

Large Firms and Internal Labor Markets

Theodore Papageorgiou*

Penn State University, Department of Economics[†]

May 2011

Abstract

This paper studies the interaction between the size of a firm and the labor market outcomes of its workers. In our model, firms are internal labor markets where workers are matched with occupations. The quality of matches is uncertain and learned over time. Larger firms offer more opportunities to workers to find a suitable occupational match. In equilibrium, workers in larger firms are employed in better matches and earn higher wages. Conditional on wages, they are less likely to separate from the firm, but more likely to switch occupations within the firm, while the wage premium is higher for workers with longer tenure. We find support for the implications of the model using data from the SIPP.

Keywords: Size-Wage Premium, Internal Labor Markets, Occupational Mobility, Separation Rates

JEL Classification: E24, J24, J31

*I am grateful to Joe Altonji, Dimitris Cheliotis, Manolis Galenianos, Ed Green, William Hawkins, Alex Monge, Giuseppe Moscarini, Andrés Rodríguez-Clare, Richard Rogerson, Russell Toth, Neil Wallace, Ruilin Zhou, as well as seminar participants at Yale, Rochester, Wisconsin-Madison, Wesleyan, New York Fed, the Search and Matching Workshop at the University of Pennsylvania, the Cornell-Penn State Macro Workshop, the 2010 SED Annual Meeting and the 2011 SOLE Annual Meeting for useful comments.

[†]418 Kern Building, University Park, PA 16802, USA. E-mail: tpapageorgiou@psu.edu. Web: <http://sites.google.com/site/theodorepapageorgiou/>

1 Introduction

It is well-known that workers in larger firms have higher wages and separate less often. The size effect is persistent and remains even after one controls for worker selection and firm characteristics (Brown and Medoff (1989)). A lesser-known fact is that almost a third of worker reallocation takes place within rather than across firms: every year 8% of all workers switch occupations within their firm. This is especially true for workers in larger firms: workers in firms with more than 100 employees are 61% more likely to switch occupations within firms compared to those in firms with less than 25 employees (see Table 1).

This paper links the reallocation of workers across occupations but within firms, and the reallocation of workers across both occupations and firms. The implied interaction between firm size and occupational choice provides a natural mechanism to account for the wage premium and the differences in reallocation rates. We introduce a theoretical model of internal labor markets and empirically investigate its implications regarding the wage premium, within firm mobility and separation probabilities. The model correctly predicts that conditional on wages, workers in larger firms are *less* likely to switch occupations across firms, but *more* likely to switch occupations within firms. Moreover, the wage premium is higher for workers with longer tenure.

In our setup, a firm constitutes a labor market without search frictions and workers can switch tasks or occupations *within* a firm at no cost. Switches *across* firms, however, are costly due to search frictions. The quality of a match between a worker and a task or occupation is uncertain and revealed gradually. Workers need to experiment in order to find a suitable match. They observe output realizations and update their beliefs in Bayesian fashion. At any point a worker has the option of leaving his current task and going to another one within his firm or to unemployment.

The optimal behavior of the worker is one of a reservation strategy: if his belief regarding the quality of his match with his current task or occupation falls below a certain threshold, the worker moves on to another one; this threshold is increasing in the number of tasks within the firm that the worker has not tried. Workers in larger firms have more tasks remaining and therefore become selective and willing to abandon unpromising matches more easily, unlike workers in small firms who have limited options. In equilibrium, workers in large firms are better matched; this leads to higher wages and lower separation rates.

The proposed theory is consistent with numerous observed empirical regularities: workers in larger firms are more productive (Idson and Oi (1999)), have higher wages

| Firm Employees: | Annual Switching Probability (Unconditional) |
|-----------------|--|
| >100 | 8.85% |
| 25-99 | 7.3% |
| <25 | 5.48% |

Table 1: Annual Percentage of Workers who Switch 3-digit Occupations without Changing Firms. 1996 Panel of the Survey of Income and Program Participation.

(Moore (1911), Brown and Medoff (1989)) and their initial wages are also higher (Barron et al. (1987)). In addition, the theory implies that workers in larger firms have more options within the firm due to the multitude of tasks, which results in a lower separation probability, even conditional on their current wage (Brown and Medoff (1989), Idson (1993)).

In the empirical section, we investigate our setup’s implications, by looking at a particular type of match, namely occupations.¹ Using the 1996 panel of the Survey of Income and Program Participation (SIPP), we document a number of novel empirical facts. For instance, we find that conditional on wages, workers in larger firms are more likely to internally switch occupations. In our framework, workers in large firms are indeed more likely to switch than workers in small firms earning the same wage, as they have more options within their firm. In addition, we document that switching occupations within a firm is associated with significant wage gains, as predicted by our setup: a worker abandons an unpromising match in order to move to an occupation that is expected to be a better match.

Moreover, we document that the size-wage premium is higher for workers with longer firm tenure. In our setup, finding a suitable match takes time, so the workers who have been with a large firm the longest have fully reaped the benefits. As a result, the size-wage premium is increasing with firm tenure. We also find that initial wages are higher for workers in larger firms, consistent with the predictions of our framework.

Our empirical exercise also confirms the documented differences in separation rates: the annual separation probability is 8.5% lower for workers in larger firms. This difference in separation rates, cannot be solely explained by the higher wages in larger firms. Indeed,

¹Our empirical exercise explores model implications regarding occupations. A recent empirical literature (see Kambourov and Manovskii (2009a), Sullivan (2010), Groes (2010)) documents the importance of occupational matching in wage formation. However our theory is more general as to what a match may constitute. Besides occupations, it could refer to a worker’s fit within a department, a product division or a team and how well he can collaborate with his co-workers there. One may also think of it as a location match as well, with larger firms having more locations.

we find that when controlling for wages, workers in larger firms are *still* almost 6% less likely to separate than workers in smaller ones. Our model is consistent with both results.

This last implication distinguishes our setup from existing labor market models of the wage premium and separation rates. For instance, Burdett and Mortensen's (1998) or Burdett and Coles's (2003) models similarly predict that workers in larger firms are paid higher wages and are therefore less likely to separate; but once one controls for the wage, firm size no longer affects the separation probability.² In contrast, our setup can naturally replicate this feature of the data. Workers in larger firms form better matches and this leads to lower separation rates. But even conditional on the wage- which here captures the quality of the worker's match- workers in large firms are less likely to separate, as they have more options within their firm. Indeed, if their match proves unpromising they will exhaust other tasks in their firm before separating to unemployment. The larger number of available occupations is also consistent with higher internal mobility of workers conditional on wages, which as discussed above, is indeed confirmed in the data.

In order to further investigate the differences in separation rates, we distinguish between the probability of separating and switching occupations versus the probability of separating and not switching occupations. In the first case, we find that both wage and firm size have large negative effects on the probability of separating *and* switching occupations. In the second case, we find that only firm size has a significant impact on the probability of separating but remaining in the same occupation. This is consistent with our setup which argues that workers who remain in the same occupation care only about the size of the internal labor market, since their occupational match, captured here by the wage, remains the same. Put differently, since the quality of the worker's occupational match does not change if he remains in the same occupation, he only switches occupations if the new firm has more occupations available.

The present paper is consistent with a large empirical literature that documents the observed size-wage premium. Brown and Medoff (1989), Idson and Feaster (1990) and Schmidt and Zimmermann (1991) find that the positive relationship between firm size and wage holds even conditional on observed and unobserved labor quality and therefore cannot be explained away by worker selection. Brown and Medoff (1989) and Schmidt and Zimmermann (1991) show that the size premium persists when controlling for a

²Postel-Vinay and Robin's (2002) framework also predicts that workers in more productive firms separate less often, even conditional on wages. In their setup, workers in more productive firms accept a lower initial wage in exchange for a steeper wage path. However in the data, initial wages are in fact significantly higher in larger firms (see Barron et al. (1987), as well as Figure 1 of the present paper). Moreover, Postel-Vinay and Robin's (2002) setup does not provide any implications regarding mobility within the firm.

multitude of firm characteristics, such as working conditions, while Brown and Medoff (1989) demonstrate that it remains even when considering firms that offer a piece-rate system. In a related paper, Idson (1993) investigates empirically labor turnover inside the firm and argues that larger firms provide more firm-specific training to their workers. Novos (1992) investigates worker turnover and firm scope in the presence of learning-by-doing and asymmetric information. Moscarini and Thomsson (2006), using data from the CPS, document that a significant fraction of occupational switches do not involve an employer switch. For a survey of the literature on internal labor markets see Gibbons and Waldman (1999).^{3,4}

Our paper is also related to Neal (1999) who argues that workers follow a two-stage search strategy: first they search for a career (occupation), and then shop for jobs within the chosen career. In Neal's setup however workers cannot search for another occupation/career within an employer.

The current paper also follows a growing literature of theoretical models that underline the importance of occupational matching for workers (Kambourov and Manovskii (2009b), Antonovics and Golan (2010), Papageorgiou (2010), Groes, Kircher and Manovskii (2010), Eechkout and Weng (2010), Alvarez and Shimer (2009, 2011)). In the above labor market models the role of firms is of lesser importance, somewhat contrary to conventional wisdom, as well as the documented differences in wages and separation rates between workers in firms of different sizes. A contribution of the present paper is that, while consistent with the above literature in emphasizing the importance of occupational matching, it introduces a non-trivial role for firms in the labor market: firms act as institutions that alleviate frictions to occupational matching.

The next section describes the economic environment, while Section 3 solves for worker behavior. Section 4 derives the model's implications regarding productivity and wages, while Section 5 discusses some of the model's assumptions. Section 6 uses data from the 1996 SIPP and examines the empirical evidence, as well as derives further implications regarding the wage premium, internal mobility, and the probability a worker separates from his firm. We offer some concluding remarks in Section 7.

³See also Pastorino (2009) for a more recent contribution.

⁴The idea of match uncertainty between a worker and a firm dates back to Jovanovic (1979), while the setup in this paper is closer to Moscarini (2005). In the current paper however, a firm may present more than one potential match for the worker.

2 The Economy

Time is continuous. There is a population of firms of fixed mass. Each firm is characterized by the number of its occupations or tasks, $m \in \{1, 2, \dots, M\}$.⁵ The distribution of tasks is exogenous and let q_m denote the fraction of firms with m tasks.

There is a fixed mass of workers. Workers are risk neutral and have discount rate $r > 0$. A worker can be either employed or unemployed. All unemployed workers are identical.

When employed, a worker works in one task at a time. When he begins employment at a firm with m tasks, he randomly picks one of those m tasks which are ex ante identical. The flow output for worker i , in task j , in firm l at time t is given by:

$$dY_t^{ijl} = \alpha^{ijl} dt + \sigma dW_t^{ijl}$$

where dW_t^{ijl} is the increment of a Wiener process, $\alpha^{ijl} \in \{\alpha^G, \alpha^B\}$ is mean output per unit of time and $\sigma > 0$. Assume that, without loss of generality, $\alpha^G > \alpha^B$ and that productivities, α^{ijl} , are independently distributed across tasks, firms and workers.

Furthermore, assume α^{ijl} is unknown. Before starting work at a task, a worker draws his productivity parameter, α^{ijl} , from a Bernoulli distribution, with parameter p_0^{ijl} , the probability that $\alpha^{ijl} = \alpha^G$. p_0^{ijl} is common knowledge and is distributed according to a known distribution $G(\cdot)$ with support $[0, 1]$.

At any point a worker can leave his current task and go to another task or to unemployment. There are no restrictions in the number of workers employed in a task, i.e. there are no congestion externalities. Once a worker leaves a task, he cannot return to it.⁶ A worker separates from his firm either endogenously, or exogenously according to a Poisson process with parameter $\delta > 0$, in which case he becomes unemployed. There is no on-the-job search in the baseline model; we introduce on-the-job search in Section 6.3.

Following the literature, we assume that the wage is determined by generalized Nash bargaining between the firm and each worker separately, with $\beta \in (0, 1)$ denoting the worker's bargaining power.

When unemployed, a worker needs to search for a firm. Search is undirected and an

⁵We use the terminology “tasks”, but the reader may choose to think of these as departments, teams, locations etc.

⁶This is an implicit cost to switching. The paper's main results go through without this assumption (see for instance the model in the Online Appendix); however if occupational switching is completely costless, then a worker on the margin will be switching between occupations multiple times within a time period.

unemployed worker meets a firm according to a Poisson process with parameter $\lambda > 0$. An unemployed worker earns $b > \alpha^B$.⁷

The sequence of actions is the following: an unemployed worker meets a firm at rate λ and observes the number of its tasks, m . He then draws his prior for the first task from $G(\cdot)$ and either begins working in that task or chooses to move on to the another task and draws a new prior there. After beginning work at a task, he observes output realizations and as will be shown, updates his beliefs regarding the quality of his match with the task. At any point he has the option of leaving his current task and going to another one within his firm or to unemployment.

In what follows we solve for the optimal behavior of a worker, i.e. when does he leave his current task and move to another one and when does he quit to unemployment. We focus on the equilibrium cross-sectional distribution of workers and discuss the model's implications regarding wages, internal mobility and separation rates.

3 Worker Behavior

Workers observe their output and obtain information regarding the quality of their match in that specific task. Let p_t^{ijl} denote the posterior probability that the match of worker i with task j in firm l is good, i.e. $\alpha^{ijl} = \alpha^G$. In particular, a worker observes his flow output, dY_t^{ijl} , and updates p_t^{ijl} , according to (Liptser and Shyryaev (1977)):

$$dp_t^{ijl} = p_t^{ijl} (1 - p_t^{ijl}) \zeta \frac{dY_t^{ijl} - \left(p_t^{ijl} \alpha^G + (1 - p_t^{ijl}) \alpha^B \right) dt}{\sigma} \quad (1)$$

where $\zeta = \frac{\alpha^G - \alpha^B}{\sigma}$. The last term on the right hand side is a standard Wiener process with respect to the unconditional probability measure used by the agents. To minimize notation, from now on, we drop the t subscript, as well as the i , j and l superscripts.

The beliefs regarding the quality of the worker-task match follow a Bernoulli distribution. The posterior probability is thus, a sufficient statistic of the worker's beliefs and a state variable for his value function. Besides their employment status and beliefs regarding their current task, workers also differ in their opportunities to work in other tasks within their firm. The number of remaining tasks, k , available to the worker in his current firm, therefore, also constitutes a state variable for the value of an employed worker.⁸

⁷To avoid a situation where a worker never chooses to be employed, we assume $1 - G\left(\frac{b - \alpha^B}{\alpha^G - \alpha^B}\right) > 0$, which implies that $\alpha^G > b$.

⁸Note that the total number of tasks in the worker's current firm is not relevant, as employment

Given the process for the evolution of beliefs (equation (1)), the Hamilton-Jacobi-Bellman equation of a worker employed in a task with posterior probability p of being in a good match, in a firm where $k \geq 0$ other tasks are available to work in, is given by:

$$rV(p, k) = w(p) + \frac{1}{2}\zeta^2 p^2 (1-p)^2 V_{pp}(p, k) - \delta(V(p, k) - U) \quad (2)$$

where $V(\cdot)$ is his value, $V_{pp}(\cdot)$ is the second derivative of $V(\cdot)$ with respect to p and $w(\cdot)$ is his wage, which as shown below, depends only on p . The flow benefit of the worker consists of his flow wage, plus a term capturing the option value of learning, which allows him to make informed decisions in the future. Finally, the worker loses his job in his current firm at rate δ and becomes unemployed.

Similarly, the flow value of an unemployed worker is then given by:

$$rU = b + \lambda \left(\sum_{m=1}^M q_m E_{p_0} V(p_0, m-1) \right) - \lambda U \quad (3)$$

where $E_{p_0} V(p_0, k)$ is the expected value of a worker, with k other tasks remaining, who is about to draw his prior, p_0 , for his current task.⁹

The solution to the generalized Nash bargaining problem results in the linear sharing rule:

$$\beta J(p, k) = (1 - \beta)(V(p, k) - U) \quad (4)$$

where $J(p, k)$ is the value to the firm of employing a worker with posterior p and k tasks remaining. As shown in Appendix A, using the above equation, we obtain:

Lemma 1 *The worker's wage is given by:*

$$w(p) = \beta \bar{\alpha}(p) + (1 - \beta) rU \quad (5)$$

where $\bar{\alpha}(p) = p\alpha^G + (1-p)\alpha^B$ is the worker's expected output.

Note that the wage is an affine transformation of the posterior, p and does not depend on k : as in the model of Mortensen and Pissarides (1994), wages only depend on the

history in previous tasks does not affect the worker's current or future payoffs.

⁹One can write:

$$E_{p_0} V(p_0, k) = \int_{\underline{p}(k)}^1 V(p_0, k) g(p_0) dp_0 + G(\underline{p}(k)) E_{p_0} V(p_0, k-1)$$

where $\underline{p}(k)$ is defined below.

current level of output and not on its future path.¹⁰

If we increase the number of tasks available to the worker, then the worker cannot be worse off, as he has the option of not trying them out, so $V(\cdot)$ is non-decreasing in k . Moreover, it cannot be the case that $V(p, k+1) = V(p, k)$, as that would imply that a worker can forever ignore one task and obtain the same value. This can not be true, as the option value of trying out one more task is positive, hence:¹¹

Lemma 2 *The value of an employed worker, $V(\cdot)$, is increasing in k .*

We look for equilibria such that $V(\cdot)$ is an increasing function of p . A worker chooses when to leave his current task and move on to either another task, or to unemployment.

Take the case of a worker with $k \geq 1$. The solution to the optimal stopping problem for this worker is given by a trigger $\underline{p}(k)$, such that the following value matching and smooth pasting conditions are satisfied:

$$V(\underline{p}(k), k) = E_{p_0} V(p_0, k-1) \quad (6)$$

and

$$V_p(\underline{p}(k), k) = 0 \quad (7)$$

Note that we allow the threshold, $\underline{p}(k)$, to depend on the number of remaining tasks in the current firm.

A worker who has exhausted all available tasks in his current firm, i.e. $k = 0$, optimally quits to unemployment when:

$$V(\underline{p}(0), 0) = U \quad (8)$$

and

$$V_p(\underline{p}(0), 0) = 0 \quad (9)$$

Note that these triggers also determine whether a worker in a new task who has just drawn his prior, p_0 , begins working there or chooses to move on to the another task or to

¹⁰The wage equation (5) is almost identical to the wage equation in the Mortensen-Pissarides (1994) model (see for instance equations (2.9) and (2.10) on page 42 of Pissarides (2000)).

¹¹Consider a worker with current posterior $p < 1$ and k tasks remaining. Assume $V(\cdot)$ is increasing in p (we show this to be true below) and that one more task becomes available. Consider the following strategy for the worker: if his current task proves unsuccessful, instead of moving on to another task, he draws a prior, p_0 , for the newly-available task and if it's higher than p (which occurs with probability $1 - G(p)$), the worker begins work there instead. In this case, he has k tasks remaining again, but a higher posterior p , so his value must be higher. Since the worker leaves his current task with positive probability, then the option value of trying out one more task is positive.

unemployment.

In Appendix B, we use the boundary conditions (6) through (9) to solve for the value function of a worker. The following proposition states the result:

Proposition 3 *The value of an employed worker, $V(\cdot)$, is increasing in p . A worker optimally leaves his current task when his posterior hits $\underline{p}(k)$, defined as :*

$$\underline{p}(0) = \frac{(\theta - 1)(rU - \alpha^B)}{(\theta - 1)(rU - \alpha^B) + (\theta + 1)(\alpha^G - rU)} \quad (10)$$

for $k = 0$ and:

$$\underline{p}(k) = \frac{(\theta - 1)((r + \delta)E_{p_0}V(p_0, k - 1) - (r + \delta - \beta r)U - \beta a^B)}{\beta(\theta + 1)(\alpha^G - a^B) - 2((r + \delta)E_{p_0}V(p_0, k - 1) - (r + \delta - \beta r)U - \beta a^B)} \quad (11)$$

for $k > 0$, where $\theta = \sqrt{\frac{8(r+\delta)}{\zeta^2} + 1}$ and the value of unemployment is defined in equation (3). Finally a worker who draws a prior, p_0 , below $\underline{p}(k)$ optimally chooses not to work in his new task.¹²

The above proposition, along with Lemma 2, immediately results in the following corollary:

Corollary 4 *The endogenous separations triggers $\underline{p}(k)$, are increasing in k .*

In other words, the better their outside option, the more likely workers are to separate from an unpromising match.

4 Model Implications

The equilibrium evolution of beliefs, employment states and remaining tasks is a positive recurrent process: starting from any posterior and any number of tasks remaining k , $(p, k) \in (\underline{p}(k), 1] \times [0, M - 1]$, the joint process returns to (p, k) infinitely many times, as

¹²A worker quits to unemployment only after exhausting all tasks available to him, i.e. $E_{p_0}V(p_0, 0) > U$. To see this note that if a worker were to never work in the last task then this would imply that $\underline{p}(0) > 1$ (in other words, no matter how high the initial prior, the worker still prefers to be unemployed). However, from eq. (10) we note that because $rU > \alpha^B$ (since $b > \alpha^B$ by assumption) and $\alpha^G > rU$ (since $\alpha^G > b$), then $\underline{p}(0) \in (0, 1)$. This also implies that an unemployed worker draws an initial prior at all firms he contacts, even if $m = 1$.

long as $\delta > 0$.¹³ Therefore there exists a stationary distribution of beliefs and remaining matches.

We show that workers are more productive, on average, in firms with more tasks, m . The proof consists of two steps: we first consider all workers with k tasks remaining and show that average productivity is increasing in k . We next establish that there is a higher percentage of workers that have few tasks remaining (low k), in firms with fewer tasks (low m).

We group tasks according to k , the number of untried tasks available to the employed worker. We consider the steady state distribution of posteriors, p , across all tasks in which the employed worker has k tasks remaining. In the steady state, the flow of workers that exits the distribution, i.e. those hit by an exogenous shock, δ , and those that reach $\underline{p}(k)$, equals the flow of workers into the distribution at various posteriors according to $\frac{g(p_0|p_0 > \underline{p}(k))}{1-G(\underline{p}(k))}$.

We want to show that the cross-sectional posterior mean of all workers with k tasks remaining is increasing in k . Consider the following system:

Let p be a diffusion process that starts at some p_0 , evolves according to equation (1), while at a Poisson rate $\delta > 0$, it returns to p_0 . Finally, let $\underline{p} \in (0, p_0)$, be a reflective boundary, such that when the process hits it, it immediately returns to p_0 .

In Appendix C, we show that the mean value of the process is increasing in i) p_0 and ii) \underline{p} . Intuitively, i) implies that the process starts off and resets at higher point, so one would expect the mean to be higher. Similarly, ii) implies that the higher boundary does not let the process move away from p_0 towards zero and shoots it back to p_0 sooner.

The process in the above system is positive recurrent. Moreover, since shocks are i.i.d. across workers, time averages equal space averages by Birkhoff's Ergodic Theorem, so the mean of the cross-sectional distribution of workers in the above system is increasing in p_0 and \underline{p} .

Since workers with more tasks, k , have both a higher \underline{p} (Corollary 4), as well as a higher p_0 on average ($G(\cdot)$ is truncated at a higher \underline{p}), then, the following lemma is proven in Appendix C:

Lemma 5 *Let $F_k(x)$ be the probability that a worker with k tasks remaining has posterior, p , less or equal to x , i.e. $F_k(x) \equiv \Pr(p \leq x | \text{tasks remaining} = k)$. Then:*

$$\int_{\underline{p}(k')}^1 p f_{k'}(p) dp > \int_{\underline{p}(k)}^1 p f_k(p) dp$$

¹³An unemployed worker also returns to being unemployed infinitely many times.

when $k' > k$.

Moreover, let:

$$\bar{\alpha}_k \equiv \int_{\underline{p}(k)}^1 (p\alpha^G + (1-p)\alpha^B) f_k(p) dp$$

denote the mean output of tasks in which the employed worker has k tasks remaining. Then Lemma 5 immediately implies that:

Lemma 6 *Workers who have more tasks available to work in, have, on average, higher output, i.e.:*

$$\bar{\alpha}_{k'} > \bar{\alpha}_k$$

whenever $k' > k$.

We next show that in firms with many tasks (high m), there is a lower percentage of workers that have few tasks remaining (low k).

Intuitively, consider two firms, one with $m = 50$ and one with $m = 5$. In the first firm it's reasonable to expect that a relatively low percentage of its workforce has $k = 3$ tasks remaining, since $k = 3$ implies that they have exhausted 46 other tasks. In the second firm however, one anticipates a higher percentage of its workers to have $k = 3$ tasks remaining, since workers hired by that firm start off with $k = 4$.

Formally, consider all firms, with m available tasks. Let s_k^m be the share of workers with k tasks remaining, in firms with m tasks in total. The average productivity of firms with m total tasks is given by:

$$\sum_{k=0}^{m-1} s_k^m \bar{\alpha}_k$$

We are interested in how the employment share of task k , s_k^m , changes as m increases. Consider a worker who has just been hired by a firm. Let $\Pr(m-1|m)$ denote the probability that a worker starting off at task m who hasn't drawn his prior, reaches task $m-1$.¹⁴ Then the probability that a worker entering a firm with m tasks, reaches task k is given by:

$$\Pr(m-1|m) \times \Pr(m-2|m-1) \times \dots \times \Pr(k|k+1)$$

¹⁴Ignoring δ shocks, the probability that a worker's posterior belief, p , reaches $\underline{p}(m-1)$, before it reaches 1, unconditionally on match quality, is given by $\frac{1-\underline{p}}{1-\underline{p}(m-1)}$ (Karlin and Taylor (1981)). Given that, we have:

$$\Pr(m-1|m) = G(\underline{p}(m-1)) + (1 - G(\underline{p}(m-1))) \int \frac{1-p_0}{1-\underline{p}(m-1)} \frac{g(p_0|p_0 > \underline{p}(m-1))}{1 - G(\underline{p}(m-1))} dp_0$$

Since $\Pr(j|j+1) \in (0, 1)$ for all j , the probability of a worker reaching task $k \leq m-1$, declines with the number of firm tasks, m (even though the probability of separating from his current task increases in the number of remaining matches). In other words, the expected amount of time a worker spends in tasks greater than k is increasing in m . However, conditional on reaching task k , the expected amount of time he spends in every task from k through 1 has not changed. The expected share of time a worker spends in every task, $k \leq m-1$, while employed in the firm, is therefore decreasing in m . By Birkhoff’s Ergodic Theorem we obtain the following lemma:

Lemma 7 *The steady state firm employment share of every task, $k \leq m-1$, is decreasing in the total number of firm tasks, m .*

Put differently, the distribution of workers across tasks is better, in a first-order stochastic dominance sense, in firms with more tasks. Lemma 7, along with Lemma 6, imply that in firms with a large number of tasks, less workers are employed in tasks where average productivity is low, leading to the following proposition:

Proposition 8 *Average firm productivity, $\sum_{k=0}^{m-1} s_k^m \bar{\alpha}_k$, is increasing in the number of total firm tasks, m .*

Since wages are an affine function of expected output, we have the following corollary:

Corollary 9 *Average wages are higher in firms with more tasks, m .*

5 Discussion

Before turning to the data and looking further at the model’s implications we discuss some of our assumptions. In our setup all movements across occupations are “lateral”, in other words a worker who switches moves to a *different* occupation. However one may worry that in the data some of these movements are in fact “promotions”, in other words, the worker moves to a *better* occupation. In the Online Appendix available on the author’s website¹⁵, we show that in fact the present paper’s predictions regarding wages and separation rates go through in a setup where workers are hierarchically ranked and learn about their unobserved productivity. In that model, workers whose ability is revealed to be better than expected move to a better position or occupation, in other

¹⁵<http://sites.google.com/site/theodorepapegeorgiou/>

words they get promoted. In that setup, workers in larger firms are also better matched: if for example, it's time for them to be promoted, a position is more likely to be available in a larger firm. Put differently, larger firms offer more opportunities for advancement. On the other hand, for workers in smaller firms a position may not always be available and it is possible that they remain mismatched, as is the case in the present paper.

In addition, in our model, the number of occupations available in each firm is exogenously determined. One may imagine that each firm specializes in producing a single product and some products are more complicated than others with their production process requiring more occupations. The goal of this paper is to analyze the mechanism through which workers in larger firms earn higher wages than workers in smaller ones, rather than explain the firm size distribution.

We should also note that the cost of moving to another firm in the above setup is modelled as search frictions. In practice, switching costs may also include other factors such as foregone firm-specific human capital, transferring over one's retirement account, changing health insurance providers etc. These additional costs strengthen the implications of our setup.

Furthermore, our framework implies that working in a large firm is particularly important in locations or time periods where impediments to external mobility are large. Consistent with this, Moscarini and Thomsson (2006) document that during the expansion of the 1990s, occupational mobility across firms increased, while occupational mobility within firms fell significantly.

Moreover, the paper assumes that search is undirected: workers cannot choose which type of firm to seek. This is not restrictive however; one can allow workers to direct their search, as in Moen (1997). In that case, there is a submarket for each firm type and workers who search for the more desirable firms (here those with more tasks), have a lower job finding probability; in equilibrium, unemployed workers are indifferent across submarkets.

Finally, large firms appear to recognize the advantage they enjoy in better matching workers, and attempt to actively exploit it. In particular, they offer "rotation programs", in which employees rotate across different tasks and departments within the firm for a period of time, in order to find their preferred match. According to the 1993 Survey of Employer Provided Training of the Bureau of Labor Statistics, 12% of all establishments report having a job rotation program, while when considering establishments with more than 50 employees this percentage jumps to 24% (Gittleman et al. (1998)). For instance, Freescale Semiconductor's rotational programs description reads: "These programs are

designed for you to define and evaluate your own career path. There are several different rotation paths from which to choose to help you find the opportunity best suited for you.” Similarly, Intel’s Rotation Engineers Program (REP) “helps you learn about your personal strengths and preferences by exposing you to various business groups and technologies in hardware design, software design, manufacturing and marketing...Your REP experience will give you the visibility and understanding of how you best fit within the company.”¹⁶

6 Empirical Evidence and Further Implications

In this section we examine the empirical evidence and derive further implications of our setup. We assume that a worker matches with an occupation and we investigate the model’s implications regarding the wage premium, internal mobility and the probability a worker separates from his firm. We use data from the 1996 panel of the Survey of Income and Program Participation (SIPP). In the 1996 SIPP, interviews were conducted every four months for four years and included approximately 36,000 households. It contains information on the worker’s wage, 3-digit occupation, current employer and employer size (three size categories).^{17,18} Our predictions so far concern the number of occupations or tasks, m , which are unobserved in our data. Following, however, Corollary 13 derived in Section 6.3 below - which shows that in our setup, firms with more occupations have more employees - we use the number of employees to capture firm size.

We first look at the implications of the model regarding the wage premium, as well as its evolution with firm tenure. We next examine internal mobility and the impact of an occupational switch on wages. We conclude by looking into how firm separation probabilities are affected by firm size and wage and also how these dependencies differ depending on whether the worker remains in the same occupation or not.

6.1 Wage Premium

In this section we examine the model’s implications regarding the wage premium.

From Corollary 9, we know that larger firms offer higher wages in equilibrium, so

¹⁶Ortega (2001) and Antonovics and Golan (2010) also argue that these rotation programs are designed to provide information regarding the worker’s fit within the firm.

¹⁷We exclude workers in the armed forces. We use hourly wages deflated to real 1996 dollars using the Consumer Price Index. The 1996 SIPP uses dependent interviewing, which is found to reduce occupational coding error (Hill (1994)).

¹⁸All following results are qualitatively the same if we consider establishments instead of employers. They also do not change if we consider only males less than 35 years old.

| | ln(wage) |
|-----------------|------------------------|
| | |
| ≥ 100 empl | 0.135 |
| | (0.002) ^{***} |
| | |
| 25-99 empl | 0.035 |
| | (0.003) ^{***} |

Table 2: Wage Premium. 1996 Panel of Survey of Income and Program Participation. Controls include gender, race, education, quadratic in age, 11 industry dummies, 13 occupation dummies. 223,748 observations.

our model is consistent with the size-wage premium. Table 2 shows the well-documented wage premium: workers in larger firms earn significantly higher wages, even conditional on observables. This is particularly true for workers in firms with more than 100 employees.

In addition, we obtain the following result:

Proposition 10 *Initial wages are higher in firms with more tasks, m .*

A newly-hired worker will not work in his new task if the prior he draws is less than $\underline{p}(m - 1)$. From Corollary 4, we know that $\underline{p}(m - 1)$ is increasing in m and therefore workers in firms with fewer tasks are more likely to accept a lower prior and therefore a lower wage. On the other hand, workers in firms with high m who choose to reject a task with a low prior realization do so because they expect to draw a higher one in their next task: since $V(p, k)$ is increasing in p and k , a worker who chooses to accept a lower k must, by revealed preference, expect to draw a higher p_0 .

Figure 1, shows the evolution of the wage premium with tenure. We note that initial wages at larger firms are higher than at smaller firms, as implied by Proposition 10 above, but that the difference is equal to a fraction of the overall wage premium. The longer the worker stays with the firm, the greater the difference between his wage and the wage of workers in smaller firms with similar tenure levels.¹⁹

These observations are consistent with our framework's implications: workers in larger firms are more likely to abandon unpromising matches and over time more likely to find an occupation that is a good match. However this process takes time and workers in larger firms only with time will fully reap the benefits of working there.

¹⁹This result casts doubt on the selection hypothesis: if more able individuals are more likely to seek employment in larger firms, one would expect high wage premia for incoming workers. Brown and Medoff (1989) argue that a significant wage premium remains, even after controlling for unobserved heterogeneity.

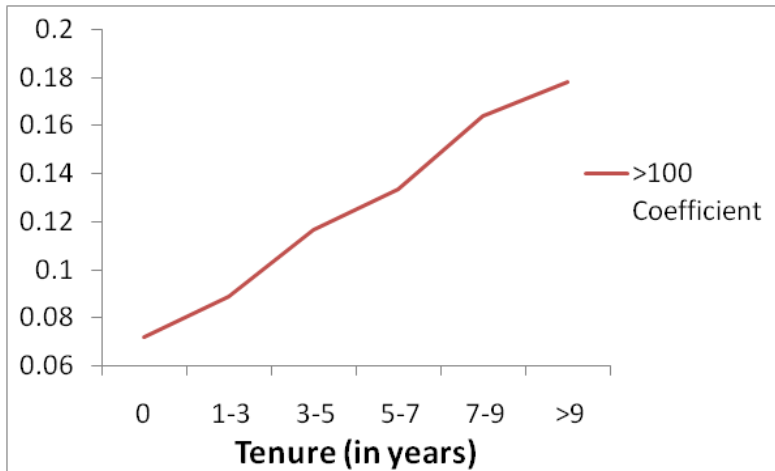


Figure 1: Wage Premium by Tenure Level. 1996 Panel of Survey of Income and Program Participation. Controls include gender, race, education, medium size firm dummy, quadratic in age, 11 industry dummies, 13 occupation dummies. All coefficients significant at one percent.

6.2 Internal Mobility

We next explore the framework’s implication regarding internal mobility. Ignoring δ shocks, the probability a worker with posterior belief p , switches occupations when he has k tasks remaining within the firm is given by (Karlin and Taylor (1981)):

$$\text{Occ Switch Prob} = \frac{1 - p}{1 - \underline{p}(k)}$$

Clearly the probability of an internal switch of occupations is declining in p and therefore the wage, $w(p)$, but increasing in $\underline{p}(k)$. Corollary 4 states that $\underline{p}(k)$ is increasing in the number of remaining tasks, k , while from Lemma 7 we know that workers in high m firms have higher k on average. Thus we have the following proposition:

Proposition 11 *Conditional on wages, workers in firms with more tasks, m , are more likely to switch occupations internally.*

Table 3 shows how wages and firm size affect the probability of switching occupations for workers who remain with the same employer.²⁰ Consistent with our setup, workers with higher wages are less likely to switch, since they have managed to find a good match within the firm. Moreover, conditioning on wages (and therefore match quality), workers

²⁰We use one observation per wave, so we avoid the seam bias (see for instance Nagypal (2008)).

| | |
|----------------------|-------------------------|
| | Occ. Switching |
| | Prob. (Probit) |
| | Same Employer |
| | |
| ≥ 100 employees | 0.0122 |
| | (0.0012) ^{***} |
| ln(wage) | -0.0059 |
| | (0.0011) ^{***} |

Table 3: Wage and Size Impact on Internal Mobility. 1996 Panel of Survey of Income and Program Participation. 4-month probabilities. Controls include gender, race, education, medium size firm dummy, quadratic in age, 11 industry dummies, 13 occupation dummies. Coefficients represent marginal effects evaluated at the average value of the 4-month probability, which equals 0.0281. 170,589 observations.

| | |
|-------------------------|------------------------|
| | ln(wage) |
| | Same Employer |
| | |
| Occ Switch Dummy | 0.024 |
| | (0.002) ^{***} |
| ln(wage) _{t-1} | 0.9 |
| | (0.001) ^{***} |

Table 4: Impact of Occupation Switch on Wage. 1996 Panel of Survey of Income and Program Participation. 4-month intervals. Controls include gender, race, education, quadratic in age, 11 industry dummies, 13 occupation dummies. 160,039 observations.

in larger firms are significantly more likely to switch, since they have better outside options and are therefore more selective.

We next explore the impact of occupational switching on wages. Our model implies that workers who switch occupations within a firm should see their wages increase, as they leave an unproductive match and move on to a (potentially) more productive one. Formally, since $V(p, k)$ is increasing in both p and k , when a worker accepts a lower k , he must, by revealed preference expect a higher p and therefore higher wages, $w(p)$. Indeed, our prediction is consistent with the data: Table 4 reveals that workers switching occupations within a firm experience significant wage gains.²¹

²¹Our framework also implies that workers switching occupations within larger firms are more likely to leave their old occupations at higher wage levels and also obtain higher wages in their new occupations. This is consistent with evidence from the 1996 SIPP. Moreover, internal mobility is decreasing with firm tenure, consistent with the model's implications.

6.3 Separation Probabilities

In this section we examine how the separation rate varies with the number of occupations or tasks, m . Following the discussion in Section 4, the probability that a worker, who just found employment in a firm with m tasks, separates endogenously, is given by (ignoring δ shocks):

$$\Pr(m-1|m) \times \Pr(m-2|m-1) \times \dots \times \Pr(Un|1)$$

where $\Pr(Un|1)$ is the probability a worker starting off in the last task quits to unemployment. Since $\Pr(j|j+1) \in (0, 1) \forall j$, the probability of endogenous separations declines with the number of firm tasks, m . It is clear that the above result does not change if $\delta > 0$, since the exogenous separations rate is independent of m . The following proposition holds:

Proposition 12 *Firms with more tasks, m , have lower separation rates.*

Note that Proposition 12 implies that average tenure in firms with more tasks firms is higher. Thus the following corollary immediately follows:²²

Corollary 13 *Firms with more tasks, m , have more employees.*

The separation probability of two workers with the same posterior p and therefore the same wage, is lower for the one with higher k , based on the derivation above. From Lemma 7, we therefore obtain:

Proposition 14 *Conditional on the wage, the separation probability declines in the total number of firm tasks m .*

We next turn to the empirical evidence. The first column of Table 5 shows that the separation rates are lower for larger firms. This negative relationship between firm size and the separation probability is well-known and consistent with other models in the literature, such as Burdett and Mortensen (1998).

In the second column of Table 5 we note however that workers in larger firms are less likely to separate, *even conditional* on the worker's wage. Indeed even though the size coefficient falls in magnitude as expected, it continues to be economically large and statistically significant: indeed, the impact of size on the separation probability is as

²²Furthermore, even though the contact rates are the same for all firms regardless of m , firms with fewer tasks, in equilibrium convert a lower percentage of these contacts into hires, reinforcing Corollary 13 that follows.

| | Probability of Separation | Probability of Separation | Pr. of Separation & Occup Sw. | Pr. of Separation & No Occup Switch |
|-----------------|----------------------------------|----------------------------------|----------------------------------|-------------------------------------|
| ≥ 100 empl | -0.028 (0.002) ^{***} | -0.019 (0.002) ^{***} | -0.012 (0.002) ^{***} | -0.005 (0.001) ^{***} |
| ln(wage) | | -0.078 (0.002) ^{***} | -0.068 (0.002) ^{***} | -0.005 (0.001) ^{***} |

Table 5: Wage and Size Impact on Separation Probability (Probit). 1996 Panel of Survey of Income and Program Participation. 4-month probabilities. Controls include gender, race, education, medium size firm dummy, quadratic in age, 11 industry dummies, 13 occupation dummies. Coefficients represent marginal effects evaluated at the average value of the 4-month probability, which equal 0.115 (prob. of separating), 0.0838 (w/ occup. switch) and 0.0236 (no occup. switch). 197,819, 191,103 and 190,948 observations respectively.

large as that of a 24% wage increase. This is in contrast to models such as Burdett and Mortensen (1998), which predict that once we control for wages, firm size should no longer affect the separation probability.

This dependence occurs naturally in our framework: comparing two workers who are similarly matched (and therefore earn the same wage), the worker in the larger firm is less likely to separate; indeed if their match doesn't work out, the worker employed in the larger firm has more alternatives *within* his firm and therefore is less likely to leave his employer (Proposition 14). Indeed, as shown above, conditional on wages workers in larger firms are more likely to move within the firm, which is consistent with them having more options available. To explore this issue further, we examine separately the probability of separating and switching occupations and the probability of separating but not switching occupations.

In order to do that, we extend our setup to allow for employer switching without switching occupations. Assume employed workers meet other firms at rate $\lambda_1 > 0$. The worker's current occupation may or may not be available in the new firm. If the worker does meet another firm, he moves to the firm where his value, $V(\cdot)$, is the highest, after receiving the wage resulting from Nash bargaining, eq. (5). In Appendix D, we show that this is the result of an ascending auction in which the current and poaching firm place bids in order to attract the worker.

We first consider the case where the worker's current occupation is not available in the new firm. The worker moves to the new firm, if $E_{p_0} V(p_0, \tilde{m} - 1) > V(p, k)$, where \tilde{m} is the number of available occupations at the new firm. The probability the worker separates from his current employer, is declining in p and therefore in his current wage $w(p)$; it is also declining in k and therefore, following Lemma 7, in the size of his current firm m (see also Corollary 13).

We consider next the case of a worker whose current occupation is available in the new firm. In that case the worker switches firms, but remains in the same occupation if $V(p, \tilde{m} - 1) > V(p, k)$, or, put differently, if $\tilde{m} - 1 > k$. Intuitively, given that the quality of the worker's occupational match does not change if he remains in the same occupation, he only switches occupations if the new firm has more occupations available. Therefore, the probability of switching firms but staying in the same occupation depends negatively on the size of his current firm, not on his wage (see Lemma 7 and Corollary 13).²³

We now return to the data. The third column of Table 5 indicates that the probability of separating and switching occupations depends negatively on both the wage and the size of the current firm. In particular, a worker employed in a large firm is 14.3% less likely to separate and switch occupations, compared to one employed in a small firm. This is consistent with the implications of the model regarding both separations to unemployment, as well as separations to other firms where the worker switches occupations: in both cases the theory implies that wage and firm size negatively affect the probability of separating and switching occupations.

On the contrary, the fourth column of Table 5 shows that essentially only the size of a worker's current firm matters for the worker's decision to separate, as our model implies: a worker in a large firm is 21.2% less likely to separate compared to a worker in a small one. The coefficient on wages, although statistically significant is now an order of magnitude lower and close to zero. Other labor models (e.g. Burdett-Mortensen (1998)) do not distinguish between switching occupations or not when separating from a firm,

²³In fact, the wage may positively, rather than negatively, affect the rate of switching firms and not switching occupations. In particular, this is given by:

$$\lambda_1 \times \Pr(\text{occupation available}) \\ \times \Pr(\tilde{m} - 1 > k | \text{occ avail}) \times \Pr(p > \underline{p}(\tilde{m} - 1) | \text{occ avail} \ \& \ (\tilde{m} - 1 > k))$$

If one were to allow for firm-specific human capital accumulation (see for instance Idson (1993)), then we would expect workers with higher firm-specific human capital and therefore higher wages to have a lower probability of switching firms, but not occupations. It is possible that the two effects cancel out leading the coefficient on wages being essentially equal to zero in the data, as shown the fourth column of Table 5.

and uniformly predict a negative dependence of the separation probability on the wage. The above result indicates that the dependence on wages is different in the two cases, while firm size continues to be important, consistent with our proposed setup.²⁴

It is also worth noting, that the probability of separating and switching occupations is significantly larger than that of separating, but not switching occupations (8.38% versus 2.36%). This is consistent with occupational switching being an important determinant in the worker's decision to separate from his employer.

7 Conclusion

This paper investigates the interaction between firm size and occupational switching. We develop a theoretical model of internal labor markets and show that if mobility within firms is less costly than mobility across, then in equilibrium workers in larger firms are better matched, earn higher wages and separate less often.

Using data from the 1996 SIPP, we document a number of facts that support the implications of our setup: conditional on wages, internal mobility is higher in large firms, while external mobility is lower, consistent with larger firms having bigger internal labor markets. As worker reallocation has implications regarding the quality of worker matches, the different reallocation rates between large and small firms imply different match qualities.²⁵ Consistent with this result, we document that occupational mobility within a firm is associated with wage gains and that the wage premium is larger for workers with longer tenure: since a suitable match takes time, workers who have been with a large firm the longest have fully reaped the benefits.

Finally, the present paper, while consistent with previous models that emphasize the importance of occupational matching for workers, also introduces a role for firms in the labor market: namely firms act as institutions that facilitate occupational matching.

²⁴In the data, for workers switching employers, but not occupations, the probability of going to a small firm is significantly lower if they are currently in a large or medium sized firm, underlining the importance of firm size for these workers. Furthermore, workers who switch firms, but not occupations are less likely to go through unemployment during an employer switch compared to workers who also switch occupations, as implied by our theory.

²⁵One could argue that worker reallocation across firms is driven exclusively by non-pecuniary reasons and has no productivity implications. However such an explanation would have difficulty reconciling the wage patterns discussed next.

Appendix

A Proof of Lemma 1

The value to the firm of employing a worker with posterior p and k is given by:

$$rJ(p, k) = \bar{\alpha}(p) - w(p, k) + \frac{1}{2}\zeta^2 p^2 (1-p)^2 J_{pp}(p, k) - \delta J(p, k) \quad (12)$$

Multiplying eq. (2) by $1 - \beta$ and subtracting $(1 - \beta)rU$ leads to:

$$(1 - \beta)r(V(p, k) - U) = (1 - \beta)w(p, k) + \frac{1}{2}(1 - \beta)\zeta^2 p^2 (1-p)^2 V_{pp}(p, k) - \delta(1 - \beta)(V(p, k) - U) - (1 - \beta)rU \quad (13)$$

Multiplying eq. (12) by β , subtracting eq. (13) from it and using the surplus sharing condition, eq. (4), leads to:

$$w(p, k) = \beta\bar{\alpha}(p) + \frac{1}{2}\zeta^2 p^2 (1-p)^2 (\beta J_{pp}(p, k) - (1 - \beta)V_{pp}(p, k)) + (1 - \beta)rU \quad (14)$$

Finally, by taking the second derivative with respect to p in the surplus sharing condition, eq. (4), and using eq. (14), results in the wage equation, (5).

B Proof of Proposition 3

The surplus of the match between the firm and the worker, $S(p, k)$, is given by:

$$S(p, k) = V(p, k) + J(p, k) - U$$

Substituting in for $V(p, k)$ and $J(p, k)$ leads to:

$$(r + \delta)S(p, k) = \bar{\alpha}(p) - rU + \frac{1}{2}\zeta^2 p^2 (1-p)^2 S_{pp}(p, k)$$

The general solution to the above differential equation is given by:

$$S(p, k) = \frac{\bar{\alpha}(p) - rU}{r + \delta} + K_1^k p^{\frac{1}{2} - \frac{1}{2}\theta} (1-p)^{\frac{1}{2} + \frac{1}{2}\theta} + K_2^k p^{\frac{1}{2} + \frac{1}{2}\theta} (1-p)^{\frac{1}{2} - \frac{1}{2}\theta}$$

where $\theta = \sqrt{\frac{8(r+\delta)}{\zeta^2} + 1}$ and K_1^k and K_2^k are undetermined coefficients that depend on k . When $p \rightarrow 1$ however, $\lim_{p \rightarrow 1} K_2^k p^{\frac{1}{2} + \frac{1}{2}\theta} (1-p)^{\frac{1}{2} - \frac{1}{2}\theta} = K_2^k \cdot 1 \cdot \lim_{p \rightarrow 1} (1-p)^{\frac{1}{2} - \frac{1}{2}\theta} = +\infty$ which follows from $\theta > 1$. Since the present discount sum of the output produced by a good match is given by $\frac{\alpha^G}{r+\delta} < \infty$, it must be the case that $K_2^k = 0, \forall k$. Thus:

$$S(p, k) = \frac{\bar{\alpha}(p) - rU}{r + \delta} + K_1^k p^{\frac{1}{2} - \frac{1}{2}\theta} (1-p)^{\frac{1}{2} + \frac{1}{2}\theta}$$

where K_1^k is an undetermined coefficient. Moreover:

$$S_p(p, k) = \frac{\alpha^G - \alpha^B}{r + \delta} + K_1^k \left(\frac{1}{2} - \frac{1}{2}\theta - p \right) p^{-\frac{1}{2} - \frac{1}{2}\theta} (1-p)^{-\frac{1}{2} + \frac{1}{2}\theta}$$

Consider the case where $k = 0$. Using equation (4), one can rewrite the value matching and smooth pasting conditions (equations (8) and (9) respectively), in terms of $S(\cdot)$ and use them to pin down $\underline{p}(0)$ and K_1^0 .

From equation (9), one immediately obtains:

$$K_1^0 = -\frac{\alpha^G - \alpha^B}{r + \delta} \left(\frac{1}{2} - \frac{1}{2}\theta - \underline{p}(0) \right)^{-1} \underline{p}(0)^{\frac{1}{2} + \frac{1}{2}\theta} (1 - \underline{p}(0))^{\frac{1}{2} - \frac{1}{2}\theta}$$

Substituting for K_1^0 into equation (8) and solving leads to equation (10).

Similarly in the case where $k > 0$, equation (7) leads to:

$$K_1^k = -\frac{\alpha^G - \alpha^B}{r + \delta} \left(\frac{1}{2} - \frac{1}{2}\theta - \underline{p}(k) \right)^{-1} \underline{p}(k)^{\frac{1}{2} + \frac{1}{2}\theta} (1 - \underline{p}(k))^{\frac{1}{2} - \frac{1}{2}\theta}$$

Substituting for K_1^k into equation (6) leads to equation (11).

To show that the value function is increasing in p , note that straightforward calculations show that $S_{pp}(p, k) > 0$, and therefore $V_{pp}(p, k) > 0$, for all p and k . Given that $V_p(\underline{p}(k), k) = 0$, this implies that $V_p(p, k) > 0$ for all $p > \underline{p}(k)$.

C Proof of Lemma 5

We want to show that $\int_{\underline{p}(k)}^1 pf_k(p) dp$ is greater when k is higher.

Consider the set of all workers with k tasks remaining who draw the same prior, $p_0 \in (\underline{p}(k), 1)$. Let p_t be a diffusion process that starts at $p_0 \in (0, 1)$, evolves according to equation (1), while at a Poisson rate $\delta > 0$, it returns to p_0 . Finally, let $\underline{p} \in (0, p_0)$, be a reflective boundary, such that when the process hits it, it immediately returns to p_0 .

Moreover define:

$$I(p_0, \underline{p}) = \int_{\underline{p}}^1 p h_{\underline{p}}^{p_0}(p) dp$$

where $h_{\underline{p}}^{p_0}(p)$ is the steady state density of the above process.

Then:

$$\int_{\underline{p}(k)}^1 p f_k(p) dp = \int_{\underline{p}(k)}^1 I(p_0, \underline{p}(k)) \frac{g(p_0 | p_0 > \underline{p}(k))}{1 - G(\underline{p}(k))} dp_0$$

For a given $p_0 > \underline{p}(k)$, the weight $\frac{g(p_0 | p_0 > \underline{p}(k))}{1 - G(\underline{p}(k))}$ is higher when $\underline{p}(k)$, and therefore k (Corollary 4), is higher. Moreover, if $I(p_0, \underline{p}(k))$ is increasing in $\underline{p}(k)$ and p_0 and therefore k , then the left hand side is increasing in k .

We proceed to show that $I(p_0, \underline{p}(k))$ is increasing in $\underline{p}(k)$ and p_0 .

We first want to prove that if $\underline{p}^1 < \underline{p}^2$, then $I(p_0, \underline{p}^1) < I(p_0, \underline{p}^2)$.

Define:

$$T_p = \inf \{t > 0 : p_t = p\}$$

Call p_t^i the diffusion with boundary \underline{p}^i .

From Athreya and Lahiri (2006), Theorem 14.2.10, part (i),²⁶ we have that:

$$I(p_0, \underline{p}) = \frac{1}{E_{p_0} T_{\underline{p}}} E_{p_0} \left(\int_0^{T_{\underline{p}}} p_s ds \right) \quad (15)$$

Since:

$$E_{p_0} T_{\underline{p}^1} = E_{p_0} T_{\underline{p}^2} + E_{\underline{p}^2} T_{\underline{p}^1}$$

and:

$$E_{p_0} \left(\int_0^{T_{\underline{p}^1}} p_s^1 ds \right) = E_{p_0} \left(\int_0^{T_{\underline{p}^2}} p_s^2 ds \right) + E_{\underline{p}^2} \left(\int_0^{T_{\underline{p}^1}} p_s^1 ds \right)$$

straightforward algebra implies that it suffices to prove that:

$$\frac{1}{E_{\underline{p}^2} T_{\underline{p}^1}} \left(E_{\underline{p}^2} \left(\int_0^{T_{\underline{p}^1}} p_s^1 ds \right) \right) < I(p_0, \underline{p}^2) \quad (16)$$

Define a process Y_t which starts from \underline{p}^2 and evolves like p_t^1 with the only difference

²⁶In our case we set $f(X_j) = X_j$. The theorem states the proof in the case where X_j is a discrete time process. To prove it in the continuous time we follow the proof of the discrete case. The sums Y_i in that proof are replaced by integrals and the sequence of regeneration times in our case is the times where p_t^i hits \underline{p}_i ($i = 1, 2$). In each of these times the diffusion p_t^i restarts and the trajectories between these times are i.i.d.

that whenever it reaches \underline{p}^1 , it resets to \underline{p}^2 , rather than p_0 . As we know, p_t^2 starts from p_0 . Run Y_t and p_t^2 with the same Brownian motion, W_t , and the same Poisson times of resetting. Now, when Y_t and p_t^2 meet, they continue together until \underline{p}^2 is hit. Then p_t^2 jumps to p_0 , while Y_t continues. Moreover, since p_t^2 starts above Y_t , it is always true that $Y_t \leq p_t^2$ (if Y_t jumps to p_0 , then p_t^2 jumps as well). However for a positive proportion of time the inequality is strict. Thus:

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t Y_s ds < \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t p_s^2 ds$$

which using eq. (15) above, leads to eq. (16).

We also need to prove that if $p_0^3 < p_0^4$, then $I(p_0^3, \underline{p}) < I(p_0^4, \underline{p})$.

Now call p_t^i the diffusion starting off from p_0^i .

As before, since:

$$E_{p_0^4} T_{\underline{p}} = E_{p_0^4} T_{p_0^3} + E_{p_0^3} T_{\underline{p}}$$

and:

$$E_{p_0^4} \left(\int_0^{T_{\underline{p}}} p_s^4 ds \right) = E_{p_0^4} \left(\int_0^{T_{p_0^3}} p_s^4 ds \right) + E_{p_0^3} \left(\int_0^{T_{\underline{p}}} p_s^3 ds \right)$$

straightforward algebra implies that it suffices to prove that:

$$I(p_0^3, \underline{p}) < \frac{E_{p_0^4} \left(\int_0^{T_{p_0^3}} p_s^4 ds \right)}{E_{p_0^4} T_{p_0^3}} \quad (17)$$

Using a similar argument as above, define a process Z_t which starts from p_0^4 and evolves like p_t^4 with the only difference that whenever it resets to p_0^4 whenever it reaches p_0^3 rather than \underline{p} . As we know, p_t^3 starts from p_0^3 and resets at \underline{p} . Running Z_t and p_t^3 with the same Brownian motion, W_t , and the same Poisson times of resetting implies as above that:²⁷

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t p_s^3 ds < \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t Z_s ds$$

which using eq. (15) above, leads to eq. (17).

²⁷Essentially now we have p_0^4 in the place of p_0 , p_0^3 in the place of \underline{p}_2 , and \underline{p} in the place of \underline{p}_1 . In the place of Y_t we have p_t^3 and in the place of p_t^2 we have Z_t .

D On-The-Job Search

In this section, we analyze the auction that takes place if an employed worker meets another firm. In particular, the current firm and poaching firm engage in an English first-price auction to attract the worker, as in Moscarini (2005): they make bids consisting of lump-sum transfers to the worker plus the promise to split the resulting match’s surplus afterwards via Nash bargaining.

Let $S(p, k)$ denote the surplus of the worker’s match with a firm where he has k tasks remaining and whose current posterior equals p . Similarly let $E_{p_0}S(p_0, \tilde{m} - 1)$ denote expected surplus of the worker with the poaching firm. If $E_{p_0}S(p_0, \tilde{m} - 1) > S(p, k)$, then the poaching firm can always outbid the worker’s current firm in a bidding war: the highest possible bid the current firm is willing to make is worth $S(p, k)$ which is strictly less than the maximum bid of the poaching firm, $E_{p_0}S(p_0, \tilde{m} - 1)$.²⁸ The current firm recognizes this and places no bid, while the worker switches to the poaching firm. The lump-sum transfer is zero and from then on the worker bargains bilaterally with the firm over the match’s surplus, receiving $V(p, k)$.²⁹ Similarly if $S(p, k) > E_{p_0}S(p_0, \tilde{m} - 1)$, the poaching firm knows it can’t outbid the current firm and places no bid. Finally, if $S(p, k) = E_{p_0}S(p_0, \tilde{m} - 1)$, we assume the worker stays with the incumbent. The above firm strategies constitute a subgame perfect Nash equilibrium.

Summarizing, since Nash bargaining implies that a worker’s value equals a constant share of the surplus, a worker ends up employed in the firm where his value, $V(\cdot)$, is the highest and receives the wage given by eq. (5).

References

- [1] Alvarez F. and R. Shimer (2009): “Unemployment and Human Capital,” Mimeo, University of Chicago, Department of Economics.
- [2] Alvarez F. and R. Shimer (2011): “Search and Rest Unemployment,” *Econometrica*, 79(1), 75-122.
- [3] Antonovics K. and L. Golan (2010): “Experimentation and Job Choice,” Mimeo, Carnegie Mellon University, Tepper School of Business.

²⁸In that case the worker receives a wage that guarantees him an expected present discounted value of $V(p, k)$, as prescribed by the Nash bargaining solution and $S(p, k) - V(p, k)$ as a one-time lump sum transfer, effectively capturing the entire surplus of the match.

²⁹Both firms bidding up to $S(p, k)$ and $E_{p_0}S(p_0, \tilde{m})$ respectively also forms a subgame Nash equilibrium of this game. However, such an equilibrium is not robust to adding a small cost to bidding.

- [4] Athreya, K. B. and S. N. Lahiri (2006): *Measure Theory and Probability Theory*, New York: Springer.
- [5] Barron, J. M., Black D. A. and M. A. Loewenstein (1987): “Employer Size: The Implications for Search, Training, Capital Investment, Starting Wages and Wage Growth,” *Journal of Labor Economics*, 5(1), 76-89.
- [6] Brown C. and J. Medoff (1989): “The Employer Size-Wage Effect,” *Journal of Political Economy*, 97(5), 1027-1059.
- [7] Burdett K. and M. Coles (2003): “Equilibrium Wage-Tenure Contracts,” *Econometrica*, 71(5), 1377-1404.
- [8] Burdett K. and D. T. Mortensen (1998): “Wage Differentials, Employer Size and Unemployment,” *International Economic Review*, 39(2), 257-273.
- [9] Eeckhout, J. and X. Weng (2010): “Assortative Learning,” Mimeo, University of Pennsylvania, Department of Economics.
- [10] Gibbons, R. and M. Waldman (1999): “Careers in Organizations: Theory and Evidence,” *Handbook of Labor Economics*, Volume 3, ed. by O. Ashenfelter and D. Card, Elsevier Science B. V., 2373-2437.
- [11] Gittleman, M., Horrigan M. and M. Joyce (1998): “Flexible” Workplace Practices: Evidence from a Nationally Representative Survey,” *Industrial and Labor Relations Review*, 52(1), 99-115.
- [12] Groes, F., Kircher P. and I. Manovskii (2010): “The U-Shapes of Occupational Mobility,” Mimeo, University of Pennsylvania, Department of Economics.
- [13] Groes, F. (2010): “Occupational Experience, Wages and Mobility: Patterns in the Danish data,” Mimeo, Copenhagen Business School.
- [14] Hill, D. H. (1994): “The Relative Empirical Validity of Dependent and Independent Data Collection in a Panel Survey,” *Journal of Official Statistics*, 10(4), 359-380.
- [15] Idson, T. L. (1993): “Employer Size and Labor Turnover,” Discussion Paper No. 673, Columbia University, Department of Economics.
- [16] Idson, T. L. and D. J. Feaster (1990): “A Selectivity Model of Employer-Size Wage Differentials,” *Journal of Labor Economics*, 8(1), Part 1, 99-122.

- [17] Idson, T. L. and W. Y. Oi (1999): “Workers Are More Productive in Large Firms,” *American Economic Review Papers and Proceedings*, 89(2), 104-108.
- [18] Jovanovic, B. (1979): “Job Matching and the Theory of Turnover,” *Journal of Political Economy*, 87(5) Part 1, 972-990.
- [19] Kambourov, G. and I. Manovskii (2009a): “Occupational Specificity of Human Capital,” *International Economic Review*, 50(1), 63-115.
- [20] Kambourov, G. and I. Manovskii (2009b): “Occupational Mobility and Wage Inequality,” *Review of Economic Studies*, 76(2), 731-759.
- [21] Karlin, S. and H. M. Taylor (1981): *A Second Course in Stochastic Processes*, New York: Academic Press.
- [22] Liptser, R. and A. Shyryaev (1977): *Statistics of Random Processes*, Vol. 2 Berlin: Springer-Verlag.
- [23] Moen, E. R. (1997): “Competitive Search Equilibrium,” *Journal of Political Economy*, 105(2), 385-411.
- [24] Moore, H. L. (1911): *Laws of Wages*, New York: MacMillan.
- [25] Mortensen, D. T. and C. A. Pissarides (1994): “Job Creation and Job Destruction in the Theory of Unemployment,” *The Review of Economic Studies*, 61(3), 397-415.
- [26] Moscarini, G. (2005): “Job Matching and the Wage Distribution,” *Econometrica*, 73(2), 481-516.
- [27] Moscarini G. and K. Thomsson (2006): “Occupational and Job Mobility in the US,” Yale Working Papers on Economic Applications and Policy, Discussion Paper No 19.
- [28] Nagypal, E. (2008): “Worker Reallocation over the Business Cycle: The Importance of Employer-to-Employer Transitions,” Mimeo, Northwestern University, Department of Economics.
- [29] Neal, D. (1999): “The Complexity of Job Mobility among Young Men,” *Journal of Labor Economics*, 17(2), 237-261.
- [30] Novos, I. E. (1992): “Learning by doing, Adverse Selection and Firm Structure,” *Journal of Economic Behavior and Organization*, 19(1), 17-39.

- [31] Ortega, J. (2001): “Job Rotation as a Learning Mechanism,” *Management Science*, 47(10), 1361-1370.
- [32] Papageorgiou, T. (2010): “Learning Your Comparative Advantages,” Mimeo, Penn State University, Department of Economics.
- [33] Pastorino, E. (2009): “Learning in Labor Markets and Job Mobility,” Mimeo, University of Iowa, Department of Economics.
- [34] Pissarides, C. A. (2000): *Equilibrium Unemployment Theory*, 2nd edition, Cambridge MA: MIT Press.
- [35] Postel-Vinay, F. and J.-M. Robin (2002): “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity,” *Econometrica*, 70(6), 2295-2350.
- [36] Schmidt, C. M. and K. F. Zimmermann (1991): “Work Characteristics, Firm Size and Wages,” *The Review of Economics and Statistics*, 73(4), 705-710.
- [37] Sullivan, P. (2010): “Empirical Evidence on Occupation and Industry Specific Human Capital,” *Labour Economics*, 17(3), 567-580.