\delta + \lambda + \gamma}{\lambda + \gamma} f_T (p) + \frac{\lambda \gamma}{\lambda + \gamma} f_C (p) = 0
\]

The solution to the above differential is given by:

\[
f_T (p) = C_T^1 p^{q_T - 2} (1 - p)^{-1 - q_T} + C_T^2 p^{-1 - q_T} (1 - p)^{q_T - 2}
\]

\[
- \frac{d}{\sqrt{c_T (4 + c_T)}} p^{q_T - 2} (1 - p)^{-1 - q_T} \int_0^p \tau^{q_T + a - 1} (1 - \tau)^{\psi - q_T} d\tau
\]

\[
+ \frac{d}{\sqrt{c_T (4 + c_T)}} p^{-1 - q_T} (1 - p)^{q_T - 2} \int_0^p \tau^{q_T + a - 1} (1 - \tau)^{\psi - q_T} d\tau
\]

where \( q_T = \frac{1}{2} - \frac{1}{2} \sqrt{\frac{8(\delta + \lambda + \gamma)\gamma}{\left( \frac{a_T^C - a_T^R}{\sigma_T} \right)^2 (\lambda + \gamma)}} + 1, c_T = \left( \frac{a_T^C - a_T^R}{\sigma_T} \right)^2 (\lambda + \gamma), d = - \frac{\lambda}{(\delta + \lambda + \gamma)B(a, \psi)} \) and \( C_T^1 \) and \( C_T^2 \) are undetermined coefficients. However since \( \int_0^1 \tau^{q_T - 2} \log \tau \log \tau \tau^{q_T - 1} d\tau = -\infty \) as \( x \to 0 \), because \( q_T < 0 \), it must be the case that \( C_T^1 = 0 \) because the mass of workers at any interval is bounded from below by zero. For the same reason we require that  \( a > 1 \).

For \( p \in (p, \bar{p}) \), for the case of occupation \( C \), since \( z_C (p) = 0 \) at that interval, equation (13) becomes:

\[
\frac{d^2}{dp^2} \left[ \frac{1}{2} \left( \frac{a_C^C - a_C^R}{\sigma_C} \right)^2 p^2 (1 - p)^2 f_C (p) \right] - (\delta + \gamma) f_C (p) - \eta \lambda f_C (p) = 0
\]

The solution to the above differential equation is given by:

\[
f_C (p) = C_5^C p^{-\frac{3}{2} - \kappa_T} (1 - p)^{-\frac{3}{2} + \kappa_T} + C_6^C p^{-\frac{3}{2} + \kappa_T} (1 - p)^{-\frac{3}{2} - \kappa_T}
\]

where \( \kappa_T = \frac{1}{4} + \frac{2(\delta + \lambda + \gamma)}{\left( \frac{a_C^C - a_C^R}{\sigma_C} \right)^2} \).

Using the boundary condition, \( f_C (p) = 0 \), we are able to substitute out for \( C_5^C \).
Now we are in a position to substitute out for $f_C(p)$ in equation (18) and obtain:

$$
\frac{d^2}{dp^2} \left[ \frac{1}{2} \left( \frac{a_C^2 - a_R^2}{\sigma_T} \right)^2 p^2 (1-p)^2 f_T(p) \right] - \frac{\delta + \lambda + \gamma}{\lambda + \gamma} f_T(p) = 0
$$

The solution to the above differential equation gives the expression for $f_T(p)$ for the $(\hat{p}, \bar{p})$ interval.

The derivation of $f_C(p)$ and $f_T(p)$ for the $[\hat{p}, \bar{p})$ interval is symmetric and otherwise identical to the one for the $(\bar{p}, \hat{p})$ interval: using the fact that $z_T(p) = 0$, equation (14) simplifies considerably and after solving the resulting differential equation and using the condition that $f_T(\bar{p}) = 0$, one obtains an expression for $f_T(p)$ with only one undetermined coefficient. Substituting into equation (17) and solving, results in the corresponding expression for $f_C(p)$.

Finally, the case where $p \in (\bar{p}, 1]$ is similarly symmetric to the case where $p \in [0, \bar{p})$. Now we know that $f_T(p) = 0$ and therefore equation (17) simplifies substantially. Solving the resulting differential equation and using the condition that the mass of employed workers in occupation $C$ is bounded from above by the one, obtains the expression for $f_C(p)$ in that interval with one undetermined coefficient.

Given the above, the resulting steady state distribution of workers employed in occupation $C$ is given by:

For $p \in (\hat{p}, \bar{p}]$:

$$f_C(p) = C_6^C (p (1-p))^{-\frac{3}{2}} \left( \left( \frac{p}{1-p} \right)^{\kappa_T} - \left( \frac{p}{\bar{p}} \right)^{2\kappa_T} \left( \frac{1-p}{p} \right)^{2\kappa_T} \right)$$

for $p \in (\hat{p}, \bar{p}]$:
\[
f_C(p) = C_5 p^{\gamma_c - 2} (1 - p)^{-1 - \gamma_c} + C_4 p^{-1 - \gamma_c} (1 - p)^{\gamma_c - 2} \\
- \frac{1}{\sqrt{c_c (4 + c_c)}} p^{\gamma_c - 2} (1 - p)^{-1 - \gamma_c} [C_5 m_c \int_\tilde{p}^p \tau^{-\frac{1}{2} + \kappa_c - \gamma_c} (1 - \tau)^{-\frac{3}{2} - \kappa_c + \gamma_c} d\tau \\
+ C_5 T n \int_\tilde{p}^p \tau^{-\frac{1}{2} - \kappa_c - \gamma_c} (1 - \tau)^{-\frac{3}{2} + \kappa_c + \gamma_c} d\tau + d \int_\tilde{p}^p \tau^{a - \gamma_c} (1 - \tau)^{\psi + \gamma_c - 1} d\tau] \\
+ \frac{1}{\sqrt{c_c (4 + c_c)}} p^{-1 - \gamma_c} (1 - p)^{\gamma_c - 2} [C_5 m_c \int_\tilde{p}^p \tau^{-\frac{3}{2} + \kappa_c + \gamma_c} (1 - \tau)^{-\frac{1}{2} - \kappa_c - \gamma_c} d\tau \\
+ C_5 T n \int_\tilde{p}^p \tau^{-\frac{3}{2} - \kappa_c + \gamma_c} (1 - \tau)^{-\frac{1}{2} + \kappa_c - \gamma_c} d\tau + d \int_\tilde{p}^p \tau^{a + \gamma_c - 1} (1 - \tau)^{\psi - \gamma_c} d\tau] \\
\]

and for \( p \in [\tilde{p}, 1] \):

\[
f_C(p) = C_4 p^{\gamma_c - 2} (1 - p)^{-1 - \gamma_c} \\
- \frac{d}{\sqrt{c_c (4 + c_c)}} p^{\gamma_c - 2} (1 - p)^{-1 - \gamma_c} \int_1^p \tau^{a - \gamma_c} (1 - \tau)^{\gamma_c + \psi - 1} d\tau \\
+ \frac{d}{\sqrt{c_c (4 + c_c)}} p^{-1 - \gamma_c} (1 - p)^{\gamma_c - 2} \int_1^p \tau^{\gamma_c + a - 1} (1 - \tau)^{\psi - \gamma_c} d\tau
\]

And similarly for occupation \( T \), in the case where \( p \in [0, \tilde{p}] \):

\[
f_T(p) = C_2 p^{-1 - \gamma_T} (1 - p)^{\gamma_T - 2} \\
- \frac{d}{\sqrt{c_T (4 + c_T)}} p^{\gamma_T - 2} (1 - p)^{-1 - \gamma_T} \int_0^p \tau^{a - \gamma_T} (1 - \tau)^{\gamma_T + \psi - 1} d\tau \\
+ \frac{d}{\sqrt{c_T (4 + c_T)}} p^{-1 - \gamma_T} (1 - p)^{\gamma_T - 2} \int_0^p \tau^{\gamma_T + a - 1} (1 - \tau)^{\psi - \gamma_T} d\tau
\]

for \( p \in (\tilde{p}, \overline{p}) \):
\[
\begin{align*}
f_T(p) &= C_3^T p^{q_T - 2} (1 - p)^{-1 - q_T} + C_4^T p^{-1 - q_T} (1 - p)^{q_T - 2} \\
 &+ \frac{1}{\sqrt{c_T(4 + c_T)}} p^{q_T - 2} (1 - p)^{-1 - q_T} \int_\tilde{p}^p \tau^{-\frac{3}{2} - \kappa_T - q_T} (1 - \tau)^{-\frac{3}{2} + \kappa_T + q_T} d\tau \\
 &+ C_6^C n \int_\tilde{p}^p \tau^{-\frac{3}{2} + \kappa_T + q_T} (1 - \tau)^{-\frac{3}{2} - \kappa_T + q_T} d\tau + d \int_\tilde{p}^p \tau^{a - q_T} (1 - \tau)^{\psi + q_T - 1} d\tau \\
 &+ \frac{1}{\sqrt{c_T(4 + c_T)}} p^{-1 - q_T} (1 - p)^{q_T - 2} \int_\tilde{p}^p \tau^{-\frac{3}{2} - \kappa_T - q_T} (1 - \tau)^{-\frac{3}{2} + \kappa_T + q_T} d\tau \\
 &+ C_6^C n \int_\tilde{p}^p \tau^{-\frac{3}{2} + \kappa_T + q_T} (1 - \tau)^{-\frac{3}{2} - \kappa_T + q_T} d\tau + d \int_\tilde{p}^p \tau^{a + q_T - 1} (1 - \tau)^{\psi - q_T} d\tau
\end{align*}
\]

and for \( p \in [\tilde{p}, p] \):

\[
\begin{align*}
f_T(p) &= C_5^T (p(1 - p))^{-\frac{3}{2}} \left( \left( \frac{1 - p}{p} \right)^{\kappa_C} - \left( \frac{1 - \tilde{p}}{\tilde{p}} \right)^{2\kappa_C} \left( \frac{p}{1 - p} \right)^{\kappa_C} \right)
\end{align*}
\]

where \( C_2^T, C_3^T, C_4^T, C_5^T, C_1^C, C_3^C, C_4^C \) and \( C_6^C \) are undetermined coefficients and:

\[
\begin{align*}
n &= -\lambda \frac{\delta + \eta \lambda + \eta \gamma}{\gamma (\delta + \lambda + \gamma)} \\
d &= -\frac{\lambda}{(\delta + \lambda + \gamma) B(a, \psi)} \\
c_C &= \frac{\left( \frac{a_G^2 - a_R^2}{\sigma_C} \right)^2 (\lambda + \gamma)}{2\gamma (\delta + \lambda + \gamma)} \\
\kappa_C &= \sqrt{\frac{1}{4} + \frac{2 (\delta + \gamma + \eta \lambda)}{\left( \frac{a_G^2 - a_R^2}{\sigma_C} \right)^2 \left( \frac{\sigma_C}{a_G^2} \right)^2}} \\
m_C &= -n \left( \frac{1 - p}{\tilde{p}} \right)^{2\kappa_C} \\
q_C &= \frac{1}{2} - \frac{1}{2} \sqrt{\frac{8 (\delta + \lambda + \gamma) \gamma}{\left( \frac{a_G^2 - a_R^2}{\sigma_C} \right)^2 (\lambda + \gamma)} + 1}
\end{align*}
\]
Moreover, for the case where \( p \in (\bar{p}, 1] \), we have:

\[
\begin{align*}
  f_C'(p) &= C_1^C (q_c - 2 + 3p) p^{q_c - 3} (1 - p)^{2 - q_c} \\
          &\quad \frac{d}{\sqrt{C_C (4 + C_C)}} (q_c - 2 + 3p) p^{q_c - 3} (1 - p)^{2 - q_c} \int_1^p \tau^{a - q_c} (1 - \tau)^{q_c + \psi - 1} d\tau \\
          &\quad + \frac{d}{\sqrt{C_C (4 + C_C)}} (-1 - q_c + 3p) p^{2 - q_c} (1 - p)^{q_c - 3} \int_1^p \tau^{q_c + a - 1} (1 - \tau)^{\psi - q_c} d\tau
\end{align*}
\]

For reasonable parameter values, it turns out that the first term of the three on the right hand side is the quantitatively most important. Indeed, if \( C_1^C \) turns out to be positive (which is true for reasonable parameter values), then \( f_C(\cdot) \) will be decreasing in \( p \) in that interval, if \( \gamma > \left( \frac{a^g - a^R}{\sigma_C} \right)^2 \) or approximately \( \gamma > \left( \frac{a^g - a^R}{\sigma_T} \right)^2 \).

Similarly one obtains that \( f_T(\cdot) \) will be increasing in \( p \) in the interval \([0, \bar{p}]\) if \( \gamma > \left( \frac{a^g - a^R}{\lambda + \gamma} \right)^2 \) or approximately \( \gamma > \left( \frac{a^g - a^R}{\sigma_T} \right)^2 \).

Given the above restrictions, the shape of the distribution of posteriors for the employed workers in both occupations is unimodal, skewed and features a fat right tail (fat left tail for the case of occupation \( T \)).

Using the 8 conditions mentioned in the main text we can solve for the values of the remaining 8 undetermined coefficients. They are given by the solution to the linear system, which can be written in matrix form as follows:

\[ A \times C = B \]

where:

\[
c_T = \frac{\left( \frac{a^g - a^R}{\sigma_T} \right)^2 (\lambda + \gamma)}{2\gamma (\delta + \lambda + \gamma)} \\
\kappa_T = \sqrt{\frac{1}{4} + \frac{2 (\delta + \gamma + \eta \lambda)}{(\frac{a^g - a^R}{\sigma_C})^2}} \\
m_T = -n \left( \frac{p}{1 - p} \right)^{2\kappa_T} \\
q_T = \frac{1}{2} - \frac{1}{2} \sqrt{\frac{8 (\delta + \lambda + \gamma) \gamma}{(\frac{a^g - a^R}{\sigma_T})^2 (\lambda + \gamma)}} + 1
\]
\[
A = \begin{bmatrix}
a_{11} & a_{12} & a_{13} & 0 & 0 & 0 & 0 & a_{18} \\
0 & 0 & 0 & a_{24} & a_{25} & a_{26} & a_{27} & 0 \\
0 & a_{32} & a_{33} & a_{34} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & a_{46} & a_{47} & a_{48} \\
0 & 0 & 0 & a_{54} & a_{55} & a_{56} & a_{57} & 0 \\
a_{61} & a_{62} & a_{63} & 0 & 0 & 0 & 0 & a_{68} \\
0 & 0 & 0 & 0 & 0 & a_{76} & a_{77} & a_{78} \\
0 & a_{82} & a_{83} & a_{84} & 0 & 0 & 0 & 0
\end{bmatrix}
\]

\[
C' = \begin{bmatrix}
C_T^2 & C_T^3 & C_T^4 & C_T^5 & C_1^C & C_3^C & C_4^C & C_6^C
\end{bmatrix}
\]

\[
B' = \begin{bmatrix}
b_1 & b_2 & 0 & 0 & b_5 & b_6 & 0 & 0
\end{bmatrix}
\]

and:

\[
a_{11} = p^{-1-q_T}(1-p)^{q_T-2}
\]
\[
a_{12} = -p^{q_T-2}(1-p)^{-1-q_T}
\]
\[
a_{13} = -a_{11}
\]
\[
a_{18} = \frac{1}{\sqrt{c_T(4+c_T)}}
\left[
p^{q_T-2}(1-p)^{-1-q_T}(m_T \int_{\tilde{p}}^{p} \tau^{-\frac{1}{2}+\kappa_T-q_T}(1-\tau)^{-\frac{3}{2}+\kappa_T+q_T} d\tau
\right.
+n \int_{\tilde{p}}^{p} \tau^{-\frac{1}{2}+\kappa_T-q_T}(1-\tau)^{-\frac{3}{2}+\kappa_T+q_T} d\tau)
\left. -p^{-1-q_T}(1-p)^{q_T-2}(m_T \int_{\tilde{p}}^{p} \tau^{-\frac{3}{2}+\kappa_T+q_T}(1-\tau)^{-\frac{1}{2}+\kappa_T-q_T} d\tau
\right]
+n \int_{\tilde{p}}^{p} \tau^{-\frac{3}{2}+\kappa_T+q_T}(1-\tau)^{-\frac{1}{2}+\kappa_T-q_T} d\tau
\]

\[
b_1 = \frac{d}{\sqrt{c_T(4+c_T)}}
\left[
p^{q_T-2}(1-p)^{-1-q_T}(m_T \int_{\tilde{p}}^{p} \tau^{a-q_T}(1-\tau)^{q_T+\psi-1} d\tau
\right.
-n \int_{\tilde{p}}^{p} \tau^{a+q_T-1}(1-\tau)^{\psi+q_T} d\tau)
\left. -p^{-1-q_T}(1-p)^{q_T-2}(m_T \int_{\tilde{p}}^{p} \tau^{a-q_T}(1-\tau)^{\psi+q_T-1} d\tau
\right]
+n \int_{\tilde{p}}^{p} \tau^{a+q_T-1}(1-\tau)^{\psi-q_T} d\tau
\]

44
\[ a_{24} = \frac{1}{\sqrt{c_c (4 + c_c)}} [p^{q_c - 2} (1 - p)^{-1 - q_c} m_c \int_{\hat{p}}^{p} \tau^{\frac{1}{2} + \kappa_c - q_c} (1 - \tau)^{-\frac{3}{2} - \kappa_c + q_c} d\tau + n \int_{\hat{p}}^{p} \tau^{\frac{1}{2} - \kappa_c - q_c} (1 - \tau)^{-\frac{3}{2} + \kappa_c + q_c} d\tau] - p^{-1 - q_c} (1 - \hat{p})^{q_c - 2} (m_c \int_{\hat{p}}^{p} \tau^{\frac{3}{2} + \kappa_c + q_c} (1 - \tau)^{-\frac{1}{2} - \kappa_c - q_c} d\tau + n \int_{\hat{p}}^{p} \tau^{\frac{3}{2} - \kappa_c + q_c} (1 - \tau)^{-\frac{1}{2} + \kappa_c - q_c} d\tau)] \]

\[ a_{25} = p^{q_c - 2} (1 - p)^{-1 - q_c} \]

\[ a_{26} = -a_{25} \]

\[ a_{27} = -p^{-1 - q_c} (1 - \hat{p})^{q_c - 2} \]

\[ b_2 = \frac{d}{\sqrt{c_c (4 + c_c)}} [p^{q_c - 2} (1 - p)^{-1 - q_c} m_c \int_{1}^{p} \tau^{a - q_c} (1 - \tau)^{q_c + \psi - 1} d\tau - \hat{p}^{-1 - q_c} (1 - \hat{p})^{q_c - 2} \int_{1}^{p} \tau^{q_c + a - 1} (1 - \tau)^{\psi - q_c} d\tau + \hat{p}^{-1 - q_c} (1 - \hat{p})^{q_c - 2} \int_{\hat{p}}^{p} \tau^{a - q_c} (1 - \tau)^{\psi + q_c - 1} d\tau + p^{-1 - q_c} (1 - p)^{q_c - 2} \int_{\hat{p}}^{p} \tau^{a + q_c - 1} (1 - \tau)^{\psi - q_c} d\tau] \]

\[ a_{32} = \hat{p}^{q_c - 2} (1 - \hat{p})^{-1 - q_c} \]

\[ a_{33} = \hat{p}^{-1 - q_c} (1 - \hat{p})^{q_c - 2} \]

\[ a_{34} = \left(\frac{1 - \hat{p}}{\hat{p}}\right)^{2q_c} \hat{p}^{\frac{3}{2} + \kappa_c} (1 - \hat{p})^{-\frac{3}{2} - \kappa_c} - \hat{p}^{\frac{3}{2} - \kappa_c} (1 - \hat{p})^{-\frac{3}{2} + \kappa_c} \]

\[ a_{46} = \hat{p}^{q_c - 2} (1 - \hat{p})^{-1 - q_c} \]

\[ a_{47} = \hat{p}^{-1 - q_c} (1 - \hat{p})^{q_c - 2} \]

\[ a_{48} = \left(\frac{p}{1 - p}\right)^{2q_c} \hat{p}^{\frac{3}{2} - \kappa_T} (1 - \hat{p})^{-\frac{3}{2} + \kappa_T} - \hat{p}^{\frac{3}{2} + \kappa_T} (1 - \hat{p})^{-\frac{3}{2} - \kappa_T} \]
\[ a_{54} = \frac{1}{2} \left( \frac{a_C^G - a_C^R}{\sigma_C} \right)^2 \frac{1}{\sqrt{c_C(4 + c_C)}} \cdot \\
\left( -(1 - q_C + 3\overline{p})\overline{p}^{-q_C} (1 - \overline{p})^{q_C-1} \right) \]
\[ \cdot [m_C \int_{\overline{p}}^{p} \tau^{-\frac{1}{2} + \kappa_C + q_C} (1 - \tau)^{-\frac{1}{2} - \kappa_C - q_C} d\tau + n \int_{\overline{p}}^{p} \tau^{-\frac{1}{2} - \kappa_C + q_C} (1 - \tau)^{-\frac{1}{2} + \kappa_C - q_C} d\tau] \]
\[ - (q_C - 2 + 3\overline{p})\overline{p}^{q_C-1} (1 - \overline{p})^{-q_C} \]
\[ \cdot [m_C \int_{\overline{p}}^{p} \tau^{-\frac{1}{2} + \kappa_C - q_C} (1 - \tau)^{-\frac{1}{2} - \kappa_C + q_C} d\tau + n \int_{\overline{p}}^{p} \tau^{-\frac{1}{2} - \kappa_C - q_C} (1 - \tau)^{-\frac{3}{2} + \kappa_C + q_C} d\tau] \]
\[ - \frac{\lambda}{\lambda + \gamma} \left( \frac{a_T^G - a_T^R}{\sigma_T} \right)^2 \kappa_C \overline{p}^{-\frac{1}{2} - \kappa_C} (1 - \overline{p})^{-\frac{1}{2} + \kappa_C} \]
\[ a_{55} = -\frac{1}{2} \left( \frac{a_C^G - a_C^R}{\sigma_C} \right)^2 (q_C - 2 + 3\overline{p})\overline{p}^{q_C-1} (1 - \overline{p})^{-q_C} \]
\[ a_{56} = -a_{55} \]
\[ a_{57} = \frac{1}{2} \left( \frac{a_C^G - a_C^R}{\sigma_C} \right)^2 (-1 - q_C + 3\overline{p})\overline{p}^{-q_C} (1 - \overline{p})^{q_C-1} \]

\[ b_5 = -\frac{1}{2} \left( \frac{a_C^G - a_C^R}{\sigma_C} \right)^2 \frac{d}{\sqrt{c_C(4 + c_C)}} (-1 - q_C + 3\overline{p})\overline{p}^{-q_C} (1 - \overline{p})^{q_C-1} \int_{\overline{p}}^{1} \tau^{a + q_C - 1} (1 - \tau)^{\psi - q_C} d\tau \]
\[ + \frac{1}{2} \left( \frac{a_C^G - a_C^R}{\sigma_C} \right)^2 \frac{d}{\sqrt{c_C(4 + c_C)}} (q_C - 2 + 3\overline{p})\overline{p}^{q_C-1} (1 - \overline{p})^{-q_C} \int_{\overline{p}}^{1} \tau^{a - q_C} (1 - \tau)^{\psi + q_C - 1} d\tau \]
\begin{align*}
a_{61} &= -\frac{1}{2} \left( \frac{a_T^G - a_T^R}{\sigma_T} \right)^2 \left( -1 - q_T + 3p \right) p^{-q_T} (1 - p)^{q_T - 1} \\
a_{62} &= \frac{1}{2} \left( \frac{a_T^G - a_T^R}{\sigma_T} \right)^2 \left( q_T - 2 + 3p \right) p^{q_T - 1} (1 - p)^{-q_T} \\
a_{63} &= -a_{61} \\
a_{68} &= \frac{1}{2} \left( \frac{a_T^G - a_T^R}{\sigma_T} \right)^2 \frac{1}{\sqrt{c_T (4 + c_T)}} \cdot \\
&\left( \left(-1 - q_T + 3p\right) p^{-q_T} (1 - p)^{q_T - 1} \right)^2 \\
&\cdot \left[ m_T \int_{\tilde{p}}^{p} \tau^{-\frac{1}{2} - \kappa_T} (1 - \tau)^{-\frac{1}{2} + \kappa_T - q_T} d\tau + n \int_{\tilde{p}}^{p} \tau^{-\frac{1}{2} + \kappa_T + q_T} (1 - \tau)^{-\frac{1}{2} - \kappa_T - q_T} d\tau \right] \\
&- \left( q_T - 2 + 3p \right) p^{q_T - 1} (1 - p)^{-q_T} \\
&\cdot \left[ m_T \int_{\tilde{p}}^{p} \tau^{-\frac{1}{2} - \kappa_T - q_T} (1 - \tau)^{-\frac{1}{2} + \kappa_T + q_T} d\tau + n \int_{\tilde{p}}^{p} \tau^{-\frac{1}{2} + \kappa_T - q_T} (1 - \tau)^{-\frac{1}{2} - \kappa_T + q_T} d\tau \right] \\
&+ \frac{\lambda}{\lambda + \gamma} \left( \frac{a_T^G - a_T^R}{\sigma_C} \right)^2 \kappa_T \tilde{p}^{-\frac{1}{2} + \kappa_T} (1 - p)^{-\frac{1}{2} - \kappa_T} \\
b_6 &= \frac{1}{2} \left( \frac{a_T^G - a_T^R}{\sigma_T} \right)^2 \frac{d}{\sqrt{c_T (4 + c_T)}} \left( -1 - q_T + 3p \right) p^{-q_T} (1 - p)^{q_T - 1} \int_{0}^{\tilde{p}} \tau^{a + q_T - 1} (1 - \tau)^{\psi - q_T} d\tau \\
&- \frac{1}{2} \left( \frac{a_T^G - a_T^R}{\sigma_T} \right)^2 \frac{d}{\sqrt{c_T (4 + c_T)}} \left( q_T - 2 + 3p \right) p^{q_T - 1} (1 - p)^{-q_T} \int_{0}^{\tilde{p}} \tau^{a - q_T} (1 - \tau)^{q_T + \psi - 1} d\tau \\
a_{76} &= \left( q_C - 2 + 3\tilde{p} \right) p^{q_C - 3} (1 - \tilde{p})^{-q_C - 2} \\
a_{77} &= \left( -1 - q_C + 3\tilde{p} \right) \tilde{p}^{2 - q_C} (1 - \tilde{p})^{q_C - 3} \\
a_{78} &= \left( \frac{\tilde{p}}{1 - \tilde{p}} \right)^{2\kappa_T} \tilde{p}^{-\frac{5}{2} - \kappa_T} (1 - \tilde{p})^{-\frac{5}{2} + \kappa_T} \left( -\frac{3}{2} - \kappa_T + 3\tilde{p} \right) \\
&- \tilde{p}^{-\frac{5}{2} + \kappa_T} (1 - \tilde{p})^{-\frac{5}{2} - \kappa_T} \left( -\frac{3}{2} + \kappa_T + 3\tilde{p} \right)
\end{align*}
\[ a_{82} = (q_T - 2 + 3\hat{p}) \hat{p}^{q_T - 3} (1 - \hat{p})^{-q_T - 2} \]
\[ a_{83} = (-1 - q_T + 3\hat{p}) \hat{p}^{-q_T} (1 - \hat{p})^{q_T - 3} \]
\[ a_{84} = \left( \frac{1 - \hat{p}}{\hat{p}} \right)^{2\kappa_C} \hat{p}^{-\frac{5}{2} + \kappa_C} (1 - \hat{p})^{-\frac{5}{2} - \kappa_C} \left( -\frac{3}{2} + \kappa_C + 3\hat{p} \right) \]
\[ -\hat{p}^{-\frac{5}{2} - \kappa_C} (1 - \hat{p})^{-\frac{5}{2} + \kappa_C} \left( -\frac{3}{2} - \kappa_C + 3\hat{p} \right) \]
References


