Large Firms and Internal Labor Markets

Theodore Papageorgiou*

Penn State University, Department of Economics†

August 2010

Abstract

This paper introduces a model of internal labor markets that is consistent with the observed differences between workers in large and small firms with respect to wages and separation rates. In particular, firms constitute labor markets with no search frictions. Workers are free to move within a firm at no cost, whereas switching across firms is costly. If the quality of a match between a worker and a occupation/department/team within a firm is uncertain, then larger firms offer more opportunities for workers to find the right match. As a result, workers abandon unpromising matches more easily and are more likely to be employed in better matches. In equilibrium, workers in larger firms are more productive, earn higher wages and are less likely to quit, even conditional on their wage. Using data from the 1996 SIPP we find support for the predictions of our framework: internal mobility is higher in larger firms and depends negatively on wages and tenure; workers in larger firms switch occupations at higher wage levels and receive higher wages in their new occupation; the size-wage premium is higher for workers with longer tenure, while workers who leave large firms continue to enjoy high wages, but only if they remain in the same occupation; and finally the wage and size effects on the separation probability are significantly larger for workers who switch occupations.

Keywords: Internal Labor Markets, Occupations, Size-Wage Differentials

JEL Classification: J24, J31

*I am grateful to Joe Altonji, Dimitris Cheliotis, Manolis Galenianos, Ed Green, Alex Monge, Giuseppe Moscarini, Andrés Rodríguez-Clare, Richard Rogerson, Russell Toth, Neil Wallace and Ruilin Zhou, as well as seminar participants at the Spring 2010 Cornell-Penn State Macro Workshop, the Search and Matching Workshop at the University of Pennsylvania and 2010 SED 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

Every year 8% of all workers who do not change employers, switch occupations. Furthermore, as shown in Table 1, workers in larger firms are much more likely to switch compared to workers in small ones. The goal of this paper is to study these internal labor markets. We argue that the size of the labor market matters: workers in larger firms form better matches, which in turn is consistent with the observed difference between large and small firms with respect to wages and separation rates.

In particular, a firm constitutes a labor market without search frictions. Workers can move *within* a firm at no cost. Switches *across* firms, however, are costly. The quality of a match between a worker and either a task, a department or a team is uncertain and is revealed only gradually.[1] Workers, therefore, need to experiment in order to find a suitable match. If a task appears to be sufficiently unpromising, a worker moves on to the next one. Larger firms offer more opportunities for workers to experiment. As a result, workers there become selective and are 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 which leads to higher output, higher wages and lower separation rates. Furthermore, in equilibrium, firms with more tasks have more workers.

For example, Google in 1999 was a small company with only 40 full-time employees. Computer engineers hired by the company at the time could only work on search engines. On the contrary, a computer engineer hired by Google today can join the division developing Google Translate-a translation service Google has introduced-or join a team working on Chrome OS, a computer operating system, or any other of the numerous products and teams there. If the engineer begins working at Google Search, but starts realizing that he may not be as good at developing search engines as he originally thought, he can switch to another product. Ten years ago, he would have thought twice before quitting, because that would have meant quitting his job and looking for another company to work in.

The proposed theory is consistent with the observed empirical regularities: workers in larger firms are more productive (Idson and Oi (1999)), they are paid higher wages (Moore (1911), Brown and Medoff (1989)) and their initial wages are also higher (Barron et al. (1987)).[2] In addition, the theory predicts that workers in larger firms have better

[1] The 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 matches for the worker.

[2] The positive relationship between size and wage holds even conditional on observed and unobserved labor quality (Idson and Feaster (1990)), hazardous working conditions and occupation. Brown and Medoff (1989) find that the size premium remains even when considering firms that offer a piece-rate

Firm Employees:	Annual Switching Probability (Unconditional)
<25	5.48%
25-99	7.29%
>100	8.88%

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

outside 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)).

This last implication distinguishes our setup from other existing models. For instance, the Burdett and Mortensen (1998) model similarly predicts that workers in larger firms are paid higher wages and are therefore less likely to quit; but once one controls for the wage, firm size no longer affects the separation probability. In contrast, our setup can naturally replicate this feature of the data: even conditioning on the perceived quality of the worker's current match, a worker in a larger firm is less likely to leave if it proves unpromising, because it has other outside options within the firm.

To investigate this implication further, we look at a particular type of match for which we have data: occupations.[3] Using the 1996 panel of the Survey of Income and Program Participation, we divide workers among those who switch occupations in the next period and those who do not. The effect of wages and firm size on the separation probability is much larger for workers who switch compared to those who remain in the same occupation, consistent with the predictions of our setup.

Our empirical investigation also support for other predictions of the model: there's more internal mobility in larger firms, while long tenure and high wages predict a lower probability of switching occupations. Furthermore, occupation switching is associated with wage gains, while workers in larger are found to switch occupations at higher wage levels and also receive higher wages in their new occupations. Finally, the size-wage premium is higher for workers with longer tenure, while workers who switch occupations continue to enjoy high wages if they remain in the same occupation. These last results are also consistent with our proposed mechanism, that large firms allow workers to find better

system.

[3] Our empirical exercise explores model implications regarding occupations. However the 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 coworkers there. One may also think of it as a location match as well, with larger firms having more locations.

occupational matches: finding a good match takes time, so the workers who have been with a large firm the longest will have fully reaped the benefits. Moreover, if a worker who has found a good match in a large firm, moves on, then he will continue to enjoy the benefit of a good match in his new firm if he doesn't switch occupations.

The paper closest to ours is Idson (1993). He argues that because large firms have lower failure rates and larger internal labor markets, workers are less likely to separate. This in turn, increases the incentives of firms to provide firm-specific training to the workers. This latter result would also hold in our setup: higher expected tenure increases the incentives for workers to acquire firm-specific human capital, further amplifying differences with respect to wages and separation rates. Idson does not model internal labor markets, but does provide evidence that internal mobility is higher in larger firms, as is on-the-job training. For a survey of the literature on internal labor markets see Gibbons and Waldman (1999).[4]

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."[5]

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

---

[4] See also Pastorino (2009) for a more recent contribution.

[5] Ortega (2001) and Antonovics and Golan (2008) also argue that these rotation programs are designed to provide information regarding the worker's fit within the firm.

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, $m \in \{1, 2..., M\}$.[6] The distribution of tasks is exogenous and let q_m denote the fraction of firms with m tasks.

There is also a fixed mass of workers. Workers are risk neutral and have discount rate r. They can be either employed or unemployed. All unemployed workers are identical.

Workers and firms need to search for each other. Search is undirected and an unemployed worker meets a firm according to a Poisson process with parameter λ. Moreover, 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.

When employed in a firm, a worker can work in only one task any time. We assume there are no congestion externalities inside the firm, i.e. there are no restrictions in the number of workers employed in a task. Once a worker leaves a task, he cannot return to it. 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 and $\alpha^{ijl} \in \left\{\alpha^G, \alpha^B\right\}$ is mean output per unit of time. 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, and let $p_0^{ijl} \in (0, 1)$ be the worker's prior belief that $\alpha^{ijl} = \alpha^G$. A worker draws his prior p_0^{ijl} for each task, before starting work there from a known distribution $G(\cdot)$ with support $[0, 1]$.

Following the literature, we assume that the wage is determined by generalized Nash bargaining, with $\beta \in (0, 1)$ denoting the worker's bargaining power. Finally, an unemployed worker earns $b \geq \alpha^B$.[7]

The sequence of actions is the following: unemployed workers meet firms at rate λ and observe the number of tasks of the firm, m. They next draw their prior for their first task from $G(\cdot)$ and either begin working in that task or choose to move on to the next task and draw a new prior there.

[6] We use the terminology "tasks", but the reader may choose to think of these as departments, teams, locations etc.

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

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 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)):

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

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]

As shown in the appendix, generalized Nash bargaining implies that the worker's wage is given by:

$$w(p) = \beta\overline{\alpha}(p) + (1 - \beta) rU \tag{2}$$

where $\overline{\alpha}(p) = p\alpha^G + (1 - p)\alpha^B$ is the worker's expected output and U is the value of an unemployed worker. Note that the wage is a affine transformation of the posterior, p and does not depend on k: as in the standard Mortensen and Pissarides (1994) model, wages only depend on the current level of output and not on its future path (except through U).[9]

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:

$$rV(p,k) = w(p) + \frac{1}{2}\zeta^2 p^2 (1-p)^2 V_{pp}(p,k) - \delta\left(V(p,k) - U\right) \tag{3}$$

[8] Note that the total number of tasks in the worker's current firm is not relevant, as employment history in previous tasks does not affect the worker's current or future payoffs.

[9] The wage equation (2) is almost identical to the wage equation in the Mortensen-Pissarides model with idiosyncratic shocks (see for instance equations (2.9) and (2.10) on page 42 of Pissarides (2000)).

where $V(\cdot)$ is his value and $V_{pp}(\cdot)$ is the second derivative of $V(\cdot)$ with respect to 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:

$$rU = b + \lambda \left(\sum_{m=1}^{M} q_m EV(p_0, m-1) \right) - \lambda U \tag{4}$$

where $EV(p_0, k)$ is the expected value of a worker, with k other tasks remaining, about to draw his prior for his current task.[10]

If we increase the number of tasks available to the worker, then the worker can not be worse off, as he has the option of not trying them out, so the $V(\cdot)$ 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.[11]

Lemma 1 *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 quit 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:

$$V\left(\underline{p}(k), k\right) = EV(p_0, k-1) \tag{5}$$

$$V_p\left(\underline{p}(k), k\right) = 0 \tag{6}$$

[10] One can write:

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

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

[11] Consider a worker with current posterior $p < 1$ and k tasks remaining. Assume $V(\cdot)$ is increasing in p (we show this to be true later) 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 the next task, draw 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 higher. Since the worker separates from his current task with positive probability, then the option value of trying out one more task is positive.

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:

$$V\left(\underline{p}(0), 0\right) = U \tag{7}$$

and:

$$V_p\left(\underline{p}(0), 0\right) = 0 \tag{8}$$

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 next task or unemployment.

In the appendix we use the boundary conditions (5) through (8) to solve for the value of a worker. The following proposition states the result:

Proposition 2 *The employed worker's value function, $V(\cdot)$, is increasing in p. Moreover a worker optimally quits his current task when his posterior hits $\underline{p}(k)$, defined as :*

$$\underline{p}(0) = \frac{(\theta - 1)\left(rU - \alpha^B\right)}{(\theta - 1)(rU - \alpha^B) + (\theta + 1)(\alpha^G - rU)} \tag{9}$$

for $k = 0$ and:

$$\underline{p}(k) = \frac{(\theta - 1)\left((r + \delta) EV(p_0, k - 1) - (r + \delta - \beta r) U - \beta a^B\right)}{\beta\left(\alpha^G + a^B + (\alpha^G - a^B)\theta\right) - 2\left((r + \delta) EV(p_0, k - 1) - (r + \delta - \beta r) U\right)} \tag{10}$$

for $k > 0$, where $\theta = \sqrt{\frac{8(r+\delta)}{\zeta^2} + 1}$ and the value of unemployment is defined in equation (4). 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 1, immediately results in the following corollary:

Corollary 3 *The endogenous separations triggers $\underline{p}(k)$, are increasing in k.*

[12] A worker quits to unemployment only after exhausting all tasks available to him, i.e. $EV(p_0, 0) > U$. 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 will still prefer to be unemployed). However, from eq. (9) 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 $m = 1$.

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

4 Model Predictions

We next use the results above to investigate the model's implications. Before investigating model's productivity implications, one result is immediate:

Proposition 4 *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 3, 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 low prior do so because they expect to draw a higher one in their next task: since $V(\cdot)$ 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.

We next 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. The next step is to establish that in firms with fewer tasks, (low m), there is a higher percentage of workers that have few tasks remaining, (low k).

The equilibrium evolution of beliefs, employment realizations 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 long as $\delta > 0$.[13] Therefore there exists a stationary distribution of beliefs and remaining matches.

We next group tasks according to k, the number of other 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 $\frac{g\left(p_0 | p_0 > \underline{p}(k)\right)}{1-G\left(\underline{p}(k)\right)}$.

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:

[13] An unemployed worker also returns to that state infinitely many times.

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 the appendix 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 a higher $\underline{p}$ (Corollary 3), as well as a higher p_0 on average ($G(\cdot)$ is truncated at a higher $\underline{p}$), then, as shown in the appendix, the following lemma is true:

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

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

when $k' > k$.

Moreover, let:

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

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.:*

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

whenever $k' > k$.

We next show that firms with many tasks (high m), have a lower percentage of workers employed at low k, and therefore unproductive, tasks.

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 is employed with $k = 3$ tasks remaining, since this implies that they have to first exhaust 46 other tasks. In the second firm however, one anticipates a higher percentage of its workers workers to have $k = 3$ tasks remaining, since workers hired by that firm start off with $k = 4$.

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

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

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:

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

Since $\Pr(k|k+1) \in (0,1)$ for all k, 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

[14] Ignoring 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 $\frac{1-p}{1-\underline{p}(m-1)}$ (Karlin and Taylor (1981)). Given that, we have:

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

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,* $\sum_{k=0}^{m-1} s_k^m \overline{\alpha}_k$*, is increasing in the number of 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*.*

We next examine how the separation rate varies with the number of tasks, m. Given the above discussion, the probability that a worker, who just found employment in a firm with m tasks, separates endogenously, is given by (ignoring δ shocks):

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

where $\Pr(Un|1)$ is the probability a worker starting off in the last task quits to unemployment. Since $\Pr(k|k+1) \in (0,1)\ \forall k$, 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 10 *Firms with more tasks,* m*, have lower separation rates.*

Note that Proposition 10 predicts that average tenure in firms with more tasks firms is higher. Thus the following corollary immediately follows:[15]

Corollary 11 *Firms with more tasks,* m*, have more employees.*

The separation probability of two workers with the same 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 12 *Conditional on wage, the worker separation probability declines in the total number of firm tasks* m*.*

[15] Furthermore, 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 the following result.

5 Empirical Evidence

In this section we further explore the predictions of our setup. We assume that a worker matches with an occupation and we investigate the model's implications regarding internal mobility, the wage premium and the probability a worker separates from his firm.[16] 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 30,000 households. It includes information on the worker's wage, occupation, current employer and employer size (three size categories).[17]

We first look at the predictions of the model regarding internal mobility and the impact of an occupational switching on wages. We next examine the wage premium, how it evolves with firm tenure and whether the effect of working in a large firm persists even after a worker has moved on. We conclude by looking into firm separation probabilities and how they are affected by firm size and wage.

5.1 Internal Mobility

Table 2 shows how wages, tenure and firm size affect the probability of switching occupations for workers who remain with the same employer. Consistent with our setup, high wages and longer tenure are both associated with fewer occupational switches, while workers in larger firms are more likely to switch: workers with higher wages are those who have managed to find a good match within the firm and are therefore less likely to switch. Moreover, even conditioning on wages (and therefore match quality), workers in larger firms should be more likely to switch (Corollary 3 and Lemma 7). Finally, even conditioning on match quality and firm size, a worker who's been in the firm longer (high tenure) is less likely to switch occupations, since he should have fewer tasks remaining and therefore a lower switch trigger.

[16] One may imagine alternatively that workers are learning about their type and they are looking for occupations in which they fit best as in Papageorgiou (2009). In that model, workers may find themselves preferring to work in another occupation, but if it is not available, they may remain in their current occupation, being mismatched (workers with posterior, $p \in (\underline{p}, \widehat{p})$ in the case of occupation W and those with posterior, $p \in (\widehat{p}, \overline{p})$ in the case of occupation B). If they were employed in a firm that had multiple occupations, then they would have immediately transferred to their preferred occupation, increasing their expected output.

[17] In our investigation, we exclude workers in the armed forces. Wages are 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)).

	Occ. Switching
	Prob. (Probit)
	Same Employer
ln(wage)	-0.0037
	$(0.0011)^{***}$
tenure	-0.0003
	$(0.00007)^{***}$
≥100 employees	0.0128
	$(0.0012)^{***}$
25-99 employees	0.0074
	$(0.0016)^{***}$

Table 2: Wage, Tenure and Size Impact on Internal Mobility. 1996 Panel of Survey of Income and Program Participation. 4-month probabilities. Controls include gender, race, education, quadratic in age, 11 industry dummies. Coefficients represent marginal effects evaluated at the average value of the 4-month probability, which equals 0.0283.

	ln(wage)
	Same Employer
Occ Switch Dummy	0.026
	$(0.002)^{***}$
$\ln(\text{wage})_{t-1}$	0.91
	$(0.001)^{***}$

Table 3: Impact of Occupation Switch on Wage. 1996 Panel of Survey of Income and Program Participation. 4-month switching probabilities. Controls include gender, race, dummy for high school graduate and dummy for college graduate, quadratic in age, 11 industry dummies.

5.2 Occupational Switching and Wages

We next explore the impact of occupational switching on wages. Our model predicts 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. Our prediction is consistent with the data: Table 3 reveals that workers switching occupations within a firm experience significant wage gains.

We investigate this issue further, by looking at the impact of occupational switching on wages of workers across firms of different sizes. Our setup predicts that workers in larger firms are both more likely to leave an unpromising much earlier (Corollary 3 and Lemma 7) and more likely to obtain higher wages in their new occupation (again from Corollary 3 and Lemma 7, they are more likely to reject a low initial draw, p_0). The

	$\ln(\text{wage})_{t-1}$
	Same Employer
	w/ Occup Switch
≥100 empl	0.143
	$(0.015)^{***}$
25-99 empl	0.052
	$(0.02)^{***}$

Table 4: Size Effect on Pre-Switch Wages. 1996 Panel of Survey of Income and Program Participation. 4-month switching probabilities. Controls include gender, race, dummy for high school graduate and dummy for college graduate, quadratic in age, 11 industry dummies.

	ln(wage)
	Same Employer
	w/ Occup Switch
≥100 empl	0.03
	$(0.008)^{***}$
25-99 empl	0.022
	$(0.01)^{**}$
$\ln(\text{wage})_{t-1}$	0.858
	$(0.007)^{***}$

Table 5: Size Effect on Post-Switch Wages. 1996 Panel of Survey of Income and Program Participation. 4-month switching probabilities. Controls include gender, race, dummy for high school graduate and dummy for college graduate, quadratic in age, 11 industry dummies.

results in Table 4 are consistent with the first of these predictions: the pre-switch wages of workers in larger firms are higher on average. To consider the second prediction, we look at the wages of workers who switch, *conditional* on their pre-switch wage: indeed, as seen in Table 5, workers in larger firms earn higher wages in their new occupation.

5.3 Wage Premium

Table 6 shows the well-documented wage premium: workers in larger firms earn significantly higher wages, even conditional on observables. Figure 1, shows the evolution of the wage premium with tenure. We note that initial wages at larger firms are higher than those of smaller firms, but that the difference equals to a fraction of the overall wage premium. The longer the workers stays with the firm, the greater the difference between

	ln(wage)
≥100 empl	0.144
	(0.002)***
25-99 empl	0.033
	(0.003)***

Table 6: Wage Premium. 1996 Panel of Survey of Income and Program Participation. 4-month switching probabilities. Controls include gender, race, dummy for high school graduate and dummy for college graduate, quadratic in age, 11 industry dummies.

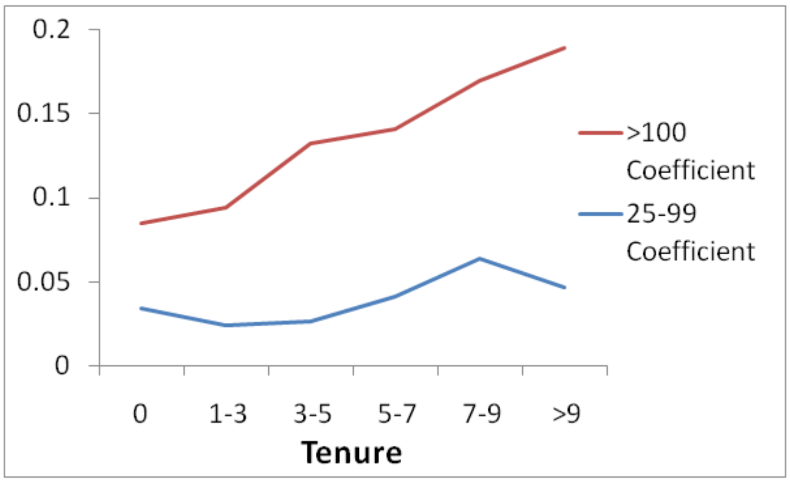

Figure 1: Wage Premium by Tenure Level. 1996 Panel of Survey of Income and Program Participation. 4-month switching probabilities. Controls include gender, race, dummy for high school graduate and dummy for college graduate, quadratic in age, 11 industry dummies. All coefficients significant at one percent.

his wage and the wage of workers in smaller firms with similar tenure levels.[18]

These observations are consistent with our framework's predictions: 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.

We next investigate whether the effect of working in a large firm persists after the worker has left. In the first column of Table 7 we observe that workers who have separated from a large firm continue to earn more in their new employer. In other words, the effect

[18] This result casts doubt on the selection hypothesis: if more able individuals are more likely to seek employment in larger firms, one expects the wage premium to be the high for incoming workers and not change over time. Brown and Medoff (1989) argue that a significant wage premium remains, even after controlling for omitted ability bias.

of large firms remains even after separation. This is inconsistent with explanations that emphasize mechanisms that allow workers to be more productive while in larger firms, such as more available capital or increased specialization, or explanations that focus on monitoring constraints. Note that this result also implies that fixed effects regressions are potentially mispecified: if workers become *permanently* more productive when employed in a large firm, then a fixed effect regressor will attribute this some of the increase in productivity to the worker fixed effect, thus underestimating the true extent of the wage premium.

The results in first column are however consistent with our story if workers are able to find similar occupations in other firms as well. In that case, if large firms allow workers to sample many occupations until they have found a good occupational match, then workers who leave the firm, but remain in that occupation, will continue to benefit having found a good match. We examine this prediction, by dividing workers, into those who switch occupations in the next period and those who do not. As seen in the second and third column of Table 7, the effect of working in a large firms persists afterwards mainly for workers who stay in the same occupation: the size effect for those that remain in the same occupation is three to four times higher than that of those who switch occupations, as our model would predict.[19] Table 8 presents the same regressions for layoffs only. The coefficients are roughly similar to the unconditional case, indicating that they are not the result of workers upgrading to better matches.

5.4 Separation Probabilities

We next turn to the probability that a worker separates from his firm. Table 9 shows that a worker is much more likely to separate from his current firm, if he switches occupations in the next period. In fact, 76.74% of all employer switches involve occupational switches as well. The first column of Table 10 looks at the dependence of the separation probability on wages and firm size. As expected, workers with higher wages are less likely to separate from their employer, consistent with our setup. This negative relationship between wages and the separation probability is well-known and consistent with other models in the literature, such as Burdett and Mortensen (1998). However those models have difficulty explaining

[19]It is interesting to note that the size effect is still significant, even though much smaller, for workers who switch occupations. This is also consistent with our setup: we have shown that larger internal labor markets imply a lower separation rate for larger firms. As Idson (1993) points out, if the separation rate is lower, then larger firms have increased incentives to provide training to incoming workers. If some of this training is general enough to be transferable across occupations and firms, then even workers who switch occupations when leaving large firms, will be more productive in their new tasks.

	ln(wage)	ln(wage)	ln(wage)
	New Employer	New Employer	New Employer
		w/ Occup Switch	No Occup Switch
$\geq$100 (Old Empl)	0.052	0.034	0.128
	(0.006)***	(0.006)***	(0.014)***
25-99 (Old Empl)	0.031	0.024	0.073
	(0.008)***	(0.009)***	(0.019)***

Table 7: Wage Premium when Switching Employers. 1996 Panel of Survey of Income and Program Participation. 4-month switching probabilities. Controls include gender, race, dummy for high school graduate and dummy for college graduate, quadratic in age, 11 industry dummies.

	ln(wage)	ln(wage)	ln(wage)
	New Employer	New Employer	New Employer
		w/ Occup Switch	No Occup Switch
$\geq$100 (Old Empl)	0.064	0.048	0.112
	(0.017)***	(0.02)**	(0.04)***
25-99 (Old Empl)	0.034	0.026	0.037
	(0.024)	(0.028)	(0.055)

Table 8: Wage Premium when Switching Employers - Layoffs only. 1996 Panel of Survey of Income and Program Participation. 4-month switching probabilities. Controls include gender, race, dummy for high school graduate and dummy for college graduate, quadratic in age, 11 industry dummies. Layoffs includes codes 1. on layoff, 8. discharged/fired, 9. employer bankrupt, 10. employer sold business and 13. slack work or business conditions.

why the separation probability is lower in larger firms, *conditional* on the worker's wage. 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 better alternatives *within* his firm and therefore is less likely to leave his employer (Proposition 12).

One may alternatively argue, that the dependence of the separation probability on firm size, conditional on wages, may be related to fringe benefits or other unobserved advantages of large firms that make them preferable to smaller ones. To explore this issue, as before, we divide workers into two groups: those who switch occupations in the next period and those who do not.

For the first group of workers the theory predicts a negative relationship between firm size, wages and the separation probability: if workers separate from their occupation at

	Separation
	Prob. (Probit)
Switch Occupation dummy	0.503
	$(0.002)^{***}$

Table 9: Separation Probability and Occupational Switching. 1996 Panel of Survey of Income and Program Participation. 4-month probabilities. Controls include gender, race, education, quadratic in age, log wage, medium and large firm size dummies, 11 industry dummies. Coefficient represents marginal effect evaluated at the average value of the 4-month probability, which equals 0.1117.

	Separation Pr.	Separation Pr.	Separation Pr.
		w/ Occup Sw.	No Occup Switch
ln(wage)	-0.079	-0.15	-0.006
	$(0.002)^{***}$	$(0.008)^{***}$	$(0.001)^{***}$
$\geq$100 empl	-0.019	-0.108	-0.006
	$(0.002)^{***}$	$(0.008)^{***}$	$(0.001)^{***}$
25-99 empl	-0.008	-0.06	-0.0006
	$(0.003)^{***}$	$(0.011)^{***}$	(0.0013)

Table 10: Wage and Size Impact on Separation Probability (Probit). 1996 Panel of Survey of Income and Program Participation. 4-month switching probabilities. Controls include gender, race, education, quadratic in age, 11 industry dummies. Coefficients represent marginal effects evaluated at the average value of the 4-month probability, which equal 0.1201 (unconditional), 0.7772 (w/ occup. switch) and 0.0274 (no occup. switch).

a high wage level, this implies that they have many occupations remaining within the firm (Corollary 3). Similarly, a larger firm offers workers more opportunities, so they are less likely to separate. For those workers (second column of Table 10) all coefficients are negative and significant. Moreover, a high wage and/or a larger firm reduces the probability of separating significantly more compared to the unconditional case.

The second group of workers are those who remain in the same occupations in the next period. Our setup would imply that these workers only separate due to exogenous reasons and wages and size should not affect the separation probability. Looking at the third column of Table 10 we notice that indeed the impact of wages and firm size on the separation probability is substantially weaker. In particular, the marginal effect of the wage, while still significant, is now an order of magnitude smaller. Similarly, the effect of firm size is significantly smaller: the dummy for medium size firms is no longer significant, while that for large firms is significant and negative, but has a much lower coefficient.

6 Conclusion

This paper argues that observed differences between workers in large and small firms can be accounted for by large firms having bigger internal labor markets.[20] In particular, our framework predicts that workers in large firms are better matched in equilibrium, which leads to higher productivity and wages and lower separation rates, even conditional on wage.

Using data from the 1996 SIPP, we find support for the predictions of our setup: internal mobility is higher in large firms and depends negatively on wages and tenure; wages are higher on average, following an occupational switch within the same employer; workers in large firms switch occupations at higher wage levels and they also obtain higher initial wages in their new occupation; the wage premium is higher for workers with longer tenure; workers who leave large firms, continue to enjoy high wages, but only if they remain in the same occupation; and finally the wage and size effects on the separation probability are significantly larger for workers who switch occupations.

Appendix

A Wage Determination

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

$$\beta J(p,k) = (1-\beta)(V(p,k) - U) \tag{11}$$

where $J(p,k)$ is the value to the firm of employing a worker with posterior p and k tasks remaining. The corresponding Hamilton-Jacobi-Bellman equation is given by:

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

Multiplying eq. (3) by $1-\beta$ and subtracting $(1-\beta) rU$ leads to:

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

[20] In our setup, learning is the driving force behind worker transitions across matches; our results, however, would similarly go through if workers switched matches due to evolving, match-specific productivity.

Multiplying eq. (12) by β, subtracting eq. (13) from it and using the surplus sharing condition, eq. (11), leads to:

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

Finally, by taking the second derivative with respect to p in the surplus sharing condition, eq. (11), and using eq. (14), results in the wage equation, (2).

B Proof of Proposition 2

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:

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

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

$$S(p,k) = \frac{\overline{\alpha}(p) - rU}{r+\delta} + K_1^k p^{\frac{1}{2}-\frac{1}{2}\sqrt{\frac{4+h}{h}}} (1-p)^{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{4+h}{h}}} + K_2^k p^{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{4+h}{h}}} (1-p)^{\frac{1}{2}-\frac{1}{2}\sqrt{\frac{4+h}{h}}}$$

where $h = \frac{\zeta^2}{2(r+\delta)}$ and K_1^k and K_2^k are undetermined coefficients that depend on k. When $p \to 1$ however, $\lim_{p\to 1} K_2^k p^{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{4+h}{h}}} (1-p)^{\frac{1}{2}-\frac{1}{2}\sqrt{\frac{4+h}{h}}} = K_2^k \cdot 1 \cdot \lim_{p\to 1} (1-p)^{\frac{1}{2}-\frac{1}{2}\sqrt{\frac{4+h}{h}}} = +\infty$ which follows from $h > 0$. 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:

$$S(p,k) = \frac{\overline{\alpha}(p) - rU}{r+\delta} + K_1^k p^{\frac{1}{2}-\frac{1}{2}\sqrt{\frac{4+h}{h}}} (1-p)^{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{4+h}{h}}}$$

where K_1^k is an undetermined coefficient. Moreover:

$$S_p(p,k) = \frac{\alpha^G - \alpha^B}{r+\delta} + K_1^k \left(\frac{1}{2} - \frac{1}{2}\sqrt{\frac{4+h}{h}} - p\right) p^{-\frac{1}{2}-\frac{1}{2}\sqrt{\frac{4+h}{h}}} (1-p)^{-\frac{1}{2}+\frac{1}{2}\sqrt{\frac{4+h}{h}}}$$

Consider the case where $k = 0$. Using equation (11), one can rewrite the value matching and smooth pasting conditions (equations (7) and (8) respectively), in terms of $S(\cdot)$ and use them to pin down $\underline{p}(0)$ and K_1^0.

From equation (8), one immediately obtains:

$$K_1^0 = -\frac{\alpha^G - \alpha^B}{r+\delta}\left(\frac{1}{2} - \frac{1}{2}\sqrt{\frac{4+h}{h}} - \underline{p}(0)\right)^{-1} \underline{p}(0)^{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{4+h}{h}}} \left(1-\underline{p}(0)\right)^{\frac{1}{2}-\frac{1}{2}\sqrt{\frac{4+h}{h}}}$$

Substituting for K_1^0 into equation (7) and solving leads to equation (9).

Similarly in the case where $k > 0$, equation (6) leads to:

$$K_1^k = -\frac{\alpha^G - \alpha^B}{r+\delta}\left(\frac{1}{2} - \frac{1}{2}\sqrt{\frac{4+h}{h}} - \underline{p}(k)\right)^{-1} \underline{p}(k)^{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{4+h}{h}}} \left(1-\underline{p}(k)\right)^{\frac{1}{2}-\frac{1}{2}\sqrt{\frac{4+h}{h}}}$$

Substituting for K_1^k into equation (5) leads to equation (10).

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 $V_p\left(\underline{p}(k), k\right) = 0$, this implies that $V_p(p,k) > 0$ for all $p > \underline{p}(k)$.

C Proof of Lemma 5

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

Consider the set of all workers with k tasks remaining who draw the same prior, $p_0 \in \left(\underline{p}(k), 1\right)$. 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:

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

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

Then:

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

For a given $p_0 > \underline{p}(k)$, the weight $\frac{g\left(p_0|p_0>\underline{p}(k)\right)}{1-G\left(\underline{p}(k)\right)}$ is higher when $\underline{p}(k)$, and therefore k (Corollary 3), is higher. Moreover, if $I\left(p_0, \underline{p}(k)\right)$ 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 $I\left(p_0, \underline{p}(k)\right)$ is increasing in $\underline{p}(k)$ and p_0.

Define:

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

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

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

From Athreya and Lahiri (2006), Theorem 14.2.10, part (i),[21] we have that:

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

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:

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

Define a process Y_t which starts from $\underline{p}^2$ and evolves like p_t^1 with the only difference that whenever it reaches $\underline{p}^1$, it is 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:

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

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

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

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

[21] In 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.

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:

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

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:[22]

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

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

References

[1] Antonovics K. and L. Golan (2008): "Experimentation and Job Choice," Mimeo, Carnegie Mellon University, Tepper School of Business.

[2] Athreya, K. B. and S. N. Lahiri (2006): *Measure Theory and Probability Theory*, New York: Springer.

[3] 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.

[4] Brown C. and J. Medoff (1989): "The Employer Size-Wage Effect," *Journal of Political Economy,* 97(5), 1027-1059.

[22] Essentially 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.

[5] Burdett K. and D. T. Mortensen (1998): "Wage Differentials, Employer Size and Unemployment," *International Economic Review,* 39(2), 257-273.

[6] 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.

[7] Gittins, J. C. and D. M. Jones (1974): "A Dynamic Allocation Index for the Sequential Design of Experiments," in *Progress in Statistics*, vol. 1, ed. by Joseph M. Gani, Karoly Sarkadi, and Istvan Vincze. Amsterdam: North-Holland, 241-266.

[8] 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.

[9] 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.

[10] Idson, T. L. (1993): "Employer Size and Labor Turnover," Discussion Paper No. 673, Columbia University, Department of Economics.

[11] 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.

[12] 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.

[13] Jovanovic, B. (1979): "Job Matching and the Theory of Turnover," *Journal of Political Economy, 87(5) Part 1, 972-990.*

[14] Karlin, S. and H. M. Taylor (1981): *A Second Course in Stochastic Processes*, New York: Academic Press.

[15] Liptser, R. and A. Shyryaev (1977): *Statistics of Random Processes,* Vol. 2 Berlin: Springer-Verlag.

[16] Moore, H. L. (1911): *Laws of Wages,* New York: MacMillan.

[17] 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.

[18] Moscarini, G. (2005): "Job Matching and the Wage Distribution," *Econometrica*, 73(2), 481-516.

[19] Ortega, J. (2001): "Job Rotation as a Learning Mechanism," *Management Science,* 47(10), 1361-1370.

[20] Papageorgiou, T. (2009): "Learning Your Comparative Advantages," Mimeo, Penn State University, Department of Economics.

[21] Pastorino, E. (2009): "Learning in Labor Markets and Job Mobility," Mimeo, University of Iowa, Department of Economics.

[22] Pissarides, C. A. (2000): *Equilibrium Unemployment Theory*, 2nd edition, Cambridge MA: MIT Press.