

# Occupational Self-Selection in a Labor Market with Moral Hazard

Berna Demiralp\*

Department of Economics

Old Dominion University

## Abstract

This paper studies the determinants and implications of self-selection when firms imperfectly observe worker effort. The effects of the resulting moral hazard problem on the self-selection mechanism are analyzed in a model in which workers simultaneously choose an employment sector and an effort level. The implications of the model reveal that in the presence of moral hazard, workers' effort decisions become an additional mechanism determining the pattern of selection into sectors. Workers' sector-specific endowments impact sectoral allocation through their effect on workers' comparative advantage as well as their effect on workers' shirking propensity. The model is then used in an empirical application that analyzes workers' self-selection into white collar and blue collar occupations. The estimation results, based on data from the National Longitudinal Survey of Youth, suggest that workers' occupational self-selection leads to higher wages and lower dismissal rates in both occupations, compared to an economy in which workers are randomly assigned to each occupation. The difference in dismissal rates between the two occupations is driven by the higher expected productivity in the white collar sector. The positive effects of occupational sorting diminish as the labor market becomes increasingly characterized by moral hazard. Results also suggest that human capital investments in skills that are most relevant to blue collar jobs may generate higher wages and lower dismissal rates in both white collar and blue collar occupations.

---

\*Old Dominion University, Department of Economics, Constant Hall, Norfolk, VA, 23529 USA (e-mail: bdemiral@odu.edu). I am grateful to Robert Moffitt and Chris Flinn for their guidance and valuable comments. I would also like to thank Matthew Shum, Insan Tunali, seminar participants at the Johns Hopkins University, New York University, SUNY Stony Brook and Koc University, participants at the 2005 European Society for Population Economics Annual Conference and the 2007 Econometric Society North American Summer Meetings and two anonymous referees for their very helpful suggestions and comments.

# 1 Introduction

In an economy with heterogeneous workers, the self-selection of workers into different sectors of the economy plays an important role in explaining the economic outcomes observed in different sectors of the economy. Starting with Roy's (1951) seminal paper, a large literature has emphasized that the distribution of workers in a given sector is endogenous on workers' sectoral choices, and it has sought to identify the determinants of these choices<sup>1</sup>. A common assumption embedded in this literature is that firms and workers have symmetric information about workers' productivity on the job. Under this assumption, the distribution of worker types in a sector represents the distribution of productivity and wages in that sector. However, when there is asymmetric information between workers and firms regarding worker productivity and information is costly, the strength of the relationship between worker type, productivity and wages may weaken, and the self-selection mechanism has to be augmented to reflect this change. The standard models of sectoral choice that are based on symmetric information do not address how the informational advantage of workers in labor markets with asymmetric information affects workers' sectoral or occupational choices and whether the informational asymmetry reinforces or diminishes the effects of self-selection in the economy.

The purpose of this paper is to investigate the determinants and consequences of self-selection in a labor market characterized by moral hazard. The framework for analysis is a labor market in which firms imperfectly observe worker output, workers have an incentive to shirk, and firms dismiss workers who are caught shirking during random monitoring. In such a labor market, workers make decisions not only on sector but also on effort level in each sector. As a result, an analysis of self-selection in such a labor market requires the examination of both sector and effort decisions simultaneously. This paper presents a model in which workers take into account their effort decisions in each sector when choosing a sector and analyzes the implications of their simultaneous decisions on economic outcomes.

The model presented in this paper is based on a shirking model by Flinn (1997) which features moral hazard and heterogeneous workers in a single sector. I extend Flinn's model to two sectors and examine workers' sectoral choice. In both sectors, the systematic dismissal of shirkers leads to an increase in the average productivity and wage of the remaining workers. In equilibrium, wages and dismissal rates experienced in each sector are consistent with both the workers' supply of effort on the job and their sectoral

---

<sup>1</sup>Roy's model has been applied to explain various labor market issues, including sectoral choice and wage distribution (Heckman and Sedlacek, 1985), wage inequality (Gould, 2002), and patterns of schooling, employment and occupational choice (Keane and Wolpin, 1997). Sattinger (1993) provides a review of the Roy's model and the related literature.

choice. The model is then used in an empirical application, examining workers' self-selection into white collar and blue collar occupations and the effect of their occupational sorting on the wages and dismissal rates observed in the two occupations.

This study contributes to the literature on sectoral choice and self-selection in several ways. First, it provides a characterization of the sectoral allocation of workers when the labor market exhibits a moral hazard problem. The Roy model of sectoral allocation, which forms the basis of much of the research on this topic, has identified workers' relative skills in different sectors as one of the main drivers of worker allocation across sectors. This mechanism, usually referred to as workers' comparative advantage, together with the return to skills in the two sectors, determines how the distribution of worker types in a sector is related to the population distribution when workers and firms have symmetric information. The model presented in this paper shows that in the presence of moral hazard, workers' effort decisions become an additional mechanism, determining the pattern of selection into sectors. When output is imperfectly observed, workers' sector-specific skill endowments affect self-selection not only through their effect on workers' comparative advantage but also through their effect on workers' effort decisions. Comparative advantage continues to be the primary allocation mechanism among workers who decide to exert effort. However, the effect of comparative advantage on self-selection is weaker among shirkers who take into account other factors, such as monitoring intensities in their sectoral choice. Therefore, when informational asymmetries are present, workers' effort decisions have to be considered in linking the population distribution of workers to the distribution of worker types in each sector.

This paper also contributes to the literature on self-selection by focusing on the relationship between sectoral choice and dismissals. The literature on occupational/industrial choice and self-selection was originally developed to explain how the sectoral allocation of workers can give rise to the occupational and industrial distribution of wages and earnings. As a result, the effect of self-selection on worker turnover behavior across occupations and industries has received less attention in this literature. There does exist a body of research that formally describes the behavioral relationship between occupational selection and turnover (e.g. Miller, 1984; McCall, 1990; Neal, 1999; Moscarini, 2001). This paper contributes to the current literature by examining the relation between self-selection and dismissals that are due to poor worker performance or malfeasance. Here, dismissals occur when workers are caught to be shirking during random monitoring by the firm. Therefore, dismissals are contingent on workers' effort decisions, which are

explicitly modeled within a moral hazard framework<sup>2</sup>.

Dismissals are used as part of a mechanism to induce worker effort when the complexity of the employment relationship makes it difficult for the firm to explicitly link pay to performance. Recent empirical work suggests that worker-firm relationships that are based on explicit contracts, specifying performance expectations, are not prevalent in the labor market. For example, MacLeod and Parent (1999) find that only about 1-5 percent of workers receive commission or piece rates in the United States. A more common type of employment relationship is one where it is impossible to fully specify performance expectations and the firm has to rely on rewards (such as bonuses and promotions) and punishments (such as dismissals) based on subjective evaluation to motivate workers. In addition, firms are more likely to use dismissals rather than bonus and promotion when the local labor market is not tight (MacLeod and Parent, 1999).

After developing the theoretical model, I present an empirical application investigating the allocation of workers across blue collar and white collar occupations. I estimate the structural parameters of the model by maximum likelihood using data from the National Longitudinal Survey of Youth (NLSY). The specification of the likelihood function requires the numerical computation of equilibrium wage sequences. The equilibrium wage sequence in each sector is computed by means of two sets of fixed point iterations, one nested within the other. The structural model performs well in fitting data on occupational choice, wage and dismissal patterns observed in the NLSY.

There are two key mechanisms governing the equilibrium wage and dismissal sequences in this model. First, in an occupation in which output is imperfectly observed, random monitoring by the firm leads to a systematic dismissal of unproductive workers in each period. This dynamic selection on the skill distribution of workers causes the average productivity among remaining workers to increase and their wages to be bid up over time. It also generates the observed dismissal rate, which depends on the prevalence of shirking and the rate at which the firm monitors its workers. As inefficient workers are eliminated and the supply of effort by remaining workers increases, the hazard rate of dismissals decreases over time. Second, given the wage and dismissal sequences in each occupation, workers make a one-time occupational choice based on their lifetime value of employment in each occupation. Workers' sorting into occupations can cause the skill distribution in each occupation to differ from the population distribution, thus affecting wages and dismissal rates. The empirical application presented in this paper demonstrates how these two mechanisms (workers'

---

<sup>2</sup>The effectiveness of the monitoring depends on its randomness; therefore, the model does not allow workers to anticipate dismissals and leave before being dismissed.

self-selection into occupations and the dynamic selection on the worker distribution due to systematic dismissal of inefficient workers) interact to bring about the observed wages and dismissal rates in the white and blue collar occupations.

The estimation results suggest that workers' self-selection of occupations leads to higher wages and lower dismissal rates in both occupations, compared to an economy in which workers are randomly assigned to each occupation. This finding is consistent with the conclusion that worker's sorting into occupations has a positive impact on the skill distribution in each occupation by increasing the average productivity in each occupation. Although workers' selection of occupations has positive effects on the worker composition in each occupation, these positive effects may be diminished by the existence of moral hazard. Analysis results suggest that if occupational selection by the same population took place in a labor market with no moral hazard, workers in both occupations would predominantly come from the higher ability tail of the population distribution, displaying a much stronger level of positive selection. Furthermore, the monitoring rate in the white collar occupation is estimated to be higher than the monitoring rate in the blue collar occupation, indicating that a white collar worker faces a higher probability of being monitored by the firm. Given that dismissal rates are determined by the monitoring rate and the prevalence of shirking in an occupation in this model, this result suggests that lower dismissal rates observed in the white collar occupation is driven by the higher expected productivity in that sector.

I also provide several comparative static exercises aimed at investigating the impact of model's parameters on wages and dismissal rates in both occupations. The results reveal that while a higher monitoring rate in the blue collar occupation leads to higher wages and lower dismissal rates in the blue collar occupation, it leads to lower wages and higher dismissal rates in the white collar occupation by making the white collar occupation relatively more attractive to workers with high propensity to shirk. Furthermore, a lower average disutility of effort for white collar jobs in the population leads to higher wages in the white collar occupation but lower wages in the blue collar occupation, while a similar decrease in the average disutility of effort for blue collar jobs in the population, leads to higher wages in both occupations. This finding suggests that policies designed to enhance human capital investments in skills relevant to white collar jobs may generate adverse selection into blue collar occupation whereas similar investments in skills used in blue collar jobs may increase productivity and decrease dismissals in both occupations.

This paper is organized in six sections. The next section presents the model and the results of representative simulations. Section 3 describes the estimation methodology and identification issues. Sections 4 and

5 present the description of the dataset and the discussion of the results, respectively, and Section 6 includes concluding remarks.

## 2 The Model

### 2.1 The Set-up

The labor market consists of a primary and a secondary sector, both of which are made up of perfectly competitive firms. The primary sector is characterized by the moral hazard problem. Firms cannot perfectly observe their workers' effort levels, possibly due to the lack of an employee-specific output measure, which is necessary for observing effort through output. The utility flow to a worker in each period is given by  $w_t - e_t$ , where  $w_t$  is the wage that he receives and  $e_t$  is the amount of effort that he exerts in period  $t$ . Due to the disutility of effort, workers have an incentive to shirk. Firms randomly monitor their workers in an attempt to detect shirking in their workforce. The punishment for shirking is dismissal from the firm.

There are two types of firms in the primary sector. Workers in primary sector firm  $j$  ( $= 1, 2$ ) face a constant probability,  $\pi_j$ , of being monitored by the firm. The monitoring rate is a measure of the intensity with which firms provide supervision of employees. Differences in monitoring rates reflect differences in monitoring technologies and monitoring costs across firms. Primary sector firm  $j$  has an exogenous output price  $\rho_j$ .

Workers who are dismissed from primary sector firms find jobs in the secondary sector of the labor market. In the secondary sector, firms perfectly observe effort via employee-specific output; therefore, there is no moral hazard. The amount of effort needed to produce one unit of output in this sector equals the output price, so the utility flow for each worker is zero. Furthermore, due to perfect competition, the secondary sector wage equals the output price.

The secondary sector consists of two types of firms. Workers who are dismissed from primary sector firm  $j$  ( $= 1, 2$ ) find employment in the secondary sector firm  $j$  and earn wage given by  $w_j^s$ . The secondary sector is an absorbing state; workers in this sector cannot be rehired by either type of primary sector firm. As I will show below, workers in the primary sector receive a positive utility flow in each period; thus, workers voluntarily start their labor market careers in the primary sector.

Each worker is endowed with two productive inefficiency indices,  $\xi_1$  and  $\xi_2$ , which determine the amount of effort needed to complete the sector-specific tasks in primary sector firms 1 and 2, respectively.

Workers are heterogeneous with respect to their productive inefficiency indices, and  $H(\xi_1, \xi_2)$  denotes the population distribution of productive inefficiency indices. Workers know their own endowments of productive inefficiency indices in the beginning of their labor market careers. However, firms do not observe worker types. The only information available to firms is whether the individual has ever been detected shirking and dismissed.

The production function for a worker of type  $\xi_j$  in the primary sector firm  $j$  in period  $t$  is

$$y_j(e_t; \xi_j) = \begin{cases} 1 & \text{iff } e_t \geq \xi_j \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where  $e_t$  is the amount of effort exerted in period  $t$ . Therefore, a worker's productive inefficiency index,  $\xi_j$ , can be interpreted as the minimum amount of effort that the worker has to exert in order to produce one unit of output. Since workers receive disutility from putting forth effort, they either exert the minimum effort possible to produce output (i.e.  $\xi_j$ ) or they exert no effort at all and shirk<sup>3</sup>. Finally, workers have infinite horizon, and they discount the future by a factor,  $\beta$ .

## 2.2 Worker's Decision

Workers in this model make two types of decisions: 1) at the beginning of their labor market careers, they decide for which primary sector firm to work, 2) in each period of their employment in the primary sector, they decide whether to exert effort or shirk. I first explain the work/shirk decision of the worker conditional on his employment in firm  $j$ . Then, I discuss worker's decision regarding the type of primary sector firm for which to work.

A worker employed in primary sector firm  $j$  decides whether to work or shirk in each period according to the following maximization problem. The value of employment for a worker of type  $\xi_j$  in firm  $j$  in period  $t$  is

$$V_{jt}(\xi_j) = \max \{w_{jt} - \xi_j + \beta V_{j,t+1}(\xi_j); w_{jt} + \beta(1 - \pi_j)V_{j,t+1}(\xi_j)\}, \quad j = 1, 2 \quad (2)$$

where the first argument is the value of working, and the second argument is the value of shirking. This

<sup>3</sup>The production function can be equivalently interpreted as stochastic in the following way:

$$y_j(e_t; \xi_j) = \begin{cases} 1 & \text{with probability } 1 \text{ if } e_t \geq \xi_j \\ 1 & \text{with probability } 1 - \pi_j \text{ if } e_t < \xi_j \\ 0 & \text{with probability } \pi_j \text{ if } e_t < \xi_j \end{cases}$$

function reflects the assumption that worker's utility flow in the secondary sector is zero. If we assume that the wage sequence is monotonically increasing over time, a worker of type  $\xi_j$  will

$$\begin{aligned} &\text{work if } \xi_j \leq \xi_{jt}, \text{ and} \\ &\text{shirk otherwise} \end{aligned} \tag{3}$$

where the threshold level of productive inefficiency in firm  $j$  in period  $t$ ,  $\xi_{jt}$ , equals

$$\xi_{jt} = \frac{\beta\pi_j(1-\beta)}{1-\beta+\beta\pi_j} \left( \sum_{s=t+1}^{\infty} \beta^{s-t+1} w_{js} \right), \quad j = 1, 2. \tag{4}$$

Furthermore, in the beginning of his labor market history, the worker makes a one-time decision on the type of primary sector firm based on his likelihood of dismissal and the wages offered by different firms. Worker chooses firm 1 if  $V_{1,t=1} > V_{2,t=1}$ ; otherwise he chooses firm 2. This condition can equally be expressed in terms of the worker's productive inefficiency indices as follows: Worker chooses firm 1 if

$$\xi_1 < \xi_1^*(\xi_2; \{w_1\}_{t=1}^{\infty}, \{w_2\}_{t=1}^{\infty}, \pi_1, \pi_2) \tag{5}$$

Based on this selection rule, the marginal distribution of  $\xi_1$  in period 1 among workers who choose firm 1 can be calculated as

$$f_{1,t=1}(\xi_1) = \int_0^{\infty} \frac{h_1(\xi_1)}{H_1(\xi_1^*(\xi_2))} dH_2(\xi_2) \tag{6}$$

The marginal distribution of  $\xi_j$  in firm  $j$  changes over time in a systematic way as a constant proportion of shirking workers are detected and dismissed in each period. In particular, if we consider a cohort of workers who start to work in firm  $j$  at the same time, the marginal distribution of workers remaining in the cohort changes over time with the mass point of the distribution moving toward lower levels of productive inefficiency. The cdf of worker types remaining in the cohort at the end of period  $t$  in firm  $j$  can be expressed in terms of the cdf of worker types in firm  $j$  in period one. When the sequence of  $\{\xi_{jt}\}$  is increasing, the relationship between  $F_{jt}(\xi_{jt})$  and  $F_{j,t=1}(\xi_j)$  is given by the following equation (Flinn, 1997):

$$F_{jt}(\xi_{jt}) = 1 - (1 - F_{j,t=1}(\xi_j)) A_{jt} \left( \{\xi_{js}\}_{s=1}^{t-1} \right)$$

where

$$A_{jt} \left( \{\xi_{js}\}_{s=1}^{t-1} \right) = \left\{ 1 + \frac{\pi_j}{(1-\pi_j)^{t-1}} F_{j,t=1}(\xi_{j,t=1}) + \cdots + \frac{\pi_j}{1-\pi_j} F_{j,t=1}(\xi_{j,t-1}) \right\}^{-1}$$

In this model, workers' endowments of  $\xi_1$  and  $\xi_2$  affect their selection of a primary sector firm in two distinct ways. First, among non-shirkers, firm choice is determined to a large extent by the relative disutility of effort associated with employment in the two firms. What derives the sorting in this case is workers' comparative advantage in the two types of firms, measured by relative disutility of effort in each firm. Second, a worker's effort decision plays an important role in his firm choice because it determines which factors are taken into account by the worker in his primary sector firm decision. For instance, for a worker who exerts effort in both firms, his relative disutility of effort in the two firms plays an important role in his firm decision. In contrast, disutility of effort does not enter a shirker's decision-making process since he exerts no effort and thus receives no disutility from it. Instead, a shirker considers how closely he will be monitored in each firm in order to assess his probability of dismissal. Therefore, workers' self-selection of firms depends in part on the sequence of work/shirk decisions that they would make in the two firms and consequently on their relative disutility of effort in these occupations.

### 2.3 Firm's Decision

In this subsection, I will continue to consider the labor market experiences of a cohort of workers, who enter the firm at the same time. The firm cannot observe the productive inefficiency index of each worker in the cohort, so it cannot observe whether each worker is working or shirking. However, the firm observes the marginal distribution of productive inefficiencies within the cohort,  $f_{jt}(\xi_j)$ , the threshold level of productive inefficiency in each period,  $\xi_{jt}$ , and thus the expected productivity in a cohort, given by  $F_{jt}(\xi_{jt})$ .

Due to the zero profit condition, the firm pays cohort members the value of the expected productivity in the cohort. Therefore, the wage that firm  $j$  offers to the members of a cohort in period  $t$  of their employment is given by Equation 7<sup>4</sup>.

$$w_{jt} = \rho_j F_{jt}(\xi_{jt}), \quad j = 1, 2. \quad (7)$$

As a result, everyone in a given cohort earns the same wage although they make different effort decisions

---

<sup>4</sup>The wage contracts described in this paper are individual wage contracts. Wage contracts that depend on group output are not considered.

based on their productive inefficiency indices.

## 2.4 Equilibrium

The Nash equilibrium wage sequence in primary sector firm  $j$  is defined as the fixed point of the following operator:

$$T(\{w_{jt}\}) = \begin{bmatrix} \rho_j F_{j,t=1}(\xi_{j,t=1}) \\ \rho_j F_{j,t=2}(\xi_{j,t=2}) \\ \vdots \\ \rho_j F_{j,t=\tau}(\xi_{j,t=\tau}) \\ \vdots \end{bmatrix}, \quad j = 1, 2. \quad (8)$$

where  $\xi_{jt}$  is given by Equation 4. The fixed point of  $T(\{w_{jt}\})$  gives the equilibrium wage sequence in firm  $j$  conditional on  $F_{j,t=1}$ , the post-selection marginal distribution of  $\xi_j$  in firm  $j$ . The parameters of  $F_{1,t=1}$  and  $F_{2,t=1}$ , are in turn fixed points of the operator given in Equation 9, which completes the characterization of the equilibrium in this model. Let  $\nu_1$  and  $\omega_1$  be the parameters that characterize  $F_{1,t=1}$ , and  $\nu_2$  and  $\omega_2$  be the parameters that characterize  $F_{2,t=1}$ . Then,

$$\begin{bmatrix} F_{1,t=1}(\nu_1, \omega_1) \\ F_{2,t=1}(\nu_2, \omega_2) \end{bmatrix} = \begin{bmatrix} \int_0^\infty \frac{h_1(\xi_1)}{H_1(\xi^*(w_1(F_{1,t=1}(\nu_1, \omega_1)), w_2(F_{2,t=1}(\nu_2, \omega_2))))} dH_2(\xi_2) \\ \int_0^\infty \frac{h_2(\xi_2)}{H_2(\xi^*(w_1(F_{1,t=1}(\nu_1, \omega_1)), w_2(F_{2,t=1}(\nu_2, \omega_2))))} dH_1(\xi_1) \end{bmatrix} \quad (9)$$

Every iteration in solving the fixed point problem in Equation 9 involves the computation of the equilibrium wage sequence in Equation 8. Therefore, the algorithm to compute the fixed points of  $T(\{w_{jt}\})$  is nested within the fixed point algorithm to compute the parameters of  $F_{1,t=1}$  and  $F_{2,t=1}$ .

This equilibrium has several important features. First, the equilibrium wage sequence in firm  $j$  depends not only on the output price, worker distribution and monitoring rate in firm  $j$ , but also on the wages and parameters in firm  $k$ . Intuitively, this result captures the fact that when workers select firm types, they take into account the wages and monitoring rates in both types of firms. Therefore, the marginal distribution of workers in a given firm depends on the wages and monitoring rates observed in the entire primary sector.

The second feature of this equilibrium is that the equilibrium wage sequence is monotonically increasing over time due to the systematic dismissal of relatively inefficient workers. In each period, a proportion

of relatively inefficient workers are dismissed. Therefore, the remaining cohort is made up of a lower proportion of relatively inefficient workers. As the expected productive inefficiency in the cohort falls, the expected worker productivity and wages rise. Furthermore, a worker's effort decision is not constant over time. With wages monotonically increasing over time, different workers stop shirking at different times depending on their productive inefficiency index with high productivity workers (low productive inefficiency workers) deciding to put forth effort earlier than others.

The systematic dismissal of inefficient workers also affects the hazard rates of dismissal in the primary sector. The separation hazard is determined by the constant monitoring rate and the proportion of shirkers, which decreases over time. The monotonic decrease in the proportion of shirkers leads to a similarly monotonic decline in the hazard rate of dismissal. Although the model does not generate the non-monotonic hazard rates of job separation noted in other studies (e.g. Jovanovic, 1979; Sauer, 1998), the data used in this study does not display strong non-monotonicity of separation hazard rate when job separations are limited to dismissals.

Finally, the system in Equation 8 is not recursive. Although  $\xi_{jt}$  depends only on the wage sequence starting in period  $t + 1$ ,  $F_{jt}$  depends on the entire wage sequence,  $\{w_{jt}\}_{t=1}^{\infty}$ . The computation of the equilibrium wage sequences is discussed in the appendix.

The structure of this model carries several assumptions that should be considered in interpreting the equilibrium results. One of these assumptions is the exogeneity of the monitoring rate. Firms are assumed to have already made decisions on monitoring intensities before workers choose firms, and they are not allowed to adjust their monitoring rates in response to workers' productivity during the period covered by the model. Although examining firms' optimal choices of monitoring intensity is beyond the scope of this paper, it is important to consider how endogenous monitoring rates might affect the equilibrium described by the model. The model predicts that the monitoring rate is one of the key parameters on the basis of which workers self-select into sectors. For instance, a low monitoring rate may lead to higher shirking not only because it is not an adequate punishment for being caught but it can attract low-ability workers who are more likely to shirk. Then, would an endogenous monitoring rate neutralize the adverse selection effects by allowing firms to respond to the negative selection by increasing their monitoring rates? Simulations performed on a version of the model where the monitoring intensity is semi-endogenous shows that the effects of self-selection do not necessarily diminish, and they can in fact become more exaggerated when firms are allowed to change their monitoring intensity.

Consider a case in which the monitoring rate in firm  $j$  is inversely proportional to the average cohort productivity, so the firm adjusts its monitoring intensity upward when the productivity is low (e.g.  $\pi_{jt} = g_j(F_{jt}(\xi_{jt}))$  and  $g'_j(F_{jt}(\xi_{jt})) < 0$ ). Simulations show that when the monitoring rate is made semi-endogenous in this way, the adverse selection effects do not necessarily disappear. In some cases, at least one of the firms can experience an increase in the proportion of shirkers in its workforce when the monitoring rate becomes semi-endogenous. In those cases, the reason for the increase in shirking propensity stems from the fact that the two types of firms have different capabilities to adjust their monitoring rates ( $g'_1(F_{1t}(\xi_{1t})) \neq g'_2(F_{2t}(\xi_{2t}))$ ). The firm that has less flexibility in responding to changes in the productivity of its workforce by changing its monitoring rate may become attractive to lower ability workers who are discouraged by the other firm's speed in adjusting its monitoring rate.

The second feature of the model that should be noted in interpreting the results is that the secondary sector is an absorbing state; workers in this sector cannot be rehired by either type of primary sector firm. This condition has two components: workers fired from firm  $j$  cannot be rehired by firm  $j$ , and workers fired from firm  $j$  cannot be rehired by firm  $k$  ( $j \neq k$ ). The former is an equilibrium outcome of the single-sector version of the model when wages are increasing monotonically over worker's life cycle (Flinn, 1997). The intuition is that since firms make zero profit on the workers remaining in the cohort, rehiring workers who have been dismissed would cause the firm to make negative profits.

The latter condition that workers dismissed from firm  $j$  cannot be hired by firm  $k$  ( $j \neq k$ ) is an assumption imposed for reasons of tractability. Relaxing this assumption presents both computational and theoretical challenges. In the absence of this assumption, the solution to the critical value,  $\xi_{jt}$ , in Equation 4 would be a function of the critical value,  $\xi_{kt}$ , significantly increasing the computational burden involved in computing the equilibrium wage sequences. Furthermore, allowing inter-occupational mobility of workers after dismissal would require additional theoretical developments since such mobility can create an opportunity for strategic behavior on the part of the other firm. For example, the fact that a worker was dismissed from firm 1 in his  $t$ th period on the job would convey information to firm 2 about worker's potential productivity in firm 2. Based on the correlation between  $\xi_1$  and  $\xi_2$ , firm 2 may choose to hire only a subgroup of workers dismissed from firm 1. Nevertheless, in general, if workers dismissed from firm  $j$  are allowed to be hired by firm  $k$ , their value of employment in  $j$  conditional on shirking increases. As a result, one can expect the critical value,  $\xi_{jt}$ , to be lower and the proportion of workers shirking in firm  $j$  to be higher when this assumption is relaxed. Extending the model to allow for inter-occupational mobility of workers after

dismissal would be an important contribution.

Another feature of this model is that the incentive not to shirk is provided by the chance to be permanently demoted to the secondary sector. This type of permanent reputation loss as a punishment for shirking is in contrast to Shapiro and Stiglitz's shirking model in which workers who are caught shirking by the firm face a transitory period of unemployment. As explained above, in a dynamic equilibrium framework in which dismissals are an equilibrium outcome, transitions back into the primary sector presents computational and theoretical challenges that are left for future work<sup>5</sup>.

In addition, by focusing on dismissals that are due to poor worker performance or malfeasance, this model abstracts from other factors that may influence dismissals, such as poor worker-firm matches and exogenous shocks to labor demand. Also, the wage growth in this model is generated by the systematic dismissal of unproductive workers from the firm. The addition of secondary sources of dismissals and wage growth to the model is considered in Demiralp (2007).

Finally, workers are assumed to perfectly observe their productive inefficiency indices in each firm, eliminating the case in which workers learn about their productive inefficiency indices.

## 2.5 Simulations

As shown in the previous section, one of the key determinants of the pattern of selection into primary sector firms is the variance-covariance structure of the population distribution of worker types. In this section, I present results of representative simulations, demonstrating how the correlation between  $\xi_1$  and  $\xi_2$  in the population affects the direction of self-selection in the labor market. In particular, I compare the outcome when the two random variables have a high positive correlation to the case when they are negatively correlated. I assume that the population distribution of worker types is characterized by a bivariate lognormal distribution<sup>6</sup>.

Figures 1 and 2 present the population marginal distributions of  $\log \xi_1$  and  $\log \xi_2$  as well as the post-selection marginal distributions in each firm when  $\log \xi_1$  and  $\log \xi_2$  have a correlation coefficient of 0.83. These results show that the pdf of  $\log \xi_1$  among people who choose firm 1 lies above the lower tail of the population distribution and below the upper tail of the population distribution, indicating a positive

---

<sup>5</sup>Albrecht and Vroman (1998) present an equilibrium model of shirking with unemployment and dismissals in equilibrium, but it is developed in a static framework.

<sup>6</sup>The full set of parameters used in the simulation are as follows:  $\rho_1 = 30$ ;  $\rho_2 = 25$ ;  $\pi_1 = \pi_2 = 0.15$ ;  $\mu_1 = 2$ ;  $\mu_2 = 1.2$ ;  $\alpha_1^2 = 1.8$ ;  $\alpha_2^2 = 0.8$ ;  $\alpha_{12} = 0.9$ .

selection into firm 1. Therefore, workers with lower productive inefficiency indices in firm 1 have a higher tendency to choose firm 1. In contrast, simulation results in Figure 2 show that workers in firm 2 come disproportionately from the right tail of the population distribution of  $\xi_2$ . Therefore, workers with relatively high levels of productive inefficiency in firm 2 exhibit a relatively higher tendency to choose firm 2. Figure 2 also reveals that workers with the highest levels of  $\xi_2$  are likely to choose firm 1; nevertheless, the majority of the firm 2 workforce consists of workers with relatively high levels of  $\xi_2$ .

These results are consistent with those of the previous research which has applied Roy's model and examined its properties in labor markets with symmetric information (e.g. Willis, 1986; Sattinger, 1993; Borjas, 1987). When  $\xi_1$  and  $\xi_2$  have a strong positive correlation, as in the standard Roy's model, self-selection tends to lead to higher average productivity in firm 1, and lower average productivity in firm 2. Since workers with low productive inefficiency in one firm also tend to have low productive inefficiency in the other firm, positive selection into firm 1 is associated with negative selection into firm 2.

When the correlation coefficient between  $\log \xi_1$  and  $\log \xi_2$  in the population distribution is -0.4, self-selection generates positive selection into both firms as shown in Figures 3 and 4. The post-selection marginal distributions of worker types in both firms lie above the lower tail of the population distribution and below the higher tail of the population distribution. These results indicate that in the case of a negative correlation between  $\xi_1$  and  $\xi_2$ , workers are likely to choose firms in which they have lower productive inefficiency. Similar to the implications of the standard Roy's model under symmetric information, negative correlation between  $\xi_1$  and  $\xi_2$  tends to lead to an outcome in which each firm contains the best workers when self-selection occurs in the presence of moral hazard.

### 3 Estimation

I estimate the structural parameters of the model using maximum likelihood. The data used in the estimation procedure consist of workers' occupational choices, wages, and information on whether they were dismissed in each period. Before specifying the likelihood function, I will discuss several empirical issues that are addressed in mapping the theoretical model to the data.

First, I translate workers' selection of firm types in the model to the selection of occupations in the estimation by assuming that firm 1 employs only white collar workers and firm 2 employs only blue collar workers. This characterization applies to a labor market in which white collar and blue collar workers

employed in a firm produce separate goods and have no interaction in the production process. Furthermore, the monitoring technology involved in monitoring white collar workers is different from the monitoring technology for blue collar workers.

Second, the model predicts that workers with the same tenure in a given occupation earn the same wage. In order to account for the variation in wages observed among workers with the same tenure in an occupation, I add measurement error to wages. Let  $p = 1$  if the wage draw is from the primary sector, and  $p = 0$  if the wage draw is from the secondary sector. Then, the measurement error takes on the following form:

$$\ln w_{jt} = p(\ln w_{jt}^*) + (1 - p) \ln w_j^s + \varepsilon_t, \quad j = 1, 2 \quad (10)$$

where  $w_{jt}$  is the reported wage,  $w_{jt}^*$  the primary sector wage predicted by the model, and  $w_j^s$  is the secondary sector wage of a worker dismissed from primary sector firm  $j$ .  $\varepsilon_t$  is independently and identically distributed over time according to a normal distribution with mean zero and standard deviation,  $\sigma_\varepsilon$ .

Another issue that should be addressed in the estimation is the number of dismissals per worker. According to the model, a worker experiences only one dismissal due to malfeasance in his labor market career. Yet, roughly 26 percent of the sample report having more than one dismissal during the sample period. In cases of multiple dismissals over one's labor market career, I assume that the first dismissal experienced by the worker during the sample period is due to shirking. Dismissals that occur after the first one are assumed to occur due to labor demand shocks when the worker is in the secondary sector. Furthermore, workers are assumed to find a new job immediately when they are dismissed in the secondary sector. Thus, they do not face any repercussions of subsequent dismissals that occur after the first one.

Finally, I assume that worker types in the population are distributed according to a bivariate lognormal distribution. The existence of a unique equilibrium wage sequence among the class of increasing wage sequences requires the distribution function of productive inefficiency indices to be concave (Flinn, 1997). The lognormal distribution satisfies the concavity condition.

The likelihood function requires the numerical computation of equilibrium wages based on the model's parameters according to the algorithm included in the appendix. The equilibrium wage sequence, together with the parameters of the model, can then be used to calculate the likelihood function.

Let  $\theta$  be the set of the model's parameters. Then, the likelihood contribution of sample member  $i$  working

in sector  $j$  ( $j \neq k$ ) is given by

$$L_i = \Pr(V_j > V_k, \{\ln w_{ijt}\}_{t=1}^T, \{d_{ijt}\}_{t=1}^T; \theta) = \Pr(V_j > V_k, \{\ln w_{ijt}\}_{t=1}^T | \{d_{ijt}\}_{t=1}^T; \theta) \cdot \Pr(\{d_{ijt}\}; \theta) \quad (11)$$

where  $V_j$  is the value of working in occupation  $j$ ,  $\{w_{ijt}\}_{t=1}^T$  is the worker's reported wage sequence, and  $\{d_{ijt}\}_{t=1}^T$  is his dismissal sequence, indicating whether the worker was dismissed or not in occupation  $j$  in period  $t$ . Nelder-Mead simplex algorithm is used in maximizing the likelihood. The full specification of the likelihood function is given in the appendix.

### 3.1 Identification

The identification of the model's parameters is obtained using data on workers' occupational choices, wages and the empirical hazard rates of dismissal in each occupation over the sample period. The parameters to be estimated are the occupation-specific monitoring rates ( $\pi_1, \pi_2$ ), output prices ( $\rho_1, \rho_2$ ), the parameters of the population heterogeneity distribution ( $\mu_1, \alpha_1, \mu_2, \alpha_2, \alpha_{12}$ ), and the secondary sector wages for workers who are dismissed from each occupation ( $w_1^s, w_2^s$ ). As described in the Model section,  $H_1(\xi_1)$  and  $H_2(\xi_2)$  denote the marginal distribution of worker types in the population while  $F_1(\xi_1)$  and  $F_2(\xi_2)$  indicate the marginal distributions in occupations 1 and 2<sup>7</sup>.

Wages and dismissal rates observed in occupation  $j$  are used in identifying the output price and the monitoring rate in occupation  $j$  as well as the parameters of  $F_j(\xi_j)$  in the following fashion. Equations 12 and 13 give the wage and hazard rate of dismissal in occupation  $j$  in period  $t$ .

$$w_{jt} = \rho_j F_{jt}(\xi_{jt}) \quad (12)$$

$$h_{jt} = \pi_j (1 - F_{jt}(\xi_{jt})) \quad (13)$$

$F_{jt}(\xi_{jt})$  indicates the percentage of workers who choose to exert effort in the cohort in period  $t$ . Then, according to Equation 12, when everyone in the cohort chooses to work, the wage reaches its upper limit at  $\rho_j$ . Although this threshold case helps with the identification of  $\rho_j$  as the limiting wage, we do not actually

<sup>7</sup> $F_j(\xi_{jt})$  was denoted as  $F_{j,t=1}(\xi_{jt})$  in the previous discussion. The " $t = 1$ " subscript is dropped here in order to simplify the notation.

observe it in the data, and therefore, we have to consider identifying the parameters of the model when  $F_{jt}(\xi_{jt})$  is between 0 and 1 for all  $t$ .

It is clear that one cross-section of the wage and dismissal rate data for occupation  $j$  would not be sufficient to identify  $\rho_j, \pi_j$ , and  $F_{jt}(\xi_{jt})$  since in that case there would be two equations and three unknowns. In order to disentangle  $\rho_j$  and  $\pi_j$  from  $F_{jt}(\xi_{jt})$ , observations from more time periods are needed. Since  $F_{jt}(\xi_{jt})$  is a function of  $F_j(\xi_{jt})$ , we can increase the number of equations by considering more time periods without increasing the number of unknowns. The feature of the model that is crucial in identification is that the time-path of wages and dismissal rates are determined solely by  $F_{jt}(\xi_{jt})$ . Therefore, the panel nature of the data can be exploited to identify the time-variant component,  $F_{jt}(\xi_{jt})$ , and the time-invariant factors,  $\rho_j$  and  $\pi_j$ , can be identified given  $F_{jt}(\xi_{jt})$  and Equations 12 and 13. For example, consider having data on two time periods,  $t$  and  $t + 1$ . In that case, the ratio of  $\frac{w_{jt}}{w_{jt+1}}$  identifies  $F_{jt}(\xi_{jt})$ , and then  $\rho_j$  and  $\pi_j$  can be backed out from Equations 12 and 13. The parametric assumption on  $F_{jt}(\xi_{jt})$  is critical in identifying these parameters. Thus, occupation-specific parameters,  $\rho_j, \pi_j, \nu_j, \omega_j$ , can be identified given information on wages and hazard rates of dismissal only in occupation  $j$ .

The identification of the parameters of  $H(\xi_1, \xi_2)$  is obtained through the set of equilibrium conditions in Equation 9 given the parameters of  $F_1$  and  $F_2$ . Therefore, the identification of the parameters of  $H(\xi_1, \xi_2)$  require data in *both* occupations. As in the case of  $F(\xi_1, \xi_2)$ , a parametric assumption on  $H(\xi_1, \xi_2)$  is needed to identify its parameters.

In order to check the performance of the estimator, I generated data based on a fixed value of the parameter vector,  $\theta$ . The estimator was able to recover the parameters, providing evidence that the model's parameters are identified. In the estimation of the model, I fix the discount rate at 0.95.

## 4 Data

### 4.1 The Sample

The sample used in the estimation is constructed from the National Longitudinal Survey of Youth (NLSY), which is a survey of individuals who were between the ages of 14 and 22 when they were first interviewed in 1979. Since then, the respondents have been interviewed annually until 1994 and once every two years after 1994. 19 waves of the NLSY from 1979 to 2000 are used in the analysis.

One of the strengths of the NLSY compared to other longitudinal datasets is the detailed employment

information that it collects. It includes the beginning and end dates of up to five jobs that the respondent has had in a year. Therefore, a relatively more accurate date of transition into the labor market can be established and job tenure can be fully captured. In addition, it includes data on usual hours worked, number of weeks worked, the hourly rate of pay, the three-digit industry and occupation codes, and the reason for separation from the job.

Following Farber (1994), I assume that a worker's labor market career starts when he makes a permanent transition into the labor force. According to this definition, a permanent labor market transition occurs in the beginning of the first 3-year spell of "primarily working," following at least one year in which the worker was "not primarily working". A worker is defined to be primarily working if he has worked at least half of the weeks since the last interview and averaged at least 30 hours per week during the weeks in which he worked. I attempt to mitigate the initial conditions problem by restricting the sample to those who have made a permanent transition into the labor market during the sample period<sup>8</sup>. Therefore, people who have never worked primarily for three consecutive years during the sample period and those who were primarily working in the first year in which they were observed in the dataset are excluded from the sample. Such a sample restriction also allows me to focus on the labor market experiences of those workers who have formed a long-term attachment to the labor market. Furthermore, I exclude workers who have started their labor market careers before the age of 16.

Only jobs that start after the worker's permanent labor market transition are included in the sample. In addition, jobs without valid data on wage, occupation and reason for separation are excluded. The occupation data are collected for jobs that last for at least 9 weeks; as a result, jobs with shorter tenure are excluded from the sample.

The discrete period of analysis is an interview year, which spans the time between two consecutive interviews and is approximately equal to one calendar year. The sample includes the first 8 years of a worker's labor market history beginning with his permanent transition into the labor force. I use the following rules to construct the variables used in the analysis.

*Occupation:* I categorize occupations into blue collar or white collar, based on one-digit census codes<sup>9</sup>.

---

<sup>8</sup>An alternative strategy to handle the initial conditions problem is to approximate the conditional probability of the initial value using a separate probit function as proposed by Heckman (1981). Wooldridge (2005) has recently proposed another alternative based on specifying the distribution of unobserved heterogeneity conditional on the initial value and exogenous variables.

<sup>9</sup>White collar occupations are 1) professional, technical, and kindred; 2) managers, officials, and proprietors; 3) sales workers; 4) farmers and farm managers; and 5) clerical and kindred. Blue collar occupations are 1) craftsmen, foremen, and kindred; 2) operatives and kindred; 3) laborers, except farm; 4) farm laborers and foremen; and 5) service workers.

The worker's occupation in a given year is the one in which he has worked the most number of hours. For multiple job holders, jobs that are not in the worker's assigned occupation are excluded from the analysis. For example, if a multiple job holder has worked the most number of hours in the white collar occupation in a given year, any blue collar job that he may have had during that year is excluded. Formulating the two types of primary sectors as white collar and blue collar occupations in the estimation carries the underlying assumption that workers use different types of skill sets in different occupations and that jobs within each occupation are homogeneous with respect to their output prices and monitoring rates. Increasing the number of occupational categories would probably capture skill heterogeneity and firm heterogeneity more accurately; however, the computational burden would also increase.

*Wages:* I use hourly wages in 2000 dollars. For multiple job holders, I calculate the weighted average of wages in the worker's occupation by multiplying each wage by the number of hours worked in each job and dividing the total earnings by the total number of hours worked in occupation.

*Dismissals:* If the worker has reported a firing or lay-off during an interview year, he is considered to have experienced a dismissal at the end of that year. I consider layoffs as dismissals in this analysis because there might be arbitrariness involved in a worker's self-reporting. He may choose to report a firing as a lay-off due to the stigma that might be associated with a firing. Furthermore, the firm might choose to dismiss its least productive workers during a lay-off instead of firing them since firing may increase the probability that the worker will be disgruntled and possibly challenge the firm's decision. Although it is quite difficult to accurately measure dismissals, this is arguably the best definition given the available data. If the worker reports a quit and takes on another job in the same occupation following the separation, I treat the tenure in that occupation as uninterrupted. If the worker becomes unemployed during the sample period, I only consider his experience until he enters unemployment.

Furthermore, about 20 percent of the individuals in the sample report switching to jobs in a different occupation before their first dismissal. The theoretical model does not include inter-occupational moves while in the primary sector; therefore, I make the following assumptions in mapping the data to the model. If the occupation switch occurs while the worker is in the primary sector, i.e. before he experiences his first dismissal, I include only his labor market experience until he switches occupations in the analysis sample.

## 4.2 Descriptive Statistics

After I impose the criteria described in the previous subsection, the resulting sample includes 5391 individuals. 2059 (38%) of these individuals are employed in the white collar occupation in the first year of their labor market careers, and they remain in the white collar sector until their first dismissal or occupation switch. The remaining 3332 (62%) are employed in the blue collar sector.

Table 1 shows the observable sample heterogeneity in terms of age and education at the start of the labor market career. These statistics suggest that blue collar workers start their long-term labor market careers earlier than white collar workers. 45 percent of blue collar workers make a permanent transition into the labor market between the ages of 16-18 while only 32 percent of white collar workers start their labor market careers before they turn 19. A related statistic is the educational composition of the labor force in the two occupations. Approximately 83 percent of the blue collar workers have at most a high school degree at the beginning of their permanent labor market careers. On the other hand, white collar employees are relatively more educated when they start their careers, with 41 percent having more than a high school degree.

Dismissal rates and average wages in each occupation conditional on the sample period are given in Table 2. Dismissal rates in both occupations follow a general downward trend over tenure in occupation, except for the second period of employment among blue collar workers. This pattern is consistent with the model's prediction that dismissal rates fall in the primary sector because over time a smaller proportion of people remaining in the cohort chooses to shirk. The blue collar sector has higher dismissal rates over the first eight years in occupation. The average white collar dismissal rate during this period is about 5 percent while the average blue collar dismissal rate is roughly 10 percent. Table 2 also shows that the wage in each occupation is monotonically increasing over tenure. Blue collar wages are lower than white collar wages at all tenure levels.

Table 3 presents ordinary least squares (OLS) regression results that show the effect of dismissals on subsequent wages. These results support several assumptions of the theoretical model. First, the statistically insignificant coefficient on the "dismissed in  $t-1$ " dummy variable suggests that dismissals in the secondary sector do not significantly affect wages. This result is consistent with the model's assumption that workers, who are dismissed in the secondary sector, immediately find another job with the same wage. On the other hand, dismissals in the primary sector seem to have a negative effect on wages although the effect is statistically insignificant among white collar workers (the sum of the coefficients on "dismissed in  $t - 1$ "

and "(dismissed in  $t - 1$ )(never dismissed until  $t - 1$ )". This result supports the model's assumption that workers dismissed in the primary sector find work in the secondary sector where they earn lower wages. Finally, the regression results show that there is a significant difference between primary and secondary sector wages of workers who were not dismissed in the previous period. In particular, the positive and statistically significant coefficient on "never dismissed until  $t - 1$ " suggests that primary sector workers earn higher wages than secondary sector workers conditional on tenure in occupation. (A worker, who has never been dismissed until period  $t - 1$  and is not dismissed in  $t - 1$ , is in the primary sector in period  $t$ .)

## 5 Results

### 5.1 Parameter Estimates

Table 4 presents the maximum likelihood estimates of the model's structural parameters and their associated asymptotic standard errors. The estimates of the output prices are similar in the two occupations: \$14.44 in the white collar occupation and \$14.77 in the blue collar occupation. The output price can be interpreted as the upper limit on the wage sequence since it equals the wage that would be earned if there were no shirking. The secondary sector wage is \$11.47 for workers who have been dismissed from a white collar job and \$10.65 for those who have been dismissed from a blue collar job. This result suggests that blue collar workers, who have been dismissed, earn less on their next jobs than the dismissed white collar workers. Furthermore, monitoring rate in the white collar sector is estimated to be 32 percent while the estimate for the blue collar sector is 25 percent. According to these findings, white collar workers face a higher probability of being monitored than blue collar workers.

The population distribution of the white collar productive inefficiency index ( $\log \xi_1$ ) has an estimated mean of 2.02 and variance of 1.2, while the estimated mean and variance of the blue collar productive inefficiency index ( $\log \xi_2$ ) are 2.5 and 0.9, respectively. Therefore, the population distribution of blue collar productive inefficiency has a larger mass at high inefficiency levels and is more concentrated around its mean compared to the population distribution of white collar inefficiency index. In addition, the covariance between  $\log \xi_1$  and  $\log \xi_2$  is estimated to be 0.89, which translates into a correlation coefficient of 0.86 between the two random variables. This result supports the hypothesis that the abilities relevant in blue collar and white collar occupations are highly correlated, indicating that the skills needed in the two occupations might be similar to each other.

Figures 5 and 6 compare the population marginal distributions with the post-selection marginal distribution in each occupation. The parameters of the post-selection distributions of worker types in each occupation are estimated by using maximum likelihood of fitting the post-selection distributions to lognormal distributions. As shown in Figure 5, the distribution of worker types in the white collar occupation is very close to the population distribution. The estimates of the mean and the variance of  $\log \xi_1$  among white collar workers are 1.96 and 1.2, resulting in a post-selection distribution of  $\log \xi_1$  in the white collar occupation that is very close to the population distribution of  $\log \xi_1$ . The distribution of  $\log \xi_1$  in the white collar sector has slightly higher mass in the lower tail and a slightly smaller mass in the upper tail, compared to the population distribution. Therefore, workers with low values of  $\xi_1$  have a higher tendency to become white collar workers while those with high values of  $\xi_1$  have a slightly lower probability of choosing the white collar occupation. These results can be interpreted as evidence of a small degree of positive selection into white collar occupation.

The pattern of selection into the blue collar occupation is more complicated as shown in Figure 6. The estimates for the mean and variance of the distribution of  $\log \xi_2$  among blue collar workers are 2.43 and 0.42, respectively. As shown in Figure 6, the marginal distribution of  $\log \xi_2$  among blue collar workers lies below the population distribution at low levels of  $\log \xi_2$ , indicating that workers with low productive inefficiency in the blue collar sector tend to choose the white collar sector. The post-selection distribution also has a smaller mass at high levels of  $\log \xi_2$  than the population distribution. Therefore, workers with high levels of productive inefficiency in the blue collar sector also have a higher probability of choosing the white collar occupation. The tendency of both high inefficiency and low inefficiency workers to avoid the blue collar sector would have opposite effects on the expected worker productivity in that sector. The forces behind this mixed pattern of selection and its impact on equilibrium wages and dismissal rates in the two occupations are discussed in the following subsections.

Dismissal rates in both occupations fall over tenure as dictated by the model (Table 5). Blue collar workers face higher dismissal rates than white collar workers during the first eight years in the occupation. The high rate of dismissals in the blue collar sector is driven by the result that relative to the white collar sector, the blue collar workforce is composed of a higher proportion of workers who are likely to shirk. In fact, the estimated average productivity in the blue collar sector over the sample period is 0.69, indicating that almost 31 percent of blue collar workers are shirking on average during the sample period. On the other hand, the average white collar productivity over eight periods is 0.86. Although the dismissal rates are

lower in the white collar occupation, they fall faster than blue collar dismissal rates, especially in the first few years in the occupation. This result suggests that shirkers are detected and dismissed faster due to the higher monitoring rate in the white collar occupation. Due to the lower monitoring rate in the blue collar occupation, it takes firms longer to eliminate shirkers in the blue collar occupation.

Table 5 also shows that the predicted wages in the white collar occupation are higher than those in the blue collar occupation. The white collar occupation exhibits higher wages in spite of a lower output price primarily because of the high level of productivity in that occupation. By the 8th period, 97 percent of the white collar workforce chooses to exert effort, and consequently, their wage approaches the upper limit of \$14.44. Furthermore, the theoretical model implies that wage growth occurs as inefficient workers are eliminated from the primary sector so that the remaining workers in the cohort become more efficient on average. Therefore, it is not surprising to see that the blue collar sector, which has higher dismissal rates, also has a higher wage growth. Blue collar wages rise at a higher rate than white collar wages because there are higher productivity gains in the blue collar occupation due to the systematic dismissal of shirkers.

## **5.2 Model's Fit to the Data**

Figures 7-10 graphically depict how well the model fits the observed data on dismissal rates and hourly wages. The results depicted in Figure 7 reveal that the model fits the white collar dismissal rates fairly well. The ratio of estimated to actual dismissal rates is very close to 1 in several periods, and in other periods, the estimated white collar dismissal rate is slightly lower than the actual. The blue collar dismissal rates estimated by the model are lower than the observed dismissal rates, except in the first year of employment, but estimated blue collar dismissal rates exhibit a similar rate of decrease as the actual dismissal rates (Figure 8). These results seem to suggest that dismissals for cause as modeled in this paper can explain most of the dismissals observed in the white collar occupation. However, relative to the white collar occupation, a greater proportion of dismissals in the blue collar occupation may be due to reasons not included in the model, such as exogenous demand shocks.

The estimated wages are slightly higher than the actual in both occupations (Figures 9 and 10). The overestimation of wages together with the underestimation of dismissal rates demonstrates the mechanics of the model as it attempts to fit the data. Although the estimated dismissal rates are lower than the actual dismissal rates, estimated wage growth is very close to the actual in both occupations. In this model, wage growth is generated by the systematic dismissal of unproductive workers. Higher dismissal rates lead to

higher wage growth as the cohort of workers becomes increasingly more productive. This close relationship brings about a difficulty in the model in simultaneously fitting the dismissal and the wage data. If the model were to generate higher dismissal rates in an attempt to yield a better fit to the dismissal data, then the model's fit to the wage growth would suffer as the resulting wage-tenure profiles would be much steeper than the ones observed in the data.

When the model is estimated separately for two education categories, model's fit to the dismissal data improves (Figures 11 and 12). The lower education category includes individuals with at most a high school degree while the higher education category includes those with some college education or more.

### **5.3 The Implications of Self-Selection on Wages and Dismissal Rates**

In this section, I compare the wages and dismissal rates that are predicted by the self-selection model with those that would result if workers were randomly assigned to each occupation. In the latter case, the equilibrium wages and dismissal rates are calculated by setting the marginal distribution parameters in each occupation to the population parameter estimates.

Results presented in Table 6 and Table 7 indicate that the random assignment of workers into occupations leads to lower wages and higher dismissal rates in both occupations compared to an economy with occupational selection. These results are consistent with the earlier finding of positive selection into the white collar occupation. Self-selection enhances the skill distribution in the white collar occupation in such a way that the expected productivity in that occupation rises. The previous subsection also documents the evidence of negative selection into the blue collar occupation at low levels of  $\xi_2$  and positive selection at high levels of  $\xi_2$ . The results in Tables 6 and 7 reveal that the net effect of this type of sorting into blue collar occupation is higher worker productivity in that occupation compared to the case of random assignment.

### **5.4 The Implications of Moral Hazard on the Pattern of Self-Selection**

Next, I study how workers would allocate between occupations when effort is perfectly observed by primary sector firms, so there is no moral hazard. In that case, all workers exert effort, wage equals the output price, and the utility flow to worker  $i$  in primary sector firm  $j$  is  $\rho_j - \xi_{ij}$ . Then, occupational choice is determined by the following: worker  $i$  chooses primary sector firm 1 if  $\rho_1 - \xi_{i1} > \rho_2 - \xi_{i2}$ , provided that  $\rho_1 - \xi_{i1} > 0$  and/or  $\rho_2 - \xi_{i2} > 0$ . The worker chooses the secondary sector if  $\rho_1 - \xi_{i1} < 0$  and  $\rho_2 - \xi_{i2} < 0$ . In

order to observe how workers select into occupations in this case, I simulate data using the estimates for the population distribution parameters and output prices given in Table 4<sup>10</sup>.

Figures 13 and 14 show the marginal distribution of worker types in each occupation when there is no moral hazard. Comparison of these figures to Figures 5 and 6 respectively reveal that the existence of moral hazard significantly diminishes the degree of positive selection into both occupations. This can be explained by the behavior of workers who have high values of  $\xi_1$  and  $\xi_2$  and thus are likely to shirk in both occupations. A comparison of Figures 13 and 14 with Figures 5 and 6 reveal that when the labor market moves from the no-moral-hazard to the moral-hazard case, these workers become more likely to choose the white collar occupation, increasing the mass at the right tail of the log  $\xi_1$  distribution in the white collar occupation and decreasing the mass at the right tail of the log  $\xi_2$  distribution in the blue collar occupation. Workers who would potentially shirk in either occupation seem to choose the white collar occupation because white collar wages are substantially higher than the blue collar wages, and the white collar monitoring rate is not high enough to discourage shirkers from choosing this occupation.

In addition, differences in the selection patterns in labor markets with and without moral hazard can also be explained by the behavior of workers with high  $\xi_1$  and low  $\xi_2$ . These workers are likely to shirk in the white collar occupation and exert effort in the blue collar occupation. A marginal worker, who has high  $\xi_1$  and low  $\xi_2$  and is indifferent between the two occupations when there is no moral hazard, might choose the white collar occupation under moral hazard if the monitoring rate in the white collar occupation is not sufficiently high to deter him from entering this occupation.

## 5.5 The Impact of Changes in the Monitoring Rate and the Output Price

Table 8 shows that a 50 percent increase in the blue collar monitoring rate leads to higher wages and lower dismissal rates in the blue collar occupation. A higher blue collar monitoring rate makes this sector less attractive for workers who have high productive inefficiency and are likely to shirk in a blue collar job. Among workers who choose the blue collar occupation, a higher monitoring rate provides a stronger incentive to supply effort, generating a corresponding increase in average productivity in the blue collar sector. According to the results in Table 8, such an increase in the blue collar monitoring rate also has an adverse

---

<sup>10</sup>Using the same parameter values as in Table 4 allows one to isolate the effects of the moral hazard assumption on workers' occupational choice decisions. Since the parameter values are identical in the moral-hazard and no-moral hazard cases, the difference in selection patterns in the two cases can be attributed to the existence of the moral hazard problem. In that sense, this exercise can be considered a comparative static exercise that is intended to compare the outcomes estimated by the model to the outcomes under the threshold case in which worker effort is perfectly observed by the firm.

effect on the composition of the white collar workforce, lowering the wages and increasing the dismissal rates in the white collar occupation.

According to the findings presented in Table 9, a 10 percent increase in the white collar output price leads to higher wages and lower dismissal rates in the white collar occupation. A higher output price has a direct positive effect on white collar wages since it raises the value of workers' marginal products. Furthermore, higher white collar wages provide work incentives, decrease shirking and thus increase expected productivity in the white collar occupation. This increase in expected productivity further increases wages in the white collar occupation, resulting in an average of 11 percent increase in white collar wages over 10 periods. A higher white collar output price also leads to higher wages and lower dismissal rates in the blue collar occupation, indicating that it has a positive effect on the composition of the blue collar workforce.

## **5.6 The Importance of Worker Heterogeneity**

As in any model of self-selection and occupational choice, the population distribution of worker types in this model plays an important role in how workers are allocated to occupations. The simulations presented in the Model section, for instance, illustrate how the correlation between  $\xi_1$  and  $\xi_2$  in the population can lead to different self-selection patterns into occupations. However, what makes this model different from other self-selection models is that it allows workers to make effort decisions simultaneously as their occupational choice decisions. As a result, the distribution of worker types impacts not only workers' occupational choice but also their productivity decisions in their chosen occupations.

The distribution of worker types in the population can be affected by policy decisions to a significant degree. Scholarships, grants, and training opportunities that are made available to students or practitioners of a given field are often driven by policy objectives to increase the population's aptitude in that field. To the extent that education and training increase one's ability and decreases his disutility of effort in performing occupation-specific tasks, the policy decision to enhance a certain set of skills in the population generates a work force with relatively higher ability and lower disutility of effort in the occupation for which that skill set is most relevant. For example, scholarships for students majoring in science and math may attract more students to these fields, enhancing their skills in science and math and ultimately decreasing their disutility of effort in occupations where these skills are prominently used. On the other hand, providing more vocational classes such as carpentry and mechanical shop and opening an automotive school are among the policy decisions that would positively affect the skills that are significantly used in many blue collar occupations.

In this section, I consider two comparative static exercises in which the population means of  $\xi_1$  and  $\xi_2$  are decreased by 20 percent. These decreases correspond to lower mean disutility of effort in white collar and blue collar jobs, respectively, and can be brought about by policies geared towards enhanced education and training opportunities related to white collar and blue collar professions. The results presented in Table 10 show that a decrease in the population mean of  $\xi_1$  leads to a decrease in dismissal rates and an increase in wages of white collar workers. In particular, a 20 percent decrease in mean  $\xi_1$  increases white collar wages by 18 percent and decreases the white collar dismissal rates by 34 percent in the first period. This result suggests that as the set of skills that are most relevant in white collar jobs are enhanced in the population, workers have lower disutility of effort for working in these jobs, and they are less likely to shirk. The decrease in mean population  $\xi_1$ , however, has the opposite effect in the economic outcomes of the blue collar occupation. As shown in Table 10, wages fall and dismissal rates rise among blue collar workers as a result of the decrease in mean  $\xi_1$ . These results indicate a certain degree of adverse selection into blue collar occupation as higher wages and lower disutility of effort in the white collar occupation attract workers with high skills in both sectors to the white collar sector.

Table 10 also shows the impact of a decrease in the population mean of  $\xi_2$ . A 20 percent decrease in mean  $\xi_2$  leads to an average of 22 percent decrease in the dismissal rates and 10 percent increase in the wages among blue collar workers over the first eight years. These changes are generated by increased productivity among blue collar workers as their disutility of effort in blue collar jobs falls. In contrast to the adverse selection effects of a lower mean  $\xi_1$  on the blue collar occupation, a lower mean  $\xi_2$  has a positive impact on the economic outcomes of the white collar occupation. As the mean  $\xi_2$  decreases by 20 percent, the white collar dismissal rate decreases by 37 percent, and the white collar wage increases by 19 percent in the first period with similar changes in subsequent periods. This result suggests that lower  $\xi_2$  in the population generates not only higher productivity in the blue collar occupation, but it also leads to positive selection into the white collar occupation. As disutility of effort falls and wages rise in the blue collar occupation, workers with low ability in the white collar occupation become more likely to choose blue collar occupation, leading to higher productivity in the white collar occupation.

These exercises reveal that in a labor market with moral hazard, investments in human capital can affect worker productivity, wages and dismissals by impacting workers' effort decisions as well as their occupational choices. The results presented here suggest that higher human capital investments in skills relevant to the white collar occupation increase the productivity of white collar workers while they lead to adverse

selection into the blue collar occupation. Similar investments in the blue collar skills, however, seem to have a positive impact on the productivity of both white collar and blue collar workers.

## 6 Conclusion

In this paper, I present and structurally estimate a model of occupational self-selection in a labor market characterized by moral hazard. The model demonstrates that in the presence of moral hazard, workers' effort decisions serve as an additional mechanism that drives self-selection, confounding the relationship between the population distribution of workers and the economic outcomes in each occupation. When workers make simultaneous occupation and effort decisions, worker types affect the occupational allocation of workers not only through their impact on workers' comparative advantage but also through their probability of shirking in each occupation.

The estimation results suggest that self-selection of workers increases expected worker productivity in both blue collar and white collar occupations. In particular, workers' occupational sorting leads to higher wages and lower dismissal rates in both occupations compared to an economy in which workers are randomly assigned to each occupation. Findings also indicate that the difference in dismissal rates between the two occupations is driven by higher expected productivity in the white collar occupation. Higher wages and higher monitoring rate in the white collar occupation provide stronger incentives for effort than those in the blue collar occupation. Furthermore, the positive effects of self-selection in terms of higher expected productivity, higher wages and lower dismissal rates diminish as the labor market becomes increasingly characterized by moral hazard. The potential for shirking that exists under moral hazard makes some workers more likely to choose an occupation in which they have high disutility of effort.

Results also indicate that a higher monitoring rate in the blue collar occupation leads to higher wages and lower dismissal rates in the blue collar occupation and lower wages and higher dismissal rates in the white collar occupation. On the other hand, a higher white collar output price leads to higher wages and lower dismissal rates in both occupations. These analyses demonstrate that an exogenous shock in one occupation can have positive or adverse effects on the other occupation's economic outcomes depending on its impact on the pattern of workers' occupational allocation.

Finally, I study the effects of a change in the distribution of worker types in the population. I find that lowering the mean disutility of effort for white collar jobs in the population leads to higher wages and lower

dismissal rates in the white collar occupation but lower wages and higher dismissal rates in the blue collar occupation. On the other hand, a similar decrease in the mean disutility of effort for blue collar jobs brings about higher wages and lower dismissal in both occupations. These results suggest that policies aimed at enhancing the set of skills that are most relevant to blue collar jobs may have a positive effect on the productivity in both white collar and blue collar occupations.

The model presented in this paper can be extended in several ways. First, it can be extended to incorporate other sources of wage growth, such as learning-by-doing or human capital investment, in order to enhance the model's ability to explain the rate of wage growth over tenure in occupation. This extension is considered in Demiralp (2007). Second, the model presented here abstracts from firms' decisions on monitoring intensity by taking the monitoring rate as exogenous. It can be extended to explain the firms' monitoring decision by specifying the monitoring technology and monitoring costs, thus allowing one to study how the firm changes its monitoring rate in response to other variables, such as worker productivity. Finally, the model focuses only on dismissals that are due to shirking or malfeasance. The model can also be formulated to include other causes of dismissals, such as exogenous demand fluctuations and low worker-firm match value. This addition would allow one to study the relative importance of different causes of dismissals in the labor market.

## 7 Appendix

### 7.1 Computation of the Equilibrium Wage Sequence

The computation of the equilibrium wage contract consists of two sets of fixed point iterations, one nested within the other. Let  $(\nu_j, \omega_j)$  be the set of parameters that characterize the post-selection marginal distribution of  $\xi_j$  in occupation  $j$  ( $= 1, 2$ ). Then,

$$\begin{bmatrix} F_{1,t=1}(\nu_1, \omega_1) \\ F_{2,t=1}(\nu_2, \omega_2) \end{bmatrix} = \begin{bmatrix} \int_0^\infty \frac{h_1(\xi_1)}{H_1(\xi^*(w_1(F_{1,t=1}(\nu_1, \omega_1)), w_2(F_{2,t=1}(\nu_2, \omega_2))))} dH_2(\xi_2) \\ \int_0^\infty \frac{h_2(\xi_2)}{H_2(\xi^*(w_1(F_{1,t=1}(\nu_1, \omega_1)), w_2(F_{2,t=1}(\nu_2, \omega_2))))} dH_1(\xi_1) \end{bmatrix} \quad (14)$$

Embedded in this fixed point algorithm is a second fixed point algorithm that computes the wage sequence in each occupation, *conditional on*  $(\nu_1, \omega_1)$  and  $(\nu_2, \omega_2)$ . The wage sequence in primary sector firm  $j$ ,

conditional on  $(\nu_j, \omega_j)$ , is the fixed point of the following operator:

$$T(\{w_{jt}\}) = \begin{bmatrix} \rho_j F_{j,t=1}(\xi_{j,t=1}) \\ \rho_j F_{j,t=2}(\xi_{j,t=2}) \\ \vdots \\ \rho_j F_{j,t=\tau}(\xi_{j,t=\tau}) \\ \vdots \end{bmatrix}, \quad j = 1, 2 \quad (15)$$

The finite approximation of this infinite horizon problem is given by the following mapping:

$$T_S(\{w_{jt}\}) = \begin{bmatrix} \rho_j F_{j,t=1}(\xi_{j,t=1}) \\ \vdots \\ \rho_j F_{j,t=S}(\xi_{j,t=S}) \\ \rho_j \\ \rho_j \\ \vdots \end{bmatrix}, \quad j = 1, 2 \quad (16)$$

The fixed point of  $T_S(\{w_{jt}\})$  gives the wage sequence in firm  $j$  conditional on the marginal distribution of  $\xi_j$  in firm  $j$ . Every iteration in solving the fixed point problem in Equation 14 involves the computation of the conditional equilibrium wage sequence; therefore, the algorithm to compute the fixed points of  $T_S(\{w_{jt}\})$  is nested within the algorithm to compute  $(\nu_j, \omega_j)$ .

The procedure to calculate the equilibrium wages and marginal distribution parameters in each firm is explained below. The execution of the following procedure relies on parametric assumptions regarding both the marginal distribution of worker types in the population and the marginal distribution of types in each occupation.  $H(\xi_1, \xi_2)$ , which describes worker heterogeneity in the population, is assumed to be a bivariate lognormal distribution.  $F_{j,t=1}(\xi_j)$  is the marginal distribution of  $\xi_j$  in firm  $j$  in the beginning of period 1, and it is also assumed to be a lognormal distribution. The algorithm to compute the equilibrium wage sequence in each firm is as follows:

**Step 1:** Choose positive constants  $\kappa$  and  $\psi$ , and set  $S$  to a large positive integer.

**Step 2:** Randomly draw  $N$  observations from the bivariate population distribution,  $H(\xi_1, \xi_2)$ .

**Step 3:** Choose initial values for the wage sequence and denote them  $\{w_{1t}\}^0$  and  $\{w_{2t}\}^0$ .

**Step 4:** Using the operator  $T(\{w_{1t}\})$ , iterate until Equation 17 is satisfied.

$$d_{\infty} \left( \{w_{1t}\}^{K+1}, \{w_{1t}\}^K \right) \leq \kappa \quad (17)$$

The value of the wage sequence at the final iteration is  $\{w_{1t}\}^*$ . Do the same for firm 2 and compute  $\{w_{2t}\}^*$ .

**Step 5:** Using  $\{w_{1t}\}^*$ ,  $\{w_{2t}\}^*$  and the parameters of the model, calculate  $V_{i,1,t=1}$  and  $V_{i,2,t=1}$  according to Equation 2, and determine which of the  $N$   $(\xi_1, \xi_2)$  pairs choose firm 1 and which ones choose firm 2. Recall that a worker chooses firm 1 if  $V_{i,1,t=1} > V_{i,2,t=1}$ .

**Step 6:** Compute  $\nu_1, \omega_1, \nu_2$ , and  $\omega_2$  by fitting the post-selection marginal distributions that are generated in Step 5 to lognormal distributions using maximum likelihood.

**Step 7:** Denote parameters estimated in Step 6,  $(\nu_1, \omega_1)^0$  and  $(\nu_2, \omega_2)^0$ .

**Step 8:** Repeat steps 4-6. Denote the parameter estimates  $(\nu_1, \omega_1)^*$ . Compute  $D_1 = d_{\infty} \left( (\nu_1, \omega_1)^*, (\nu_1, \omega_1)^0 \right)$  for firm 1. Similarly, compute  $D_2$  for firm 2.

**Step 9:** Iterate (repeat steps 4-6) by setting  $(\nu_1, \omega_1)^0 = (\nu_1, \omega_1)^*$  and  $(\nu_2, \omega_2)^0 = (\nu_2, \omega_2)^*$  until  $D_1 \leq \psi$  and  $D_2 \leq \psi$ . If  $D_1 \leq \psi$  and  $D_2 \leq \psi$ , the approximate equilibrium wage sequence in firm  $j$  is  $\{w_{jt}\}^*$  and the parameter estimates of the distribution of  $\xi_j$  in firm  $j$  is  $(\nu_j, \omega_j)^*$  for  $j = 1, 2$ .

## 7.2 The Likelihood Function

Let  $\theta$  be the set of the model's parameters. Then, the likelihood contribution of sample member  $i$  working in sector  $j$  ( $j \neq k$ ) is given by

$$L_i = \Pr(V_j > V_k, \{\ln w_{ijt}\}_{t=1}^T, \{d_{ijt}\}_{t=1}^T; \theta) \quad (18)$$

or equivalently

$$L_i = \Pr(V_j > V_k, \{\ln w_{ijt}\}_{t=1}^T | \{d_{ijt}\}_{t=1}^T; \theta) \cdot \Pr(\{d_{ijt}\}; \theta) \quad (19)$$

where  $V_j$  is the value of working in occupation  $j$ ,  $\{w_{ijt}\}_{t=1}^T$  is the worker's reported wage sequence, and  $\{d_{ijt}\}_{t=1}^T$  is his dismissal sequence, indicating whether the worker was dismissed or not in occupation  $j$  in period  $t$ . Conditional on  $\{d_{ijt}\}_{t=1}^T$ , worker  $i$ 's productive inefficiency vector  $(\xi_1, \xi_2)$  and the measurement

error in wages ( $\varepsilon_t$ ) are independent. Therefore worker  $i$ 's likelihood contribution can be stated as

$$L_i = \Pr(V_j > V_k | \{d_{ijt}\}_{t=1}^T; \theta) \cdot \Pr(\{\ln w_{ijt}\}_{t=1}^T | \{d_{ijt}\}_{t=1}^T; \theta) \cdot \Pr(\{d_{ijt}\}_{t=1}^T; \theta) \quad (20)$$

The probability that a worker has chosen occupation  $j$ , conditional on having no dismissals over  $T$  periods is given by

$$\begin{aligned} \Pr(V_j > V_k | \sum_{t=1}^T d_{ijt} = 0) &= (1 - \pi_j)^T \iint I(\xi_j^*(\xi_k) \geq \xi_j > \xi_{j,t=T} | \xi_k) \cdot dH(\xi_j, \xi_k) \\ &+ (1 - \pi_j)^{T-1} \iint I(\xi_j^*(\xi_k) \geq \xi_j | \xi_k) I(\xi_{j,t=T} < \xi_j \leq \xi_{j,t=T-1}) \cdot dH(\xi_j, \xi_k) \\ &+ \dots + \iint I(\xi_j^*(\xi_k) \geq \xi_j | \xi_k) * I(\xi_j < \xi_{j,t=1}) dH(\xi_j, \xi_k) \end{aligned} \quad (21)$$

The probability that a worker has chosen occupation  $j$ , conditional on being dismissed in period  $T$  is expressed as

$$\Pr(V_j > V_k | d_{ij,t=T} = 1) = (1 - \pi_j)^{T-1} \pi_j \iint I(\xi_j^*(\xi_k) \geq \xi_j > \xi_{j,t=T} | \xi_k) dH(\xi_j, \xi_k) \quad (22)$$

The bivariate integral in the above equation is numerically evaluated for each individual in the sample using the trapezoid rule.

The probability of a worker's reported wage sequence conditional on having no dismissals over  $T$  periods is

$$\Pr(\{\ln w_{ijt}\}_{t=1}^T | \sum_{t=1}^T d_{ijt} = 0) = \sigma_\varepsilon^{-T} \prod_{t=1}^T \phi\left(\frac{\ln w_{ijt} - \ln w_{ijt}^*}{\sigma_\varepsilon}\right) \quad (23)$$

where  $\phi$  is the pdf of a standard normal variable. The probability of the wage sequence of a worker who has been dismissed at the end of period  $T$  and has spent  $T^s$  periods in the secondary sector is given by

$$\begin{aligned} \Pr(\{\ln w_{ijt}\}_{t=1}^T | d_{ij,t=T} = 1) &= \sigma_\varepsilon^{-T} \prod_{t=1}^T \phi\left(\frac{\ln w_{ijt} - \ln w_{ijt}^*}{\sigma_\varepsilon}\right) \\ &\cdot \prod_{t=T+1}^{T^s} \phi\left(\frac{\ln w_{ijt} - \ln w_{ij}^s}{\sigma_\varepsilon}\right) \end{aligned} \quad (24)$$

Finally, the probability of the dismissal sequence,  $\{d_{ijt}\}_{t=1}^T$  is

$$\begin{aligned} \Pr\left(\sum_{t=1}^T d_{ijt} = 0\right) &= F_j(\xi_{j,t=1}) + (1 - \pi_j) [F_j(\xi_{j,t=2}) - F_j(\xi_{j,t=1})] + \dots \\ &+ (1 - \pi_j)^{T-1} [F_j(\xi_{j,t=T}) - F_j(\xi_{j,t=T-1}) + (1 - \pi_j)^T [1 - F_j(\xi_{j,t=T})]] \end{aligned} \quad (25)$$

and

$$\Pr(d_{ijt} = 1) = \pi_j (1 - \pi_j)^{t-1} [1 - F_j(\xi_{jt})] \quad t = 1, \dots, T - 1 \quad (26)$$

As shown in the equations above, the post-selection marginal distributions of worker types in each occupation,  $F_1(\xi_1)$  and  $F_2(\xi_2)$ , are needed for the construction of the likelihood function. The parameters of these distributions are evaluated numerically by means of simulations because the truncation point in Equation 5,  $\xi^*$ , cannot be solved analytically. The parameters of  $F_1(\xi_1)$  and  $F_2(\xi_2)$  are computed according to the fixed point algorithm that is explained in the previous subsection of the appendix.

## References

- [1] Albrecht, James W. and Susan Vroman, 1998. Nash Equilibrium Efficiency Wage Distributions. *International Economic Review*, 39, 1, 183-203.
- [2] Borjas, George, 1987. Self-Selection and the Earnings of Immigrants. *American Economic Review*, 77, 4, 531-553.
- [3] Demiralp, Berna, 2007. A Model of Occupational Choice with Moral Hazard and Human Capital Accumulation, manuscript.
- [4] Farber, Henry S., 1994. The Analysis of Interfirm Worker Mobility. *Journal of Labor Economics*, 12, 554-593.
- [5] Flinn, Christopher J., 1997. Equilibrium Wage and Dismissal Processes. *Journal of Business and Economic Statistics*, 15, 221-236.
- [6] Gould, Eric D., 2002. Rising wage Inequality, Comparative Advantage, and the Growing Importance of General Skills in the United States. *Journal of Labor Economics*, 20, 1, 105-147.
- [7] Heckman, James, 1981. The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process. In: C. Manski and D. McFadden, (Eds.), *The Structural Analysis of Discrete Data*, MIT Press, Cambridge, MA.
- [8] Heckman, James J., and Guilherme Sedlacek, 1985. Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market. *Journal of Political Economy*, 93, 1077-1125.
- [9] Jovanovic, B., 1979. Job Matching and the Theory of Turnover. *Journal of Political Economy*, 87, 972-990.
- [10] Keane, M., and K. Wolpin, 1997. The Career Decisions of Young Men. *Journal of Political Economy*, 105, 3, 473-522.
- [11] MacLeod, W. B. and D. Parent, 1999. Job Characteristics and the Form of Compensation. *Research in Labor Economics*, 18, 177-242.

- [12] McCall, Brian P., 1990. Occupational Matching: A Test of Sorts. *Journal of Political Economy*, 98, 1, 45-69.
- [13] Miller, Robert A., 1984. Job Matching and Occupational Choice. *Journal of Political Economy*, 92, 6, 1086-1120.
- [14] Moscarini, Giuseppe, 2001. Excess Worker Reallocation. *Review of Economic Studies*, 68, 593-612.
- [15] Neal, Derek, 1999. The Complexity of Job Mobility among Young Men. *Journal of Labor Economics*, 17, 2, 237-261.
- [16] Roy, Andrew D., 1951. Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, 3, 135-146.
- [17] Sattinger, Michael, 1993. Assignment Models of the Distribution of Earnings. *Journal of Economic Literature*, 31, 2, 831-880.
- [18] Sauer, Robert M., 1998. Job Mobility and the Market for Lawyers. *Journal of Political Economy*, 106, 1, 147-171.
- [19] Shapiro, C., and J. Stiglitz, 1984. Equilibrium Unemployment as a Worker Discipline Device. *American Economic Review*, 74, 433-444.
- [20] Willis, Robert, 1986. Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions. In: Ashenfelter, O. and R. Layard, (Eds.), *Handbook of Labor Economics*, North Holland, Amsterdam.
- [21] Wooldridge, Jeffrey M., 2005. Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity. *Journal of Applied Econometrics*, 20, 39-54.

**Table 1: Age and Education Distribution of the Sample**

	<i>White Collar</i>	<i>Blue Collar</i>
<i>Age at the start of labor market career</i>		
16-18	652 (32%)	1488 (45%)
19-21	795 (39%)	1168 (35%)
21-25	384 (19%)	390 (12%)
26-30	146 (7%)	174 (4%)
31-42	82 (4%)	112 (3%)
Total	2059	3332
<i>Education at the start of labor market career</i>		
Less than high school	545 (26%)	1690 (51%)
High school	658 (32%)	1071 (32%)
Some college	619 (30%)	521 (16%)
College	189 (9%)	45 (1%)
More than college	48 (2%)	5 (0.2%)

Column percentages are given in parentheses.

**Table 2: Dismissal Rates and Wages by Period**

Period	Dismissal Rates		Hourly Wage	
	<i>White Collar</i>	<i>Blue Collar</i>	<i>White Collar</i>	<i>Blue Collar</i>
1	9.47%	13.30%	9.16	7.97
3	7.95%	12.55%	10.19	8.80
5	4.32%	8.33%	13.28	9.27
7	1.63%	6.21%	15.64	10.77

**Table 3: OLS Regressions of Wages**

<i>Dependent variable:</i>		
<i>ln(wage)</i>	<i>White Collar</i>	<i>Blue Collar</i>
Dismissed in t-1	0.013 (0.030)	0.041 (0.021)
Never dismissed until t-1	0.049* (0.011)	0.038* (0.010)
(Dismissed in t-1) x (Never dismissed until t-1)	-0.063 (0.039)	-0.086* (0.032)
Tenure in occupation	0.079* (0.002)	0.057* (0.002)
Constant	1.935 (0.009)	1.863 (0.008)

Huber-White standard errors are in parentheses.

Asterisk indicates significance at the 5% level.

**Table 4: Maximum Likelihood Estimates of the Parameters**

<i>Description</i>	<i>Parameter</i>	<i>ML Estimates</i>	<i>Asymptotic Standard Errors</i>
WC output price	$\rho_1$	14.4442	0.1526
WC monitoring rate	$\pi_1$	0.3203	0.0127
WC sec sector wage	$w_1^s$	11.4665	0.0793
mean of $\log(\xi_1)$	$\mu_1$	2.0237	0.0337
var of $\log(\xi_1)$	$\alpha_1$	1.2013	0.0157
BC output price	$\rho_2$	14.7744	0.2449
BC monitoring rate	$\pi_2$	0.2493	0.0069
BC sec sector wage	$w_2^s$	10.6489	0.0537
mean of $\log(\xi_2)$	$\mu_2$	2.4908	0.1027
var of $\log(\xi_2)$	$\alpha_2$	0.8986	0.0433
std dev of $\varepsilon$	$\sigma_\varepsilon$	0.4705	0.0006
cov( $\xi_1, \xi_2$ )	$\alpha_{12}$	0.8902	0.0895

**Table 5: Equilibrium Dismissal Rates and Hourly Wages**

Period	Dismissal Rates		Hourly Wage	
	<i>White Collar</i>	<i>Blue Collar</i>	<i>White Collar</i>	<i>Blue Collar</i>
1	10.85%	13.68%	9.55	6.67
3	6.01%	9.67%	11.74	9.04
5	3.05%	6.36%	13.07	11.01
7	1.48%	3.95%	13.78	12.43

Simulations are performed using 10,000 draws.

**Table 6: Equilibrium Wages under Random Assignment**

Period	White Collar		Blue Collar	
	<i>With Self-Selection</i>	<i>Random Assignment of Workers</i>	<i>With Self-Selection</i>	<i>Random Assignment of Workers</i>
1	9.55	9.28	6.67	6.45
3	11.74	11.55	9.04	8.75
5	13.07	12.96	11.01	10.74
7	13.78	13.72	12.43	12.23

**Table 7: Dismissal Rates under Random Assignment**

Period	White Collar		Blue Collar	
	<i>With Self-Selection</i>	<i>Random Assignment of Workers</i>	<i>With Self-Selection</i>	<i>Random Assignment of Workers</i>
1	10.85%	11.46%	13.68%	14.04%
3	6.01%	6.41%	9.67%	10.15%
5	3.05%	3.28%	6.36%	6.81%
7	1.48%	1.60%	3.95%	4.29%

**Table 8: The Impact of a 50% Increase in Blue Collar Occupation's Monitoring Rate**

<i>Period</i>	% change in hourly wage	% change in dismissal rate
White Collar		
1	-5.99%	5.76%
3	-3.34%	7.33%
5	-1.71%	8.23%
7	-0.83%	8.12%
Blue Collar		
1	61.17%	-27.35%
3	42.83%	-53.60%
5	27.12%	-70.93%
7	16.30%	-81.35%

**Table 9: The Impact of a 10% Increase in White Collar Occupation's Output Price**

<i>Period</i>	% change in hourly wage	% change in dismissal rate
White Collar		
1	13.74%	-11.61%
3	12.02%	-13.73%
5	11.01%	-14.94%
7	10.49%	-16.25%
Blue Collar		
1	1.92%	-4.06%
3	1.14%	-6.40%
5	0.68%	-8.52%
7	0.39%	-9.79%

**Table 10: The Impact of a 20% Decrease in the Population Mean of  $\xi_j$** 

<i>Period</i>	20% decrease in mean( $\xi_1$ )		20% decrease in mean( $\xi_2$ )	
	% change in hourly wage	% change in dismissal rate	% change in hourly wage	% change in dismissal rate
White Collar				
1	17.66%	-34.47%	18.78%	-36.68%
3	8.97%	-38.94%	9.50%	-41.26%
5	4.35%	-41.31%	4.59%	-43.61%
7	2.06%	-42.57%	2.17%	-44.59%
Blue Collar				
1	-29.98%	24.63%	20.74%	-17.03%
3	-22.79%	35.99%	12.97%	-20.48%
5	-16.05%	46.86%	7.88%	-23.11%
7	-10.55%	55.95%	4.67%	-24.81%

Figure 1: Marginal Distributions of  $\xi_1$  in the Population and in Firm 1  
When Correlation Coefficient is 0.83

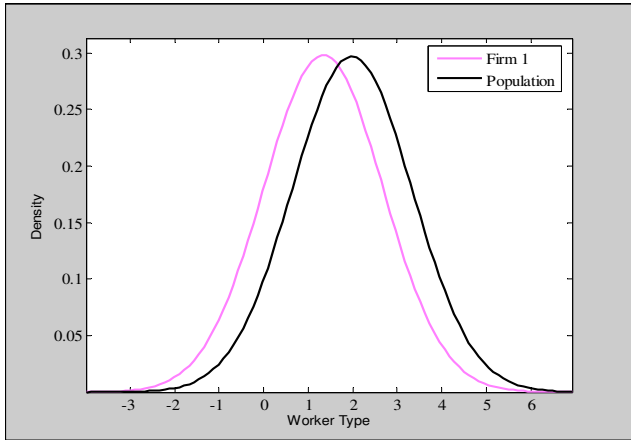


Figure 2: Marginal Distributions of  $\xi_2$  in the Population and in Firm 2  
When Correlation Coefficient is 0.83

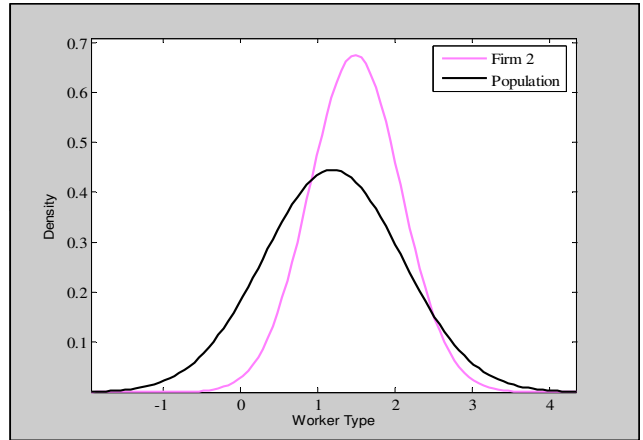


Figure 3: Marginal Distributions of  $\xi_1$  in the Population and in Firm 1  
When Correlation Coefficient is -0.4

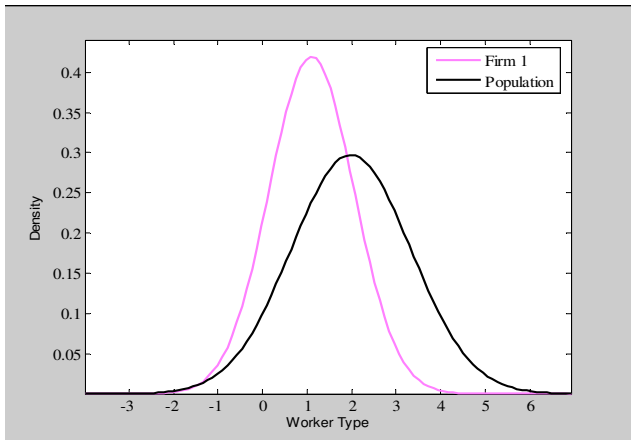


Figure 4: Marginal Distributions of  $\xi_2$  in the Population and in Firm 2  
When Correlation Coefficient is -0.4

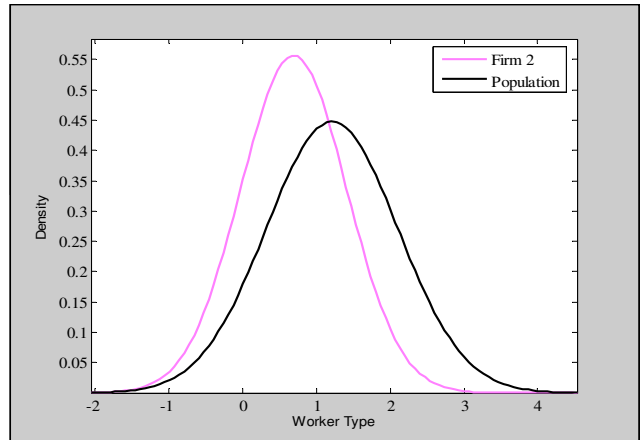


Figure 5: Marginal Distributions of  $\xi_1$  in the Population and Among White Collar Workers

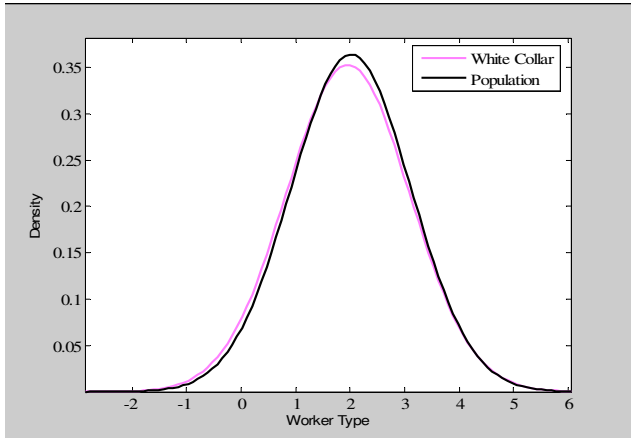
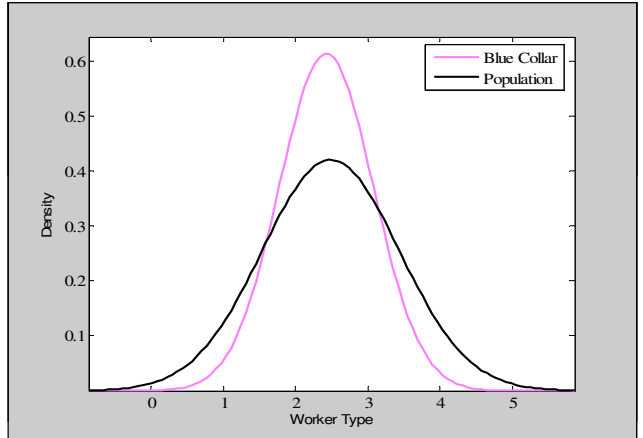
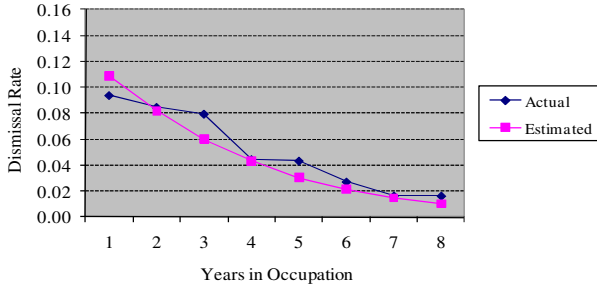


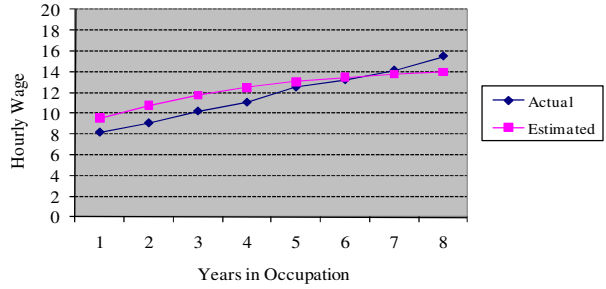
Figure 6: Marginal Distributions of  $\xi_2$  in the Population and Among Blue Collar Worker



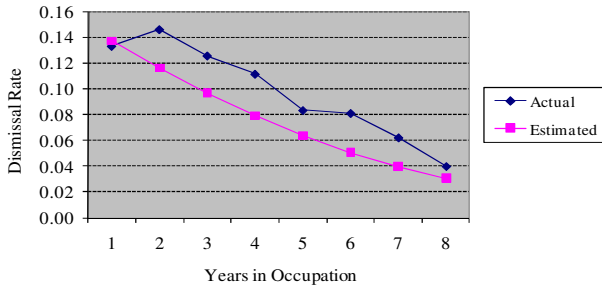
**Figure 7: Dismissal Rates in White Collar Occupation**



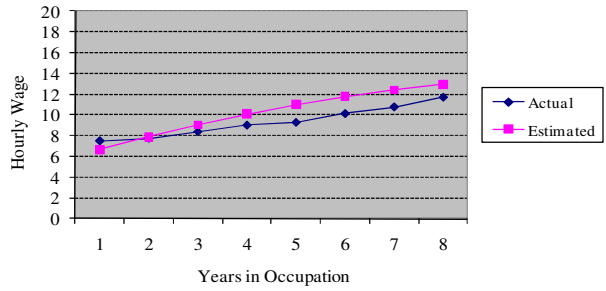
**Figure 9: Wages in White Collar Occupation**



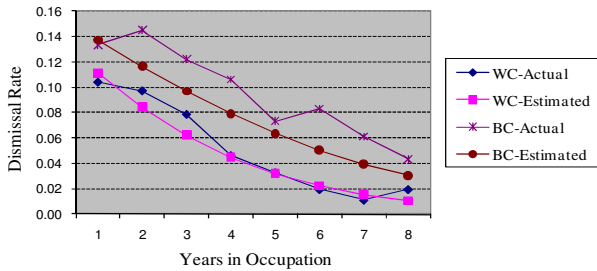
**Figure 8: Dismissal Rates in Blue Collar Occupation**



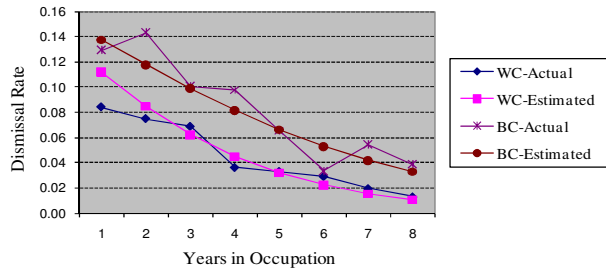
**Figure 10: Wages in Blue Collar Occupation**



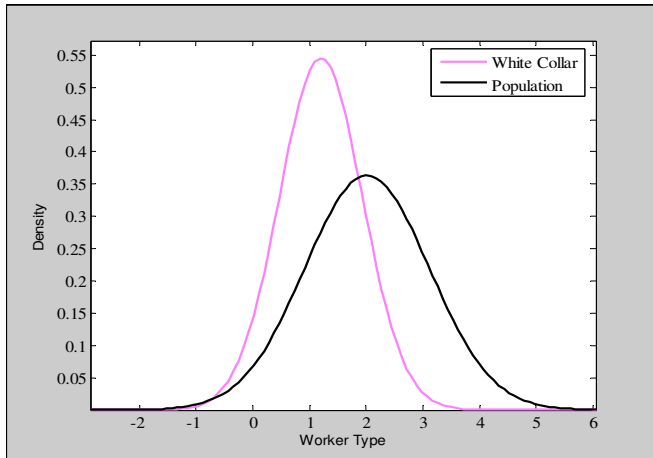
**Figure 11: Dismissal Rates Among the Lower Education Group**



**Figure 12: Dismissal Rates Among the Higher Education Group**



**Figure 13: Marginal Distributions of  $\xi_1$  in the Population and Among White Collar Workers When There is No Moral Hazard**



**Figure 14: Marginal Distributions of  $\xi_2$  in the Population and Among Blue Collar Workers When There is No Moral Hazard**

