Collegio Carlo Alberto

Labor Market Search and Schooling Investment

Christopher Flinn

No. 212

June 2011

Carlo Alberto Notebooks

www.carloalberto.org/working_papers

Labor Market Search and Schooling Investment*

Christopher Flinn
Department of Economics
New York University

Collegio Carlo Alberto
Moncalieri, Italy

First Draft: June 2010
This Draft: June 2011

Abstract

We generalize the standard search, matching, and bargaining framework to allow individuals to acquire productivity-enhancing schooling prior to labor market entry. As is well-known, search frictions and weakness in bargaining position contribute to under-investment from an efficiency perspective. In order to evaluate the sensitivity of schooling investments to "hold up," the model is estimated using Current Population Survey data. We focus on the impact of bargaining power on schooling investment, and find that the effects are large. A brief exploration of the two-sided investment model suggests that something akin to a "Hosios condition" result regarding the socially optimal surplus division rule may be attainable.

JEL Classifications: J24, J3, J64
Keywords: Labor market search; schooling choice; hold-up; Nash bargaining.

*The C.V. Starr Center for Applied Economics at New York University has partially funded this research.. James Mabli greatly aided in the development of the modeling strategy. I am grateful to Cristian Bartolucci for many useful comments and discussions, as well as to Fabien Postel-Vinay for a stimulating (formal) discussion of the paper at the SOLE Meetings in London, June 2010, and to participants in the Milton Friedman Institute Conference in honor of James Heckman held in Chicago, November 2010 and a seminar given at Collegio Carlo Alberto. I am responsible for all errors, omissions, and intpretations.

1 Introduction

A large number of papers, both theoretical and applied, have examined labor market phenomena within the search and matching framework, with some embedded in a simple general equilibrium setting.[1] Virtually all of the empirical work performed using this framework has assumed that individual heterogeneity is exogenously determined at the time of entry into the labor market. Perhaps the most important observable correlate of success in the labor market is schooling attainment. In this paper we extend the standard search and matching framework to allow for endogenous schooling decisions.[2]

We develop a simple model of schooling investment decisions, where higher levels of schooling investments are (generally) associated with better labor market environments. Individuals are differentiated in terms of initial ability, a, and the heterogeneity in this characteristic, along with the structure of the labor market, is what generates equilibrium schooling distributions. As is standard, we utilize axiomatic Nash bargaining to determine the division of the surplus between workers and firms. For simplicity, and due to the nature of the data we utilize, we assume that employed individuals do not receive alternative offers of employment, i.e., there is no on-the-job search.[3]

There is a long-standing literature examining the essence of the hold-up problem and the role contracts play to reduce, or altogether avoid, hold-up (see Malcomson 1997 and Acemoglu 1996 and 1997 for a number of citations to the relevant literature). At the core of the problem is the notion that investments must be made before agents meet and, thus, greater market frictions generally lead to more serious hold-up problems. Acemoglu and Shimer (1999) examine the potential for hold-up problems in frictional markets and investigate the manner in which markets can internalize the resulting externalities. Their focus is on identifying ways in which hold-up and inefficiencies can be mitigated in labor markets characterized by ex-ante worker and firm investments and search frictions and find

[1] A large number of macroeconomic labor applications are cited in Pissarides (2000) and the recent survey by Shimer et al. (2005). In terms of econometric implementations of the model, examples are Flinn and Heckman (1982), Eckstein and Wolpin (2005), Postel-Vinay and Robin (2002), Dey and Flinn (2005), Cahuc et al (2006), and Flinn (2006).

[2] There are a number of ambitious empirical papers which estimate life cycle individual decision rule models of schooling choice and labor market behavior, such as Keane and Wolpin (1997) and Sullivan (2010). This approach has been extended to allow for the endogenous determination of rental rates for various types of human capital, e.g., Heckman et al. (1998), Lee (2005), and Lee and Wolpin (2006). These frameworks do not allow investigation of surplus division issues and the hold-up problem since they are based on a competitive labor market assumption. Eckstein and Wolpin (19995) estimate a search and matching model for various demographic groups in order to evaluate the "return to schooling" along a number of dimensions (e.g., contact rates, matching distributions, bargaining power), but do not explicitly consider the schooling choice decision.

[3] Adding on-the-job search alters the details of what constitutes "bargaining power" in the market, but not the fact that a lack of "generalized" bargaining power, which may include the possibility of renegotiation of contracts as in Postel-Vinay and Robin (2002), Dey and Flinn (2005), and Cahuc et al (2006), will negatively impact the individual's incentive to invest in human capital.

that this can be achieved in wage-posting models with directed search.[4]

The generalized Nash bargaining power parameter has a direct impact on the extent of the hold-up problem the worker faces vis-a-vis pre-market schooling investment decisions. While there are a number of estimates of the bargaining power parameter within models of Nash bargaining and matching, the estimates tend to vary significantly with the assumptions made regarding the presence of on-the-job (OTJ) search, and given OTJ search, the nature of the renegotiation process, as well with respect to the data set used in estimation. In their search, matching, and Nash bargaining frameworks, Dey and Flinn (2005), Cahuc et al. (2006), and Flinn and Mabli (2009) found that allowing for OTJ search substantially reduced the estimate of the worker's bargaining power parameter in comparison with the case in which OTJ search was not introduced (e.g., Flinn 2006). To some degree, this is a result of allowing for Bertrand competition. When competition between firms is introduced, substantial wage gains over an employment spell can be generated simply from this phenomenon, even when the individual possesses little or no bargaining power in terms of the bargaining power parameter. Indeed, the (approximately) limiting case of this is that considered by Postel-Vinay and Robin (2002), in which workers possessed no bargaining power whatsoever. While the hold-up problem would seem to be particularly severe in this case, even to the extent that individuals would have no incentive to invest in human capital, this is not the case when Bertrand competition between competing potential employers occurs, which is when the individual can recoup some of the returns to her pre-market investment. Incentives to invest in their model are directly related to the contact rates with other potential employers in the course of an employment spell, most importantly, as well as the other rates of event occurrence (i.e., the offer arrival rate in the unemployed state and the rate of exogenous separation).

As our model structure makes clear, simply estimating separate behavioral models of the labor market for different schooling classes is at a minimum inefficient, and, more seriously, may lead to misinterpretations of labor market structure. For this reason, whenever possible, potentially endogenous individual characteristics acquired before or after entry into the labor market should be incorporated into the structure of the search, matching, and bargaining model. In order to do so in a tractable manner requires stringent assumptions regarding the productivity process, bargaining, etc., as is evident in what follows. Using our simple and reasonably tractable model, we are able to make some preliminary judgements regarding the impact of hold-up on schooling investment. We find that bargaining power has a strong impact on the incentive to invest in schooling.

Given data limitations, we are not able to estimate a version of the model in which firms make investments just as do individuals. However, we do develop such a model and attempt to indicate the manner in which Hosios-type (1990) conditions could be generated in within this framework. That is, the incentives to invest in schooling depend on the

[4] It is well-known that wage-posting models have their requirements of commitment to mitigate the incentives of firms to renegotiate contracts with individual workers.

share of the surplus given to each party, and a social planner with the ability to fix this parameter can do so as to maximize the expected match surplus in the economy. Given the endowment distributions, productivities of investment on each side of the market, etc., we expect that efficiency will require implementing a surplus sharing rule that gives each side an incentive to invest. A future research goal is to estimate such a model, but to have any hope of identification matched employer-employee data will have to be utilized.

The plan of the paper is as follows. In Section 2, we develop a bargaining model in a partial equilibrium framework, with education decisions made prior to entering the labor market. Section 3 extends the basic model to allow schooling submarkets to be characterized by different vectors of primitive parameters, such as contact and dissolution rates. In Section 4 we describe the sample used to estimate the model and discuss identification of model parameters and the estimator used. Section 5 presents model estimates and (empirical) comparative statics exercises. In Section 6 we extend the model to allow for two-sided investment and holdup, where new issues arise when examining efficiency within the context of a Nash bargaining framework. Section 7 concludes.

2 Model with Homogeneous Schooling Markets

2.1 Overview

By now there have been a number of models that have been estimated within the search, matching, and bargaining framework (e.g., Postel-Vinay and Robin (2002), Dey and Flinn (2005), Cahuc et al. (2006), and Flinn (2006)). These models posit random matching between workers and firms, at least within observationally differentiated labor markets. The flow (in continuous time) productivity of a match between worker i and firm j is assumed to be given by

$$y_{ij} = \tilde{a}_i \theta_{ij} \tilde{p}_j, \tag{1}$$

where $\tilde{a}_i$ is individual $i's$ time- and match-invariant productivity, $\tilde{p}_j$ is firm $j's$ time- and match-invariant productivity, and θ_{ij} is a random match component that is assumed to be independently and identically distributed (i.i.d.) over all potential (i, j) matches according to the distribution function G. Analyses that utilize worker-firm matched data (e.g., Postel-Vinay and Robin (2002) and Cahuc et al. (2006)) typically assume that G is degenerate with $\theta_{ij} = 1\ \forall(i, j)$. Analyses that have used only observations from the supply side of the market (e.g., Dey and Flinn (2005) and Flinn (2006)) instead assume that $\tilde{a}_i = 1\ \forall\ i$ and $\tilde{p}_j = 1\ \forall\ j$.

We can think of the specification of flow productivity in (1) terms of the standard linear model, and this is particularly clear when we consider the logarithm of the expression

$$\ln y_{ij} = \ln \tilde{a}_i + \ln \tilde{p}_j + \ln \theta_{ij}.$$

The terms $\ln \tilde{a}_i$ and $\ln \tilde{p}_j$ represent "main effects," in the language of linear models, while $\ln \theta_{ij}$ represents a higher-order interaction effect. One common specification of flow productivity restricts the (logarithmic) model to include only main effects, whereas the other specification restricts the model to include no main effects. There is no reason to think that either restriction is entirely appropriate, so that the estimation of a model that potentially includes both types of contributions to worker-firm output may produce interesting empirical implications and a more general data generating process.

The main contribution of the paper, however, is to broaden the interpretation of the "main" effects, $\tilde{a}_i$ and $\tilde{p}_j$. The nonparametric estimation of the distributions of these is a compelling contribution of the Postel-Vinay and Robin (2002) and the Cahuc et al. (2006) analyses. In this paper, we attempt to extend the standard labor market search framework to include pre-market investments.[5] For empirical tractability, we limit attention to the case in which workers and firms can decide, prior to entering the market, whether to make a costly investment that will improve their (idiosyncratic) productivity by some fixed amount. In the case of workers, we assume that the individual first is able to observe their ability endowment, a. Prior to entering the labor market, the individual can either stop their schooling at the mandatory level or continue on to more advanced competency. A student of type a who stops schooling at the mandatory level enters the labor market with idiosyncratic ability $\tilde{a}_i = a_i h_1$, where we adopt the normalization $h_1 = 1$. If they were to complete an advanced degree program, the individual would enter the market with ability level $\tilde{a}_i = a_i h_2$, with $h_2 > h_1 = 1$. Similar possibilities exist on the firm side of the market, so that a firm with a productivity endowment of p_j can undertake costly investment so as to make its productivity $\tilde{p}_j = p_j k_2$ or can enter the market without undertaking productivity-enhancing investment so that $\tilde{p}_j = p_j k_1 = p_j$, where we have adopted the normalization that $k_2 > k_1 = 1$.

In our analysis we examine the role of labor market characteristics, particularly bargaining power, on the investment decisions of workers, and to a much lesser degree (due to data limitations) those of firms. One of the contributions of the analysis is to demonstrate that when workers and firms are able to invest prior to market entry, the distributions of worker and firm productivities cannot properly be considered as "primitives," that is, these distributions are responsive to changes in labor market parameters and policy interventions.

2.2 No Firm Heterogeneity

Due to data limitations, we are not able to properly consider the general case of two-sided investment, so that the vast majority of the paper will be devoted to the situation in which firms are homogeneous, i.e., $\tilde{p}_j = 1 \; \forall \; j$. In this case, the output at a match is given by

$$y = a h_s \theta,$$

[5] There has been work on the effect of the hold-up problem on pre-marital investments, with a recent contribution being Chiappori et al. (2009). However, most contributions, such as this one, are primarily theoretical in nature. To my knowledge, no labor market search model with bargaining that includes pre-market investments has been estimated.

where θ is i.i.d. with c.d.f. G, h_s is the individual's human capital level, and a is individual ability, which is a permanent draw from the distribution F_a with corresponding density f_a, which has support $(\underline{a}, \bar{a})$, with $0 \leq \underline{a} < \bar{a} < \infty$. (We now drop the individual and firm subscripts - they are redundant since a and h refer to individuals, p and k refer to firms, and θ is a match value.) As mentioned above, we restrict our attention to the case of $S = 2$, where $s = 1$ corresponds to high school and partial college, roughly, and $s = 2$ to college completion.[6]

The analysis, but theoretical and empirical, can be made much more tractable if we make the following set of assumptions.

1. All parameters describing the labor market are independent of schooling status with the exception of h_s. (This can easily be weakened, which is done in the next section.)

2. The flow value of unemployment to a type a individual with schooling level s is given by
$$b(a, s) = b_0 a h_s.$$

This last assumption is similar to that made in Postel-Vinay and Robin (2002) and in Bartolucci (2009).

Since we use data from the Current Population Survey, and thus the information only consists of a point sample of the labor market process, we assume no on-the-job (OTJ) search. We begin by considering the case in which, across schooling "submarkets," all job search environments are identical (i.e., they have identical parameters $\alpha_1 = \alpha_2$, $\eta_1 = \eta_2$, etc.). In this case, the value of search to an individual of type (a, h_s) can be summarized solely in terms of the product $\nu \equiv a h_s$, and the value of unemployed search to such an individual is given by $V_U(\nu)$. In terms of the Nash bargaining problem, the worker-firm pair solves
$$\max_w (V_E(w, \nu) - V_U(\nu))^\alpha V_F(w, \theta)^{1-\alpha},$$

where

$$\begin{aligned} V_E(w, \nu) &= \frac{w + \eta V_U(\nu)}{\rho + \eta} \\ V_F(w, \theta, \nu) &= \frac{\theta\nu - w}{\rho + \eta}. \end{aligned}$$

[6] This classification was determined to some extent empirically. Our original classification scheme grouped together all those sample members who had completed some level of schooling beyond high school. We found that those who had attended college but not completed it were far more similar, in terms of labor market outcomes, to those with only a high school education than to those who had completed four years of college. As a result, we grouped together all those who had not completed at least a four-year college degree. Even with this classification, over 1/3 of our sample of 30-34 year old males fell into schooling class $s = 2$.

Note that we have assumed that the firm's outside option under Nash bargaining is equal to 0, which is consistent with the common free entry condition that drives the value of an unfilled vacancy to 0. The solution to the Nash bargaining problem yields

$$w(\theta, \nu) = \alpha\theta\nu + (1-\alpha)\rho V_U(\nu),$$

and since

$$\rho V_U(\nu) \equiv y^*(\nu) = \nu\theta^*(\nu),$$

we have

$$w = \nu(\alpha\theta + (1-\alpha)\theta^*(\nu)). \tag{2}$$

In terms of the value of unemployed search given ν, we have

$$\begin{aligned} \rho V_U(\nu) &= b_0\nu + \lambda \int_{\theta^*(\nu)} (V_E(\nu,\theta) - V_U(\nu)) dG(\theta) \\ \Rightarrow \nu\theta^*(\nu) &= b_0\nu + \frac{\lambda\alpha\nu}{\rho+\eta} \int_{\theta^*(\nu)} (\theta - \theta^*(\nu)) dG(\theta). \end{aligned} \tag{3}$$

Since this last equation is independent of ν, we have

$$\theta^*(\nu) = \theta^* \text{ for all } \nu,$$

which means that the reservation output value for an individual of ability a with schooling level s is simply

$$y^*(a,s) = ah_s\theta^*. \tag{4}$$

This result makes the consideration of the schooling choice problem straightforward. When an individual of type a has schooling level s and enters the labor market, the expected value of the labor market career is given by $V_U(ah_s)$. Then for a type a individual, the value of schooling level s at the time of entry into the labor market is

$$V_U(ah_s) = \rho^{-1}ah_s\theta^*.$$

There is no monetary cost[7] associated with completing schooling level 1, and the present value of the monetary cost associated with completing schooling level 2 is given by $c_2 > 0$ at time τ_1, when schooling level 1 is completed. The first time that the individual can decide to exit school is at the completion of compulsory schooling, which is $s = 1$. The additional time it takes to complete schooling level 2 is given by τ_2. We assume for simplicity that the cost c_2 includes all monetary and psychic costs incurred during the completion of schooling level 2. Then an individual of type a will choose schooling level 2 if and only if

$$\begin{aligned} \exp(-\rho\tau_2)\rho^{-1}ah_2\theta^* - \rho^{-1}a\theta^* &\geq c_2 \\ \Rightarrow a\{\exp(-\rho\tau_2)h_2 - 1\} &\geq \frac{\rho c_2}{\theta^*}. \end{aligned}$$

[7] At least, there is no monetary cost that is avoidable. If there is such a cost, it is fixed and has no influence on the schooling decision.

The left hand side is linear in a, and strictly increasing when

$$\exp(-\rho\tau_2)h_2 > 1. \tag{5}$$

Now assume that

$$\underline{a}\{\exp(-\rho\tau_2)h_2 - 1\} \quad < \quad \frac{\rho c_2}{\theta^*} \tag{6}$$

$$\bar{a}\{\exp(-\rho\tau_2)h_2 - 1\} \quad > \quad \frac{\rho c_2}{\theta^*}, \tag{7}$$

We have the following result.

Proposition 1 *Under conditions (5), (6), and (7), there exists a unique value*

$$a^* = \frac{\rho c_2}{\theta^*\{\exp(-\rho\tau_2)h_2 - 1\}},$$

such that an individual of type a chooses $s = 2$ if and only if

$$a^* \leq a \leq \bar{a}.$$

Proof. The fact that the payoff to college is linearly increasing in a and that for the least able person choice $s = 1$ dominates and for the most able person choice $s = 2$ dominates ensures that there exists a marginal ability type, a^*, who is indifferent between the two choices and that the sets of individuals choosing each schooling level are connected. ■

There are two comments we wish to make regarding this result. First, (5) is required for any agent to acquire schooling beyond the mandatory level. If this condition is not satisfied, no individual completes college. Since we do see individuals completing college in the data, a substantial number of them, this condition is required for the model to have any prima facie validity. Second, given the satisfaction of (5), the more able individuals go to college. In this sense, the net payoff to college attendance is supermodular in the two arguments (a, s), where a is continuous and s is discrete (in fact, binary in the case we are currently considering).

2.3 Comparative Statics Results

Given the simplicity of the decision rule, comparative statics results are easily derived. For the most part, they are intuitively reasonable, which is a strength of this modeling setup.

The focus of the paper is schooling decisions. In our two schooling class model, we can summarize the schooling distribution in terms of the probability that a population member graduates from college, the likelihood of which is

$$P_2 \equiv P(s = 2) = \tilde{F}_a(a^*),$$

where $\tilde{F}_a$ denotes the survivor function associated with the random variable a. The results are:

1. $\partial P_2/\partial c_2 < 0$. The proportion of the population attending college is decreasing in the direct costs of college attendance.

2. $\partial P_2/\partial \tau_2 < 0$. The proportion of the population attending college is decreasing in the time it takes to complete college, which is simply another form of (opportunity) cost associated with continuing education.

3. $\partial P_2/\partial h_2 > 0$. This is perhaps the most intuitive result. The greater the impact on labor market productivity, the more individuals complete college.

4. $\partial P_2/\partial \theta^* > 0$. Now θ^* is not a primitive parameter of course, but most primitive parameters characterizing the labor market only affect the schooling decision through θ^*, which is a determinant of the value of search for all agents (recall that the critical output level for job acceptance is $ah_s\theta^*$). Through this value, we can determine the impact of the most of the various labor market parameters on the schooling decision.

   (a) $\partial P_2/\partial \lambda > 0$. Increases in the arrival rate of offers increase θ^*, and hence increase the value of having a higher productivity distribution.

   (b) $\partial P_2/\partial \eta < 0$. Increases in the (exogenous) separation rate decrease θ^* and hence the value of becoming more productive when matched with an employer.

   (c) $\partial P_2/\partial b_0 > 0$. Increases in the "baseline" flow value of occupying the unemployment state increase the value of that state and the value of going to college.

5. $\partial P_2/\partial \rho < 0$. To be consistent with the definitions of the utility flow associated with employment, which is equal to the wage, the cost measure c_2 is defined as

$$\begin{aligned} c_2 &= \int_0^{\tau_2} e^{-\rho t}\tilde{c}_2 dt \\ &= \frac{\tilde{c}_2}{\rho}(1-\exp(-\rho\tau_2)). \end{aligned}$$

   Then the critical schooling ability level is given by

$$a^* = \frac{\tilde{c}_2(1-\exp(-\rho\tau_2))}{\theta^*\{\exp(-\rho\tau_2)h_2 - 1\}}.$$

   It follows that

$$\begin{aligned} \operatorname{sgn}\left(\frac{\partial a^*}{\partial \rho}\right) &= \operatorname{sgn}\{\tau_2\exp(-\rho\tau_2)\tilde{c}_2\theta^*[\exp(-\rho\tau_2)h_2 - 1] \\ &\quad -\tilde{c}_2(1-\exp(-\rho\tau_2))(\exp(-\rho\tau_2)h_2 - 1)\frac{\partial\theta^*}{\partial\rho} \\ &\quad +\tilde{c}_2(1-\exp(-\rho\tau_2))\tau_2\exp(-\rho\tau_2)h_2\theta^*\}. \end{aligned}$$

   The terms on the first and third line of the right hand side are unambiguously positive. Since θ^* is decreasing in ρ, we have our result.

The main comparative statics result, which is the focus of the paper, concerns the effect of bargaining power α on schooling. While the result is obvious at this point, we state it more formally than the other results.

Proposition 2 *Increases in bargaining power on the workers' side of the market result in increases in schooling level. or*

$$\frac{\partial P_2}{\partial \alpha} > 0.$$

2.4 Empirical Implications

Here we consider the model's implications for the labor market outcomes of individuals in the two schooling classes. In particular, unemployment rates and wage distributions for the two schooling classes.

2.4.1 Unemployment Experiences

Under our modeling assumptions, the steady state unemployment rate for an individual of type ν $(= ah_s)$ is independent of ν. This is due to the fact that the likelihood that any job is acceptable to an individual of type ν is simply $\tilde{G}(\theta^*)$, which is obviously independent of ν. The proportion of time and individual of type ν spends in unemployment, or the steady state probability that they will occupy the unemployment state, is simply

$$P(U|\nu) = \frac{\eta}{\eta + \lambda \tilde{G}(\theta^*)} = P(U).$$

Thus, the assumption that the primitive parameters are identical across schooling groups produces the implication that there is no difference in unemployment experiences across schooling groups.

2.4.2 Wage Distributions

We assume that the support of the matching distribution G is the nonnegative real line, and that G is everywhere differentiable on its support with corresponding density g. We have established that the schooling continuation set is defined by $[a^*, \bar{a}]$. Now, from (2) we know that

$$\theta = \frac{\frac{w}{\nu} - (1-\alpha)\theta^*}{\alpha},$$

where $\nu = ah_s$. and the lower limit of the wage distribution for an individual of type ν is $\underline{w}(\nu) = \nu\theta^*$. Then the cumulative distribution function of wages for a type ν individual is

$$F_w(w|\nu) = \frac{G(\alpha^{-1}(\frac{w}{\nu} - (1-\alpha)\theta^*)) - G(\theta^*)}{\tilde{G}(\theta^*)}, \ w \geq \nu\theta^*,$$

and the corresponding conditional wage density is given by

$$f_w(w|\nu) = \frac{1}{\alpha\nu}\frac{g(\alpha^{-1}(\frac{w}{\nu} - (1-\alpha)\theta^*))}{\tilde{G}(\theta^*)},\ w \geq \nu\theta^*.$$

Now we consider the wage densities by schooling class. For this purpose, we write

$$f_{w|a,s}(w|a,s) = \frac{1}{\alpha a h_s}\frac{g(\alpha^{-1}(\frac{w}{a h_s} - (1-\alpha)\theta^*))}{\tilde{G}(\theta^*)},\ w \geq a h_s \theta^*.$$

Then the marginal density of wages in schooling class s is given by

$$f_{w|s}(w|s) = \frac{1}{\alpha h_s \tilde{G}(\theta^*)}\int a^{-1} g(\alpha^{-1}(\frac{w}{a h_s} - (1-\alpha)\theta^*))dF(a|s),\ w \geq h_s \underline{a}(s)\theta^*,$$

where $\underline{a}(s)$ denotes the lowest ability individual who makes schooling choice s. Given the simple form of the schooling continuation decision, the density of wages among those with a high school education is

$$f_{w|s}(w|s=1) = \frac{1}{\alpha \tilde{G}(\theta^*)}\int_{\underline{a}}^{a^*} a^{-1} g(\alpha^{-1}(\frac{w}{a} - (1-\alpha)\theta^*))\frac{dF(a)}{F(a^*)},\ w \geq \underline{a}\theta^*, \tag{8}$$

while the density of wages among the college-educated population is

$$f_{w|s}(w|s=2) = \frac{1}{\alpha h_2 \tilde{G}(\theta^*)}\int_{a^*}^{\bar{a}} a^{-1} g(\alpha^{-1}(\frac{w}{a h_2} - (1-\alpha)\theta^*))\frac{dF(a)}{\tilde{F}(a^*)},\ w \geq a^* h_2 \theta^*. \tag{9}$$

The conditional wage densities for the two schooling groups differ, then, not only because college education improves the productivity of any individual who acquires it, but also through the systematic selection induced on the unobserved ability distribution F_a by the option of going to college. In terms of the conditional (on s) wage distributions, we note that the upper limit of the support of both distributions is ∞. The distributions do differ in their lower supports, with this lower bound equal to $\underline{a}\theta^*$ for those with high school education and $a^* h_2 \theta^*$ for those with college. Since $a^* h_2 > \underline{a}$, the lower support of the distribution of the college wage distribution lies strictly to the right of the high school wage distribution.

Proposition 3 *The wage distribution of the college educated first order stochastically dominates that of the high school educated.*

Proof. Since $h_2 > h_1 = 1$, for any a, $F_w(w|a,h_2)$ first order stochastically dominates $F_w(w|a,h_1)$. For any s, $F_w(w|a',s)$ first order stochastically dominates the distribution $F_w(w|a,s)$ whenever $a' > a$. Since $a' \geq a^* > a$ for all $a' \in [a^*, \bar{a}]$ and $a \in [\underline{a}, a^*)$, the mixture distributions are strictly ordered in the sense

$$F_{w|1}(w|1) \geq F_{w|2}(w|2) \text{ for all } w \geq \underline{a}\theta^*.$$

■

From this result, it immediately follows that the average wage is greater among the college educated. More importantly, the wages of the college-educated exceed those with a high school education at every quantile of the respective distributions.

Before proceeding to investigate some extensions of the basic model, we want to examine some descriptive evidence regarding the empirical implications of the model. The data used in all of the empirical analysis below will be described in more detail in the sequel. In terms of the general characteristics of the sample, it is drawn from monthly Current Population Survey samples from 2005, and consists of males living in CPS households who were between the ages of 30 to 34, inclusive when interviewed. The "high schooling" category, corresponding to $s = 2$, consists of individuals who have completed (at least) a four year college program. The "low schooling" category is all others. The hourly wage data are taken from sample members who were employed at the time of the interview, and are the actual hourly wages if the individual is paid on this basis or are imputed by dividing usual weekly earnings by usual weekly hours. We eliminated outliers by trimming the lowest and highest 2.5 percent of wage observations from both schooling subsamples.

Figure 1.a and 1.b contain the plots of the wage distributions by schooling group. The minimum wage observed (after trimming) for the low schooling group is 6.00 and for the high schooling group is 7.50. We see from these figures that the wages of the low schooling group members are highly concentrated in the range 6 to 20 dollars, while the high schooling group wages show considerably more dispersion. Figure 1.c displays the distribution of total wages. Approximately 1/3 of those in the wage sample have completed college, so that the marginal wage distribution closely resembles the non-college wage distribution at low values of w. This is not the case at high wage levels, where virtually all observations are associated with college-completers.

Proposition 3 implies that the high schooling wage distribution first order stochastically dominates the low schooling wage distribution. Figure 1.d presents evidence regarding this implication. The figure plots

$$F(w_{(i)}|s = 1) - F(w_{(i)}|s = 2)$$

for an increasing sequence of wages, $w_{(1)} < ... < w_{(K)}$. First order stochastic dominates implies that all values in this sequence should be nonnegative, and the figure strongly bears out this claim.

The homogeneous labor market assumption is not consistent with the observed unemployment rate for the two education groups. From Table 1, we note that the unemployment rate for college-completers is only .017, while among those with less schooling it is 0.041. A simple adjustment allows the model to produce such an outcome, and we now turn to this generalization.

3 Separate Schooling Sub-Markets

We continue within the partial equilibrium setting of the previous section, but consider relaxing some of the more restrictive (from an empirical perspective) features of that model. In particular, we know from the large number of structural estimation exercises involving search models that the primitive parameters across sub-markets are often found to be markedly different (see, for example, Flinn (2002)). In particular, it is often noted that the unemployment rate differs across schooling groups, with those with lower completed schooling yields often having lengthier and more frequent unemployment spells. As we saw above, such a result is not consistent with the assumption that all primitive labor market parameters are the same across schooling classes.

The situation we consider is one in which each schooling class inhabits a sub-labor market, which has its own market-specific parameters $(\lambda_s, \eta_s, \alpha_s)$. The parameter ρ, being a characteristic of individual agents (individuals and firms), is assumed to be homogeneous across labor markets, as is the baseline unemployment utility flow parameter b_0. The match productivity distribution G is also identical across markets. In terms of the productivity of an individual, nothing has changed from the previous case, since $y(a, s, \theta) = ah_s\theta = \nu\theta$, so that the distribution of y is a function of the scalar ν and the common (to all matches) distribution G. However, it is no longer the case that the critical match value will be the same across schooling sub-markets. Because primitive parameters differ across markets, ν is no longer a sufficient statistic for the value of search of an individual; instead a minimal sufficient statistics is the pair (ν, s). This is clear if we reconsider the functional equation determining the value of search in the homogeneous sub-markets case, which was given in (3), adapted to the heterogeneous case. Then we have

$$\nu\theta^*(\nu, s) = b_0\nu + \frac{\lambda_s\alpha_s\nu}{\rho + \eta_s}\int_{\theta^*(\nu,s)}(\theta - \theta^*(\nu, s))dG(\theta).$$

The solution $\theta^*(\nu, s)$ now clearly is independent of ν, but is not independent of s. Thus there is a common critical value $\theta^*(s)$ shared by all individuals with schooling choice s, which is independent of their ability a.

We now turn to the schooling choice decision in this case. The critical match value for an individual of type a in schooling market s is given by $ah_s\theta^*(s) = \nu\theta^*(s)$, so that the value of unemployed search in this submarket is given by $\rho^{-1}\nu\theta^*(s)$. Then the net value of college education to an individual of type a is

$$\exp(-\rho\tau_2)\rho^{-1}ah_2\theta^*(2) - c_2 - \rho^{-1}a\theta^*(1).$$

We have the following result.

Proposition 4 *In the heterogeneous schooling submarkets case, a necessary and sufficient condition for a measurable set of individuals to go to college is*

$$\exp(-\rho\tau_2)h_2\theta^*(2) > \theta^*(1). \tag{10}$$

Given this condition, if

$$\begin{aligned}\underline{a}\{\exp(-\rho\tau_2)h_2\theta^*(2) - \theta^*(1)\} &< \rho c_2 \\ \bar{a}\{\exp(-\rho\tau_2)h_2\theta^*(2) - \theta^*(1)\} &> \rho c_2,\end{aligned}$$

then there exists a unique value of a^* *such that*

$$\exp(-\rho\tau_2)\rho^{-1}a^*h_2\theta^*(2) - c_2 - \rho^{-1}a^*\theta^*(1) = 0,$$

and an individual of type a *chooses college if and only if* $a \in [a^*, \bar{a}]$.

The proof is an obvious extension of the homogeneous schooling sub-markets case that we have just discussed.

The condition (10) is required for anyone, of any ability, to go to college. Note that in terms of primitive labor market parameters, we could have that the "normalized" value of search, determined solely from $\theta^*(s)$, was worse in the college market than in the high school market so long as the length of college completion was not too long, direct costs of attending college are not too large, and the productivity enhancement of college, h_2, is sufficiently large. We think of the condition (10) as representing a restriction on the parameter space involving the labor market parameters that make the model consistent with our model of schooling selection. If this condition is not satisfied, then clearly our model is fundamentally misspecified.

3.1 Comparative Statics Results

We can rewrite the critical value property in a more transparent manner to investigate comparative statics,

$$a^* = \frac{\rho c_2}{\{\exp(-\rho\tau_2)h_2\theta^*(2) - \theta^*(1)\}}.$$

Comparative statics results are fundamentally different in this case in the sense that certain market-specific primitive parameters only impact the value of unemployed search within their particular submarket. By simple extension of the homogeneous results above, the results regarding $\partial P_2/\partial c_2 < 0$ and $\partial P_2/\partial \tau_2 < 0$ remain the same, since the cost structure of acquiring schooling is identical in the two cases. It is also clearly the case that $\partial P_2/\partial h_2 > 0$. The main departure from the previous case regards the presence of $\theta^*(1)$ and $\theta^*(2)$. We note that

1. $\partial P_2/\partial\theta^*(1) < 0$. As before, $\theta^*(1)$ is not a primitive parameter, but the primitive parameters specific to submarket 1 only affect the schooling decision through $\theta^*(1)$. Then

   (a) $\partial P_2/\partial\lambda_1 < 0$. Increases in the arrival rate of offers in the low-schooling market increase $\theta^*(1)$, and increase the relative value of a low schooling level.

(b) $\partial P_2/\partial \eta_1 > 0$. Such an increase decreases the value of a low schooling level.

2. $\partial P_2/\partial \theta^*(2) > 0$.

   (a) $\partial P_2/\partial \lambda_1 > 0$

   (b) $\partial P_2/\partial \eta_1 < 0$

3. Perhaps most interesting is the impact of market-specific bargaining powers α_s on the schooling decision. When there is one bargaining power parameter that holds throughout all educational labor markets, the meaning of hold-up is relatively unambiguous. When there are market-specific bargaining power parameters, a relative notion of hold-up is more appropriate. Clearly we have

$$\begin{aligned} \frac{\partial P_2}{\partial \alpha_1} &< 0. \\ \frac{\partial P_2}{\partial \alpha_2} &> 0. \end{aligned}$$

   It is important to note that α_2 could be quite low, and yet a substantial proportion of agents may choose the high schooling level if α_1 is significantly lower yet.

3.2 Empirical Implications

There are a few obvious differences in the empirical implications of the homogeneous and heterogeneous labor market models.

3.2.1 Unemployment

The characteristic scalar ν is no longer sufficient for describing an individual's probability of labor market events, in general, except when ν implies a unique value of s. While this may be the case as long as there exists a unique a^* that determines the schooling continuation set and $h_2 > 1$, we will condition on both ν and s to make the situation a bit more transparent. Now, in general, we will have different steady state unemployment rates for the two schooling groups, since

$$P(U|\nu, s) = \frac{\eta_s}{\eta_s + \lambda_s \tilde{G}(\theta_s^*)} = P_s(U), \;\; s = 1, 2.$$

As before, within a schooling group unemployment probabilities are homogeneous. Without further restrictions on the event rate parameters and the bargaining power parameters, it is not possible to order the unemployment probabilities across schooling sectors.

3.2.2 Wage Distributions

The lower bound on the support of the wage distribution associated with schooling type s is now given by $\underline{w}(1) = \underline{a}\theta^*(1)$ for the low schooling group and by $\underline{w}(2) = a^* h_2 \theta^*(2)$ for the college completers. The conditional density of wages for the low schooling group is given by

$$f_{w|a,1}(w|a,1) = \frac{1}{\alpha_1 a} \frac{g(\alpha_1^{-1}(\frac{w}{a} - (1-\alpha_1)\theta^*(1)))}{\tilde{G}(\theta^*(1))}, \ w \geq \underline{a}\theta^*(1),$$

while the wage density for the high schooling group is

$$f_{w|a,2}(w|a,2) = \frac{1}{\alpha_2 h_2 a} \frac{g(\alpha_2^{-1}(\frac{w}{ah_2} - (1-\alpha_2)\theta^*(2)))}{\tilde{G}(\theta^*(2))}, \ w \geq a^* h_2 \theta^*(2).$$

Since the model with heterogeneous schooling submarkets continues to imply that those who continue to schooling level $s = 2$ form a connected set $[a^*, \overline{a}]$, the unconditional (on a) wage densities have the simple forms

$$f_{w|s}(w|s=1) = \frac{1}{\alpha_1 \tilde{G}(\theta^*(1))} \int_{\underline{a}}^{a^*} a^{-1} g(\alpha_1^{-1}(\frac{w}{a} - (1-\alpha_1)\theta^*(1))) \frac{dF(a)}{F(a^*)}, \ w \geq \underline{a}\theta^*(1),$$

and

$$f_{w|s}(w|s=2) = \frac{1}{\alpha_2 h_2 \tilde{G}(\theta^*(2))} \int_{a^*}^{\bar{a}} a^{-1} g(\alpha_2^{-1}(\frac{w}{ah_2} - (1-\alpha_2)\theta^*(2))) \frac{dF(a)}{\tilde{F}(a^*)}, \ w \geq a^* h_2 \theta^*(2).$$

4 Econometric Issues

We begin this section by discussing the data utilized to estimate the model(s). We then proceed to a discussion of the estimation method and identification of model parameters given the nature of the data available. Model estimates and comparative statics exercises are presented in the following section.

4.1 Data

Because identification of model parameters hinges on rather "fine" features of the conditional (on schooling) and unconditional empirical wage distributions, it is essential to have precise estimates of these distributions which can only be obtained using samples with large numbers of observations. For this reason, we utilize the Current Population Survey Outgoing Rotation Groups (CPS-ORG) from all of the months in the calendar year 2005. ORGs include households in their 4th or 8th survey month, and in these months detailed earnings and employment information is ascertained. We selected males between the ages of 30 to 34, inclusive. We made no further restrictions on sample inclusion that were unrelated to missing information or labor market status. In particular, we did not exclude individuals based on race, ethnicity, or region of residence. Thus, while our sample inclusion criteria are relatively restrictive, a considerable amount of heterogeneity remains.

To be included in the final sample, an individual had to either be employed or unemployed, and had to have valid schooling completion information. If an individual was employed, to be included in the final sample there had to have been enough information available that would allow an hourly wage rate to at least be imputed.[8] If an individual was unemployed, we required that they report the weeks of the on-going search spell to be included in the estimation sample. Our final sample consists of 9,985 individuals.

After experimenting with various schooling classifications systems, we determined the one that seemed to maximize differences in schooling group outcomes. This involved assigning all those who had completed college to the high schooling group and all those with partial college or less to the low schooling group. We began by assigning all of those with any college to one group, but found that those with less than four years of college were far more similar in their labor market outcomes to those who had not attended college at all than they were to those who had completed at least four years of college.

The wage distribution is a complicated function of the underlying sources of heterogeneity in match values and ability. While it is surely the case that reported wage data contain considerable amounts of measurement error[9], how this measurement error should be introduced into an econometric model is controversial. Moreover, adding another random variable to the two already determining the "actual" wage further complicates the deconvolution problem. For this reason, we have utilized symmetric trimming as a partial solution to deleting some forms of measurement error that are likely to severely distort model estimates. Originally, there were 6,416 employed individuals at the low schooling level and 3,238 at the high schooling level. From each set, we eliminated the top and bottom 2.5 percent of wage observations. All the wage information used in the estimator was taken from the trimmed samples. We weight the likelihood contributions of the high school and college employed to compensate for the cases that were trimmed from the final sample.

Table 1 contains the descriptive statistics from the final estimation sample. We see that 33 percent of males in this age range have completed at least a four year college program. In the year 2005, well before the "Great Recession," the unemployment rate of this group was quite low, at 3.3 percent, with a significant difference between the unemployment rates of the low school group (4.1 percent) and of the high schooling group (1.7 percent), upon

[8] That is, an individual who reported being paid on an hourly basis and who reported their hourly wage rate would be included in the sample. Most males of this age range are not paid on an hourly basis, however. In these cases, if usual weekly hours and usual weekly earnings were reported, we could impute a "usual" hourly wage rate. Thus both types of individuals were included in the employed subsample.

[9] For one validation study using the Panel Study of Income Dynamics (PSID) survey, see Bound et al (1994). The authors find evidence of substantial amounts of measurement error, although there are a number of methodological problems with the validation sample that raise some concerns regarding the numerical results.

which we have already commented. On average, an unemployed sample member had been searching for work for 5.122 months, and those who have finished college on average have searched 0.8 months less.

The average hourly wage (imputed for most observations) in the sample is \$18.38, with a substantial difference between those with less than college completion (\$15.13) and those who have completed college (\$24.83). There is also a substantial degree of dispersion in these distributions (even after having trimmed the lowest and highest 2.5 percent of wage observations from each conditional wage distribution). As we shall see below, the minimum observed wages in the conditional wage distributions will be important to our identification strategy. We see that the minimum observed wage for the less-than-college-completion group is \$6.00, while the minimum wage among the college completers is \$7.50.

4.1.1 Supplemental Data

Due to the challenges of identifying the underlying match value and ability distributions, even under strong parametric assumptions, we utilize some additional data sources when estimating the two specifications of the model. The first is simply information on the labor share of the surplus. We will utilize this information in the same way as it was used in Flinn (2006); he showed that this information was virtually essential to enable identification of the bargaining power parameter, α. The discussion in Krueger (1999) led us to believe that 0.67 was a reasonable value to use for labor's share for this group of labor market participants. In Flinn's (2006) study of minimum wage effects on labor market outcomes, labor's share was computed from the Consolidated Income Statement of McDonald's corporation for 1996 and was found to be about 53 percent. This was deemed reasonable since the CPS data used were for workers between the ages of 16 to 24, inclusive. Thus for workers in the age range 30 to 34, the value of 0.67 seems to be in the right range of values.

The CPS-ORG can be used to construct short-panels. Each household is a member of the ORG subsample in their 4th and 8th month in the sample. Given the rotation design of the CPS, the 4th and 8th sample months are separated by one year. Thus, in principle, we have up to two observations on the earnings and wages of each sample member. A problematic feature of the data is that we do not know with certainty whether an individual occupies the same job or not at the two points in time.[10] For the moment, assume that we know that the an individual who is employed in sample month 4 and sample month 8 is not in the same job. Then, under the model, the difference in the log wage

[10] In fact, we do not even know with certainty the identify of the household members at the two points in time. We must rely on a statistical matching procedure, which looks at age, race, sex, and education at the two points in time, to determine which records should be paired. While the procedure should work in most cases, there almost certainly are some erroneous matches.

rates is given by

$$\begin{aligned}\ln w_{i,j',t+12}-\ln w_{i,j,t} &= \ln(a_i h_i[\alpha\theta_{i,j'}+(1-\alpha)\theta^*]) \\ &\quad -\ln(a_i h_i[\alpha\theta_{i,j}+(1-\alpha)\,\theta^*]) \\ &= \ln(\alpha\theta_{i,j'}+(1-\alpha)\theta^*)-\ln(\alpha\theta_{i,j}+(1-\alpha)\,\theta^*),\end{aligned}$$

so that the distribution of the log wage differences is only a function of α, θ^*, and the parameters describing the distribution of θ, which under our lognormal assumption are μ_θ and σ_θ. Note that

$$E\Delta\ln w_i = 0$$

since the successive θ draws are assumed to be identically and independently distributed. However,

$$E\,(\Delta\ln w_i)^2 = 2E[\ln(\alpha\theta+(1-\alpha)\theta^*)]^2 > 0,$$

and we also find that the fourth noncentral moment of the log wage differences takes a positive value. Then we have

$$\begin{aligned}E\,(\Delta\ln w_i)^2 &= d_1(\mu_\theta,\sigma_\theta;\alpha,\theta^*) > 0 \\ E(\Delta\ln w_i)^4 &= d_2(\mu_\theta,\sigma_\theta;\alpha,\theta^*) > 0.\end{aligned}$$

We utilize the analogous second and fourth noncentral sample moments to consistently estimate the left hand side of this function. Given values for α and θ^*, we can then solve the two equation system for unique values of μ_θ and σ_θ.

In order to maintain independence from our cross-sectional (2005) sample, we use CPS data from the years 2006 and 2007, adjusted for inflation. As we have noted, we do not access to information that would allow us to determine with certainty whether an individual is at the same job or not at the two points in time. Instead, we utilize information on the characteristics of the individual's job as a crude indicator of whether the job has changed. In particular, only individuals in the CPS-ORG who are employed in the 4th and 8th sample months and who do not work in the same 3-digit occupation at those two sampling times are assumed to have changed jobs. Undoubtedly some people who change jobs over the one year interval remain in the same 3-digit industry, so ours would seem to be a conservative criterion upon which to define a job-changing sample. On the other hand, there are certainly response and coding errors that result in our including individuals in this sample who have not changed jobs. We hope the misclassification errors are not too large, though there is no validating information available in the CPS that allow us to say anything further on this point.

4.2 Identification

The primitive parameters in the homogeneous markets case are $\rho, b, \lambda, \eta, F_a, G, h_2$, and $\tilde{c}_2$. Much of the identification analysis can be conducted using results from Flinn and Heckman

(1982) after noting which of the parameters determine labor market outcomes explicitly once we condition on the observed value of schooling, s. For most of our discussion, we assume that the distribution of flow costs of schooling, $\tilde{c}$, is degenerate. At the end of this section, we consider the case in which it is not.

As we showed above, conditional on s, randomness in labor market outcomes (across individuals and over time) is generated by the two independent random variables ν and θ. As we have shown, under our model assumptions the critical match value θ^* is independent of s. We also showed that the model implies that all individuals with an ability level less than a^* chose $s = 1$, while all others choose $s = 2$ (college completion). Under the normalization $h_1 = 1$, the lower bound of the support of the wage distribution of the low-schooling group is

$$\underline{w}_1 = \theta^* \underline{a},$$

while the corresponding value for the high-schooling group is

$$\underline{w}_2 = \theta^* a^* h_2.$$

Just as Flinn and Heckman (1982) showed that parametric assumptions were, in general, necessary to recover the parameters of the wage offer distribution in the partial-partial equilibrium search case, they will also be necessary here for similar reasons. Parametric assumptions on F_a also include the specification of the support of the distribution, of course. In this case, we assume that $\underline{a} = 1$, so that $\underline{w}_1 = \theta^*$. Then from Flinn and Heckman (1982), we know that

$$\hat{\theta}^* = \min\{w_i\}_{i \in S_1},$$

where S_1 is the set of sample members in the low-schooling group, is a superconsistent estimator of θ^* when there is no measurement error in wages.

The value of a^*, which characterizes the schooling decision rule, is a function of all of the parameters in the model, including the the flow cost of attending school, $\tilde{c}_2$. It is clear that this "free" primitive parameter only enters the schooling decision directly, and thus for estimation purposes we can treat a^* as a parameter to be estimated. If all other parameters determining a^* are identified, then the estimated value of a^* can be inverted to yield an estimate of $\tilde{c}_2$ since a^* is monotone in $\tilde{c}_2$.

Identification discussions are typically facilitated when a likelihood function is available, and we utilize a likelihood-based estimation method. We really have three separate groups of observations defined in terms of likelihood contributions in the homogeneous labor market case. Since the unemployment experiences of individuals are independent of their schooling level, the likelihood contribution of a sample member from this group is given by

$$L_k(\tilde{t}_{U,k}, U_k) = h_U \exp(-h_U \tilde{t}_{U,k}) \times \frac{\eta}{\eta + h_U}, \ k \in S_U,$$

where the first term on the right hand side of the expression is the likelihood of being in an on-going search spell of duration $\tilde{t}_{U,k}$ given the individual is unemployed at the sampling time and the second term in the product is the probability of being unemployed at a random sampling time in the steady state, S_U is the set of unemployed individuals in the point sample, and the hazard of leaving unemployment is

$$h_U = \lambda \tilde{G}(\theta^*).$$

The second group of contributors to the likelihood function are those in the low-schooling group who are employed at the time of the sample. For these individuals, we know their wage rate, their schooling level, and the fact that they were employed, with their likelihood contribution given by

$$\begin{aligned}
L_k(w_k, s_k &= 1, E_k) = f_{W|S=1}(w_k|s_k = 1, E_k) \times p(s_k = 2|E_k) \times p(E_k) \\
&= f_{W|S=1}(w_k|s_k = 1) \times p(s_k = 2) \times p(E_k) \\
&= \frac{1}{\alpha\tilde{G}(\theta^*)} \int_1^{a^*} a^{-1} g(\alpha^{-1}(\frac{w}{a} - (1-\alpha)\theta^*)) \frac{dF(a)}{F(a^*)} \times F(a^*) \times \frac{h_U}{\eta + h_U} \\
&= \frac{1}{\alpha\tilde{G}(\theta^*)} \int_1^{a^*} a^{-1} g(\alpha^{-1}(\frac{w}{a} - (1-\alpha)\theta^*)) dF(a) \times \frac{h_U}{\eta + h_U},\ w \geq \underline{a}\theta^*,\ k \in S_E \cap S_1.
\end{aligned}$$

Note that there is an additional restriction on the limits of integration. If the value w is observed for an individual at schooling level 1, then the value of a for that individual can be no larger than $\bar{a}(w, s = 1)$, which is given by

$$\begin{aligned}
w &= \bar{a}(w, s = 1)\theta^* \\
\Rightarrow \bar{a}(w, s = 1) &= \frac{w}{\theta^*}.
\end{aligned}$$

We can either adjust the upper limit of integration directly or multiply the integrand by the indicator function $\chi[a \leq w/\theta^*]$.

The third group of contributors to the likelihood function are those in the high-schooling group who are employed at the time they are sampled. The construction of the likelihood contribution for this group is analogous to that of the low-schooling group, so that

$$L_k(w_k, s_k = 2, E_k) = \frac{1}{\alpha h_2 \tilde{G}(\theta^*)} \int_{a^*}^{\bar{a}} a^{-1} g(\alpha^{-1}(\frac{w}{ah_2} - (1-\alpha)\theta^*)) dF(a) \times \frac{h_U}{\eta + h_U},\ w \geq a^* h_2 \theta^*,\ k \in S_E \cap S_2.$$

Also for this group there is a restriction on the limits of integration of a, with the upper bound on a given by

$$\begin{aligned}
w &= \bar{a}(w, s = 2) h_2 \theta^* \\
\Rightarrow \bar{a}(w, s = 2) &= \frac{w}{h_2 \theta^*}.
\end{aligned}$$

As shown in Flinn (2006), for example, knowledge of the proportion unemployed and the average duration of unemployed search from a point sample sufficient to identify the rate parameter η and the hazard function h_U. Now the hazard rate associated with the unemployment state can be inverted to yield is given by

$$\lambda = \frac{h_U}{\tilde{G}(\theta^*)}.$$

Since we have a superconsistent estimator of θ^* and a $\sqrt{N}$ consistent estimator of h_U, if consistent estimators of the finite-dimensional parameter vector characterizing G are available, then

$$\hat{\lambda} = \frac{\hat{h}_U}{1 - \hat{G}(\widehat{\theta^*})}$$

is a consistent estimator of λ.

The identification of α is primarily achieved through the use of labor share information, so for purposes of discussion we will assume that a consistent estimator of α is available. Identification of the distributions of the components determining total match productivity is extremely challenging using only point sample wage data. As is evident from (8) and (9), the schooling-specific wage distributions are mixtures of a truncated lognormal distributions, $G(\theta|\theta \geq \theta^*)$ with the mixing distribution, F_a, representing the (truncated, under the model) distribution of abilities within schooling level s. As stated above, parametric assumptions are required for the identification of $G(\theta)$, and in the empirical work below we make the common assumption that the match values are lognormally distributed, so that

$$G(\theta; \mu_\theta, \sigma_\theta) = \Phi(\frac{\ln \theta - \mu_\theta}{\sigma_\theta}),\ \ \theta \in R_+,$$

where Φ is the standard normal c.d.f. Thus G is assumed to be completely characterized given knowledge of the two parameters μ_θ and σ_θ.

It is a practical necessity to make parametric assumptions regarding what is essentially a mixing distribution, F_a. The literature on the estimation of models that involve mixing without parametric assumptions on the mixing distribution (e.g., Heckman and Singer (1984)) make clear that in general the task is enormously difficult even when there exist vast quantities of data. Experimentation with various assumptions on F_a led us to a two-parameter power distribution with support $[1, 1+m]$, $m > 0$, where

$$\begin{aligned} F_a(a) &= \left(\frac{a-1}{m}\right)^{\gamma}, \\ f_a(a) &= \frac{\gamma}{m}\left(\frac{a-1}{m}\right)^{\gamma-1},\ \text{with } 1 \leq a \leq 1+m,\ \gamma > 0,\ m > 0. \end{aligned}$$

Under our model specification and the nature of the decision rules used by agents, the distribution of wages by schooling level are functions of the parameters (both primitives and decision rules): $\mu_\theta, \sigma_\theta, \gamma, m, \alpha, h_2, \theta^*, a^*$. We already have access to a superconsistent estimator of θ^* and we are assuming that a consistent estimator of α is available. We find that even with access to a relatively large number of wage observations (over 9000), identification of the parameters and decision rules characterizing the wage distributions is not possible without using additional information. In particular, we use the fact that a superconsistent estimator of the product a^*h_2 is given by

$$\begin{aligned} \widehat{a^*h_2} &= \frac{\min\{w_i\}_{i\in S_2}}{\min\{w_i\}_{i\in S_1}} && (11) \\ &\equiv \nabla \min w && (12) \end{aligned}$$

to reduce the dimensionality of the parameter space by one. Then we will define the estimator of a^* by

$$\widehat{a^*} = \frac{\nabla \min w}{\hat{h}_2}.$$

Even using after reducing the dimensionality of the parameter space, identification of the parameters characterizing G and F_a is problematic with wage samples of our size. It is for this reason that we employ the information on wage change information for those in the CPS-ORG who change their jobs between the 4th and 8th sample month and which was discussed in the previous subsection. Given consistent estimators for θ^* and α, we saw that noncentral sample moments involving the wage differences could be used to identify μ_θ and σ_θ. Then we can further reduce the dimensionality of the parameters to be estimated by maximum likelihood by substituting

$$\begin{aligned} \hat{\mu}_\theta &= d_1'(\zeta_2, \zeta_4; \widehat{\theta^*}, \hat{\alpha}) && (13) \\ \hat{\sigma}_\theta &= d_2'(\zeta_2, \zeta_4; \widehat{\theta^*}, \hat{\alpha}), && (14) \end{aligned}$$

where ζ_2 and ζ_4 are the sample analogs of the noncentral second and fourth moments of log wage changes among those changing jobs.

After making all of these substitutions into the log likelihood function, we are left with the task of estimating γ, m, h_2 and α. In order to estimate α, we follow the procedure utilized in Flinn (2006). This involves conditioning on a value of α, $\hat{\alpha}$, and determining $\hat{\mu}_\theta$ and $\hat{\sigma}_\theta$ using d_1' and d_2', and then estimating the parameters γ, m, h_2, and $a^*(h_2, \nabla \min w)$. Given the values of the maximum likelihood estimates, we then use the labor share equation to solve for the implied value of α. The estimate of α is updated to this new value, and the maximum likelihood estimates are obtained given this value. The process is repeated until convergence of the sequence of $\hat{\alpha}$ values.

4.2.1 An Alternative Specification

Even with all of the substitutions and use of superconsistent estimators, estimates of model parameters are very sensitive to the specification of the mixing distribution F_a. For this reason, we also estimated a second model in which the distribution of ability is degenerate with all mass at $a = 1$. In this case, to have a well-defined schooling decision rule, it is necessary to assume that the instantaneous cost of schooling is nondegenerate. We denote the distribution of these schooling values by $F_{\tilde{c}}$.

In this case, the decision to attend school is simply given by

$$c \leq \exp(-\rho\tau_2)\rho^{-1}h_2\theta^* - \rho^{-1}\theta^*,$$

where $c = \tilde{c} \times \frac{1}{\rho}[1 - \exp(-\rho\tau_2)]$. Thus, there exists a critical value of the instantaneous cost, $\tilde{c}^*$, such that all individuals with $\tilde{c} \leq \tilde{c}^*$ invest in higher education and all those with $\tilde{c} > \tilde{c}^*$ do not.

We assume that the distribution of costs, which are not observable, is exponential, with

$$F(\tilde{c}) = 1 - \exp(-\xi\tilde{c}),\ \tilde{c} \geq 0,\ \xi > 0.$$

The probability of completing higher education is then given by

$$\begin{aligned} P_2 &= F(\tilde{c}^*) \\ &= 1 - \exp(-\xi\tilde{c}^*). \end{aligned}$$

The wage distribution is now simply generated by the lognormal draws from the distribution G and the human capital level h_2. This simplifies the identification of μ_θ and σ_θ, which can now be precisely estimated simply from the cross-section CPS-ORG observations. In terms of identification of the cost distribution, we recover the parameter ξ in the following way. First of all, without any ability heterogeneity, we have

$$\operatorname*{plim}_{N\to\infty} \nabla \ln w = h_2,$$

and we have a superconsistent estimator of h_2. Then note that

$$\tilde{c}^* = \theta^* \times \frac{h_2 \exp(-\rho\tau_2) - 1}{1 - \exp(-\rho\tau_2)}.$$

Given that we have access to superconsistent estimators of θ^* and h_2, and given knowledge of ρ and τ_2 (which is 48 months in our application), we have access to consistent estimators of all unknown parameters or decision rules appearing on the right hand side, and thus we can consistently estimate $\tilde{c}^*$. The distribution of $\tilde{c}$ only influences the likelihood of completion of college completion, so that the (conditional) maximum likelihood estimator of ξ is

$$\hat{\xi} = -\frac{\ln(1 - \hat{P}_2)}{\widehat{\tilde{c}^*}}.$$

Estimation of α depends crucially on the labor share information, as before, and the estimation method iterates between solution of the labor share equation and conditional (on $\hat{\alpha}$) maximum likelihood estimation of the remaining parameters.

5 Model Estimates

We begin by describing the estimates from the heterogeneous ability model. Due to the sensitivity of the estimates to the specification of the distribution of a, we will quickly move on to consider the homogeneous ability model estimates, and we will be conducting the comparative statics exercises using those estimates. We will be able to see, from inspection of the fitted wage distributions under the homogeneous ability specification, why adding heterogeneous ability is somewhat "gratuitous," at least under our distributional assumptions. As we have mentioned previously, we only estimate the homogeneous labor market model given the severe identification problems we face using the data at our disposal.

5.1 Estimates of the Heterogeneous Ability Model

The estimates of this model appear in Table 2. A few things to bear in mind before examining the estimates are the following. Throughout the entire estimation exercise, the (super) consistent estimate of θ^* is 6.00, obtained after trimming the data. The (super) consistent estimate of $\alpha^* h_2$ is equal to $\nabla \min w = 7.50/6.00 = 1.25$. Thus given the manner in which the model is estimated, it will always be the case that $\widehat{a^*} \times \hat{h}_2 = 1.25$. Since for the model to make sense it is necessary for $h_2 > 1$, this means that the estimated value of a^* cannot be be greater than 1.25. This imposes some restriction on the distributions of a that can be considered. This led us to the choice of a (truncated) power distribution with support $[1, 1+m]$. Since a^* must be strictly less than 1.25, and since approximately 1/3 of sample members are college graduates, we should have an estimate of $1 - \hat{F}_a(\widehat{a^*})$ in the vicinity of 1/3 if we are to fit the education margin reasonably well. This imposes some restrictions on the value of m and γ. Of course, the shape of the distribution of a has ramifications for the wage distribution as well as the schooling distribution, which is what makes the estimation of this specification so challenging.

The estimated arrival rate of job offers (0.195) implies that the average wait for a contact is approximately 5 months, and the estimated rate of job destruction (0.006) implies that jobs last almost 14 years on average.[11] The estimated values of the parameters characterizing the lognormal match distribution are reasonable, as well, though we should bear in mind that they are mainly determined by the values of the supplementary information contained in the wage change information for job leavers. When we attempted to estimate this model without using this information, we were not able to find well-behaved maximum likelihood estimates of the matching and ability distributions.

Although imprecisely estimated, the point estimate of the upper support of the ability distribution is 1.213 (recall that the lower support point is normalized to 1). The estimated value of the power distribution parameter is 0.812 with an asymptotic standard error of

[11] The estimated destruction rate would be much higher, of course, if more recent CPS data were used. Moreover, this group of males between 30 and 34 years of age can be expected to experience relatively stable employment.

0.155. At conventional significance levels, one could not reject the null hypothesis that ability was uniformly distributed.

The estimated impact on match productivity of college completion (h_2) is 13.4 percent, and this productivity enhancement is estimated reasonably precisely. Of course, not all of this enhancement passes through to the worker, with the point estimate of the bargaining power equal to 0.571. As we have noted previously, this value is heavily influenced by the labor share value we use in the estimation, which is equal to 0.67. The point estimate of a^* is equal to 1.102. Given the point estimates of F_a and a^*, the predicted probability of college completion is 0.450, while the sample value is 0.330. This poor fit is most likely due to the fact that a serves dual roles in the model. One is as an important determinant of the wage distribution, and the other is as the only determinant of the schooling decision (at the individual level). In the next model, without individual heterogeneity in a, the model fit on the school dimension will be "perfect," by construction, with little loss in the ability to fit the schooling-specific wage distributions.

5.2 Estimates of the Homogeneous Ability Model

The estimates of the model in which schooling costs are heterogeneous in the population while $a_i = 1 \; \forall \; i$ are given in Table 3. Once again, θ^* is estimated to be 6.000, though in this case the estimated value of h_2 is $\nabla \min w = 1.25$, since the shift in the wage distribution is attributed solely to human capital improvements instead of being a "mixture" of these improvements and composition effects.

The rate parameters λ and η are estimated from essentially the same sample information as was the case in the heterogeneous ability model, so that the point estimates of these parameters change little. The estimated bargaining power parameter also changes little from the previous specification, in large part due to the labor share information having such a dominant influence on the estimate.

The biggest changes relate to the estimates of the match and schooling cost distributions. Of course, the match distribution now has to supply all of the heterogeneity to generate the observed wage distributions, so in principle the fit of these distributions could be substantially worse than was the case when ability heterogeneity was also present. We see from Figure 2 that there is not a substantial cost to pay in terms of fitting the cross-sectional wage distributions by restricting ability to be homogeneous. The top panel contains the distribution of wages for individuals who have not completed a college education and the fitted wage distribution. The fit is extremely good for such a parsimonious model. The bottom panel contains the wage distribution and the fitted one for the college-educated subsample. Here the fit is not as good, though we still might deem it reasonably acceptable given the simplicity of the model structure. The fitted distribution places too much mass in the left tail of the distribution. Allowing a limited amount of heterogeneity in ability has the potential of shifting some of this mass to the right, though if inappropriate distributional assumptions are made there is the downside risk of worsening the fit in terms of its generic "shape."

The main gain in this specification is the ability to fit the educational decision precisely, since the distribution of the costs of schooling only determine this decision, and the single parameter of this distribution is estimated from the sample proportion of college completers. Though the units are largely meaningless, the point estimate of the parameter of the schooling cost distribution is $\hat{\xi} = 0.517$, and the estimated critical cost value is 0.775. These estimates must imply the sample college completion rate of .330. In assessing the impact of bargaining power on the schooling decision, we use estimates from this model in large part because of its (built-in) consistency with schooling investment choices.

5.3 The Impact of Bargaining Power on Educational Investment

In order to look at the impact of changes in bargaining power on schooling investment, we first must use the estimates reported in Table 3 to retrieve the remaining primitive parameter required to solve the educational choice problem, b_0. We estimate b_0 given the point estimates of all of the primitive parameters (and the assumed value of the discount rate, ρ) as

$$\hat{b}_0 = \hat{\theta}^* - \frac{\hat{\lambda}\hat{\alpha}}{\rho + \hat{\eta}} \int_{\hat{\theta}^*} (\theta - \hat{\theta}^*) d\hat{G}(\theta).$$

Given consistency of these parameter estimates, our estimator for b_0 is consistent as well.

Given the estimates of all of the primitive parameters required to solve for the decisions in the labor market and schooling choice, we vary the bargaining power parameter from 0.01 to 0.99. For any given "counterfactual" value of bargaining power, α' say, we compute $\theta^*(\alpha')$, which is then used to compute $\tilde{c}^*(\alpha')$. Using this decision rule and the estimated schooling cost distribution, we then have

$$P_2(\alpha') = 1 - \exp(-\hat{\xi}\tilde{c}^*(\alpha')).$$

The results of the exercise are presented in Figure 3. They show that, under our model assumptions, the college completion rate is very sensitive to the bargaining power parameter. In particular, we see that when bargaining power drops below 0.42, no one completes college. By the structure of the estimator, we know that at $\alpha = 0.565$, 33 percent of the population completes college (the sample proportion). To obtain a college completion rate of 50 percent, the bargaining power of workers would have to increase to approximately 0.78. Given the structure of the cost distribution, even as $\alpha \to 1$ not all individuals will complete college. This seems to be a reasonable implication.

We have only conducted this comparative statics exercise for the parameter α. Of course, all of the parameters of the model will impact the college completion rate through their impact on $\tilde{c}^*$, as we saw when deriving comparative statics results from the model. The model estimates can be used to determine the quantitative impact of changes in the contact rate, the distribution G, etc., in an obvious manner.

6 Two-Sided Investment

As we have just seen, the proportion of individuals investing in a college education is a function of all labor market parameters, in particular the bargaining power parameter α. No matter what the values of the other parameters (given that there is some positive gain to employment), investment in college education will be a monotonically increasing function of α.This leads to the conclusion that, in the absence of the possibility of firm investment, efficiency would require that $\alpha = 1$.[12]

Of course, this implication will only hold, in general, as long as firms do not have the opportunity to invest themselves. This is clearly not the case. Just as individuals can make investments in general education, firms can invest in infrastructure, technology, and other types of capital goods to make any match generically more productive. In such a case, it will not generally be efficient to assign the entire surplus of the match to either party. We sketch the manner in which an optimal bargaining power parameter would be determined in such a case.

We return to the original specification of match productivity, that is,

$$\begin{aligned} y_{ij} &= a_i h_i \theta_{ij} p_j k_j \\ &= \tilde{a}_i \theta_{ij} \tilde{p}_j, \end{aligned}$$

with $\tilde{p}_j = p_j k_j$. The most significant modeling change from the previous sections involves the heterogeneity among firms in their productive contributions. Our earlier analysis produced heterogeneity in the value of unemployed search of individual labor supplies, denoted $V_U(\tilde{a})$. Firms instead occupy two states, one in which they are "unemployed" in the sense of holding a vacant position, and the other when the position is occupied, where the value of a vacancy to a type $\tilde{p}$ firm is denoted $V_V(\tilde{p})$. We attempt to make the problem facing the firm and individual as symmetric as possible, so that we assume the flow value to a firm of type $\tilde{p}$ when it is holding a vacant position is given by $l_0\tilde{p}$. Now the Nash bargaining problem is stated as

$$\max_w (V_E(w, \tilde{a}) - V_U(\tilde{a}))^{\alpha} (V_F(y, w) - V_V(\tilde{p}))^{1-\alpha},$$

which becomes

$$(\rho + \eta)^{-1} \max_w (w + \eta V_U(\tilde{a}) - (\rho + \eta) V_U(\tilde{a}))^{\alpha} (y - w + \eta V_V(\tilde{p}) - (\rho + \eta) V_V(\tilde{p}))^{1-\alpha}.$$

After some simplification, the wage equation is given by

$$w^*(\tilde{a}, \tilde{p}, \theta) = \alpha(\tilde{a}\theta\tilde{p} - \rho V_V(\tilde{p})) + (1 - \alpha)\rho V_U(\tilde{a}). \tag{15}$$

[12] This point was originally raised in Fabien Postel-Vinay's discussion of this paper.

Using this expression for w^*, the values of unemployed search and unfilled positions are given by

$$\rho V_U(\tilde{a}) = b_0\tilde{a} + \frac{\alpha\lambda}{\rho+\eta}\int\int_{\theta^*(\tilde{a},z)}\{\tilde{a}\theta z - \rho V_U(\tilde{a}) - \rho V_V(z)\}dG(\theta)dF_{\tilde{P}}(z) \tag{16}$$

$$\rho V_V(\tilde{p}) = l_0\tilde{p} + \frac{(1-\alpha)\lambda}{\rho+\eta}\int\int_{\theta^*(z,\tilde{p})}\{z\theta\tilde{p} - \rho V_U(z) - \rho V_V(\tilde{p})\}dG(\theta)dF_{\tilde{A}}(z), \tag{17}$$

respectively. The lower limit of integration with respect to θ in both integrals has been denoted $\theta^*(\tilde{a},\tilde{p})$. A match is formed whenever the flow productivity exceeds the sum of the outside options of the worker and the firm, or

$$\begin{aligned} y &\geq \rho V_U(\tilde{a}) + \rho V_V(\tilde{p}) \\ \Rightarrow \tilde{a}\theta\tilde{p} &\geq \rho V_U(\tilde{a}) + \rho V_V(\tilde{p}), \end{aligned}$$

so that

$$\theta^*(\tilde{a},\tilde{p}) = \frac{\rho V_U(\tilde{a}) + \rho V_V(\tilde{p})}{\tilde{a}\tilde{p}}.$$

Let $X(\tilde{a},\tilde{p}) = \rho V_U(\tilde{a}) + \rho V_V(\tilde{p})$. Then sum (16) and (17) to get

$$\begin{aligned} X(\tilde{a},\tilde{p}) &= b_0\tilde{a} + \tau_0\tilde{p} + \frac{\alpha\lambda}{\rho+\eta}\int\int_{\frac{X(\tilde{a},z)}{\tilde{a}z}}\{\tilde{a}\theta z - X(\tilde{a},z)\}dG(\theta)dF_{\tilde{P}}(z) \\ &\quad + \frac{(1-\alpha)\lambda}{\rho+\eta}\int\int_{\frac{X(z,\tilde{p})}{z\tilde{p}}}\{z\theta\tilde{p} - X(z,\tilde{p})\}dG(\theta)dF_{\tilde{A}}(z). \end{aligned}$$

Given a unique solution for $X(\tilde{a},\tilde{p})$ it is straightforward to determine the individual valuations for the two sides of the market, since

$$\begin{aligned} \rho V_U(\tilde{a}) &= b_0\tilde{a} + \frac{\alpha\lambda}{\rho+\eta}\int\int_{\frac{X(\tilde{a},z)}{\tilde{a}z}}\{\tilde{a}\theta z - X(\tilde{a},z)\}dG(\theta)dF_{\tilde{P}}(z) \\ \rho V_V(\tilde{p}) &= \tau_0\tilde{p} + \frac{(1-\alpha)\lambda}{\rho+\eta}\int\int_{\frac{X(z,\tilde{p})}{z\tilde{p}}}\{z\theta\tilde{p} - X(z,\tilde{p})\}dG(\theta)dF_{\tilde{A}}(z). \end{aligned}$$

Given these values, the wage function (15) follows immediately.

By the structure of the model, the value of investment is increasing in endowments for each side of the market, so that there exist conditional investment rules on each side of the market that possess a critical value property, with these critical values given by

$$\begin{aligned} a^* &= a^*(p^*,\alpha,\Omega) \\ p^* &= p^*(a^*,\alpha,\Omega), \end{aligned}$$

where Ω is the vector containing all of the primitive parameters of the model with the exception of the bargaining power parameter α. A Nash equilibrium is characterized by

$$\begin{aligned} a^*_{NE} &= a^*(p^*_{NE}, \alpha, \Omega) \\ p^*_{NE} &= p^*(a^*_{NE}, \alpha, \Omega). \end{aligned}$$

For purposes of discussion, we assume that there exists such an equilibrium though we do not assume uniqueness. The sum of worker and firm expected values of market participation is given by $X(\tilde{a}, \tilde{p})$, so an efficient equilibrium corresponds to

$$\max_{\alpha} \int \int X(\tilde{a}(a, a^*(p^*_{NE}, \alpha, \Omega)), \tilde{p}(p, p^*(a^*_{NE}, \alpha, \Omega))) dF_P(p) dF_A(a). \tag{18}$$

This way of selecting a value of α is reminiscent of the Hosios condition (1990), though given the simplicity of the matching function formulation, that condition can be represented in a much more elegant and transparent manner. Given the matching function set-up, the Hosios condition proposes rewarding each side of the market (in terms of α) as a function of their contribution to match formation. In our case, we could think of the parties being rewarded as a function of the productivity of their investments. For example, if $k_1 = 1$ (as $h_1 = 1$), $b_0 = \tau_0$, and if $F_P(p) = F_A(a)$, the worker should be rewarded with a value of $\alpha > 0.5$ if $h_2 > k_2$. That is, the side for which investment is most valuable to the match should be given the greatest encouragement to invest. In fact, when the firm was incapable of making investments, it was optimal to set $\alpha = 1$.

While we have only sketched an argument in support of a social planner's solution to the surplus division problem that would yield an optimal value of α, much remains to be done in terms of proving existence and (possibly) uniqueness of investment strategies and in terms of characterizing the solution to (18). However, it seems a promising avenue to pursue since productivity-enhancing investments are potentially more amenable to measurement than is the degree of search effort on the supply side or vacancy creation on the demand side of the market.[13]

7 Conclusion

In this paper we have developed a labor market model of search, matching, and bargaining that allows for pre-entry productivity enhancing investments by workers and firms. Given data limitations, the bulk of the discussion was devoted to the case in which only workers were able to make such investments, and their investment choices were restricted to be binary: college completion or not. The model estimates allowed us to examine the impact of hold-up (which was generated by search frictions) on the investment choices of agents, and in particular we focused on the impact of the Nash bargaining power parameter on

[13] For some of the difficulties of measuring vacancies, see Petrongolo and Pissarides (2001).

college completion. We found that the investment decision of workers was highly sensitive to the bargaining power parameter value. At values less than 0.42, our prediction was that no one would have completed college. A value of 0.78, on the other hand, would have resulted in a 50 percent college completion rate, which is to be contrasted with the 33 percent completion rate in the data.

By analyzing the schooling investment decision within such a highly structured and stylized model, we were able to derive a number of comparative statics results from the theory. These strong assumptions are quite restrictive when taking the model to data. In particular, the preeminent role of ability as a determinant of wages (particularly over the life cycle) and in the schooling decision makes it difficult to simultaneously fit both. Moreover, the fact that the wage distribution must be "deconvoluted" to determine the ability and matching distributions adds a further complication to model identification. This led us to estimate a simpler version of the model in which individuals all had the same ability endowment but differed in their costs of schooling investment. For the purposes of conducting an initial quantitative assessment of the sensitivity of schooling investment on bargaining power given the value of the other estimates of primitive parameters, we found this specification adequate. However, ruling out initial ability differences did not permit us to demonstrate empirically our point that the distribution of individual (or firm) productivities is endogenous in a general setting. Of course, it is still the case that the schooling distribution is endogenous, so that changes in primitive parameters or economic policy will change the schooling type distribution as well as labor market outcomes conditional on schooling type.

Given the difficulty of identifying the model parameters even in the homogeneous labor markets case, we were not able to estimate the heterogeneous labor markets model. This is a future research goal, since some of the descriptive evidence (e.g., unemployment rates by schooling level) points to the inadequacy of the homogeneous markets assumption. One particularly interesting feature of the heterogeneous labor markets assumption is that college completion rates may be high even if bargaining power among college graduates (α_2) is low when bargaining power among those with low schooling is even lower. Then the schooling decision is influenced also by the relative amounts of bargaining in the different labor markets. Pursuing the goal of estimating such a model will probably entail using matched employer-employee data. This will also have the advantage of investigating the two-sided investment problem in a more satisfactory manner at the theoretical and empirical level.

While the models estimated are very stylized, we think that there are potentially important policy lessons to be learned from the approach adopted here. In recent years we have seen large changes in the unemployment rates among all demographic groups. To the extent that these reflect long-lasting declines in arrival rates of job offers or increases in job destruction rates, there may be a significant impact on the value of schooling and the resulting schooling distribution. It is appears that a model with search unemployment is necessary to examine these impacts.

Table 1

Descriptive Statistics

(Entire Sample and by Schooling Level)

Characteristic	All	Schooling Level H	C
$P(s = C)$	0.330	0.000	1.000
$P(U)$	0.033	0.041	0.017
$E(t_U \| U)$	5.122	5.270	4.412
$SD(t_U \| U)$	6.249	6.254	6.233
$\min\{w_i\}$	6.000*	6.000*	7.500*
$E(w \| E)$	18.384*	15.128*	24.833*
$SD(w \| E)$	9.524*	5.981*	11.704*
N	9985		

Supplemental Data

$E(\ln w)$	0.013
$E(\ln w)^2$	0.191
$E(\ln w)^4$	0.143
N	788

Note: * denotes that the sample statistic was computed after trimming the top and bottom 2.5 percent from the conditional (on schooling) wage distributions for the cross-sectional CPS-ORG data

Table 2

Maximum Likelihood Estimates of Model with Ability Heterogeneity

(Asymptotic Standard Errors)

Parameter	*Estimate*
λ	0.195
	(0.012)
η	0.006
	(0.001)
μ_θ	2.765
	(0.103)
σ_θ	0.488
	(0.096)
γ	0.812
	(0.155)
m	0.213
	(0.101)
h_2	1.134
	(0.021)
α	0.565
	-
a^*	1.102
	(0.038)
θ^*	6.000
	-
$\overline{\ln L}$	-3.572

Table 3

Maximum Likelihood Estimates of Model with No Ability Heterogeneity

(Asymptotic Standard Errors)

Parameter	*Estimate*
λ	0.197
	(0.011)
η	0.007
	(0.001)
μ_θ	3.065
	(0.005)
σ_θ	0.549
	(0.005)
h_2	1.250
	-
α	0.565
	-
ξ	0.517
	(0.058)
$\tilde{c}^*$	0.775
	-
θ^*	6.000
	-
b	-179.348
$\overline{\ln L}$	-3.318

References

[1] Acemoglu, D. 1996. “A Microfoundation for Social Increasing Returns in Human Capital,” *Quarterly Journal of Economics* 111: 779–804.

[2] Acemoglu, D. 1997. “Training and Innovation in an Imperfect Labor Market,” *Review of Economic Studies* 64: 445–64.

[3] Acemoglu, D. and R. Shimer. 1999. “Holdups and Inefficiencies with Search Frictions,” *International Economic Review* 40: 827-849.

[4] Bound, J., C. Brown, G. Duncan, and W. Rodgers. 1994. “Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data, *Journal of Labor Economics* 12: 345-368.

[5] Cahuc, P., F. Postel-Vinay, and J.-M. Robin. 2006. “Wage Bargaining with On-the-Job Search: Theory and Evidence.” *Econometrica* 74: 323-64.

[6] Chaippori, P., Iyigun, M. and Weiss, Y. 2009. “Investment in Schooling and the Marriage Market, *American Economic Review*, 99: 1689-1713.

[7] Dey, M. and C. Flinn. "An Equilibrium Model of Health Insurance Provision and Wage Determination,” *Econometrica* 73 (March 2005): 571-627.

[8] Eckstein, Z., and K. Wolpin. 1995. “Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model,” *Review of Economic Studies* 62: 263-286.

[9] Flinn, C. 2006. “Minimum Wage Effects on Labor Market Outcomes under Search, Bargaining, and Endogenous Contact Rates.” *Econometrica* 74: 1013-1062.

[10] Flinn, C., and J. Heckman. 1982. “New Methods for Analyzing Structural Models of Labor Force Dynamics.” *Journal of Econometrics* 18: 115-68.

[11] Heckman, J. and B. Singer. 1984. “A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data,” *Econometrica* 52: 271-320.

[12] Heckman, J., L. Lochner, and C. Taber. 1998. “Explaining Rising Wage Inequality: Explanations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents,” *Review of Economic Dynamics* 1: 1-58.

[13] Hosios, A. 1990. “On the Efficiency of Matching and Related Models of Search and Unemployment,” *Review of Economic Studies* 57: 279-298.

[14] Keane, M and K. Wolpin. 1997. “The Career Decisions of Young Men," *Journal of Political Economy* 105: 473-522.

[15] Krueger, A. 1999. “Measuring Labor’s Share,” *American Economic Review* 89(2): 45-51.

[16] Lee, D. 2005. “An Estimable Dynamic General Equilibrium Model of Work, Schooling and Occupational Choice,” *International Economic Review*, 46: 1-34.

[17] Lee, D and K. Wolpin. 2006. “Intersectoral Labor Mobility and the Growth of Service Sector,” *Econometrica* 47: 1-46.

[18] Malcomson, J. 1997. “Contracts, Hold-Up, and Labor Markets,” *Journal of Economic Literature* 35: 1916-1957.

[19] Petrongolo, B. and C. Pissarides. 2001. “Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature* Vol. XXXIX: pp. 390–431.

[20] Pissarides, C. 2000. *Equilibrium Unemployment Theory,* 2nd edition. Cambridge, MA: MIT Press.

[21] Postel-Vinay, F., and J-M Robin. 2002. “Equilibrium Wage Dispersion with Heterogeneous Workers and Firms.” *Econometrica* 70: 2295-1350.

[22] Shimer, R., R. Rogerson and R. Wright. 2005. “Search Theoretic Models of the Labor Market: A Survey,” *Journal of Economic Literature*, 43: 959-988.

[23] Sullivan, P. 2010. “A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search,” *International Economic Review* 51: 289-317.

Figure 1.a
Wage Distribution
Low Schooling

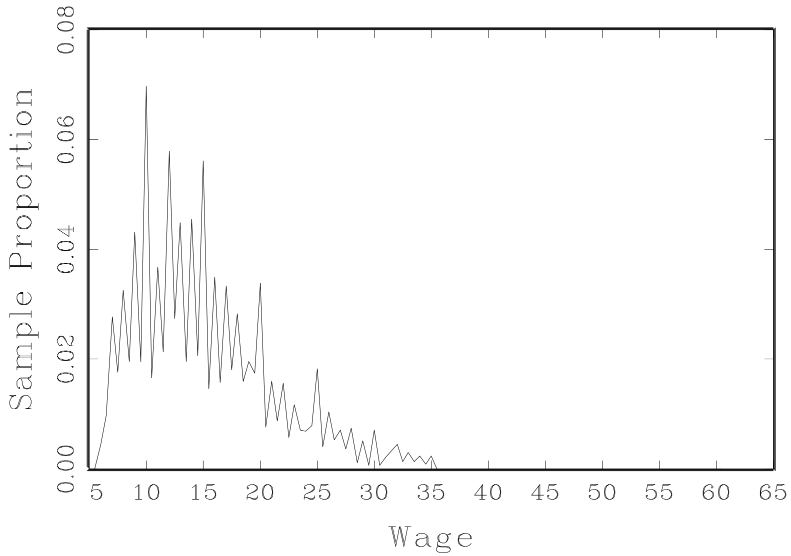

Figure 1.b
Wage Distribution
High Schooling

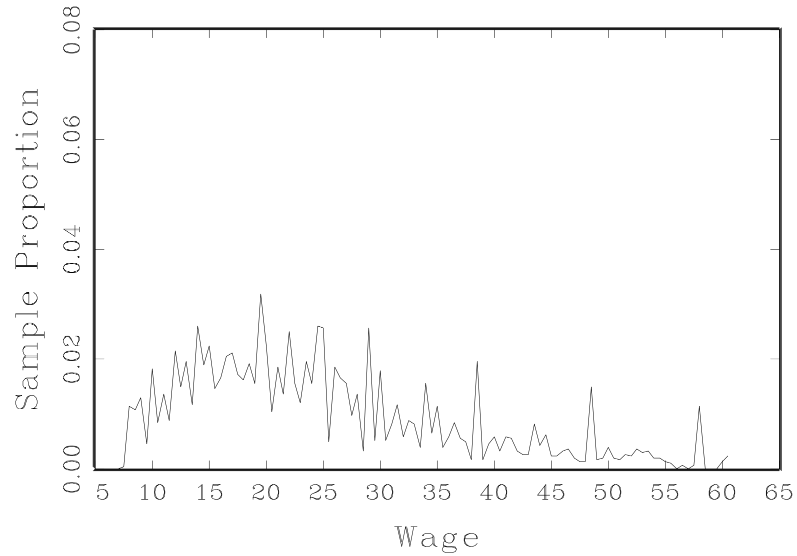

Figure 1.c
Wage Distribution
Total Sample

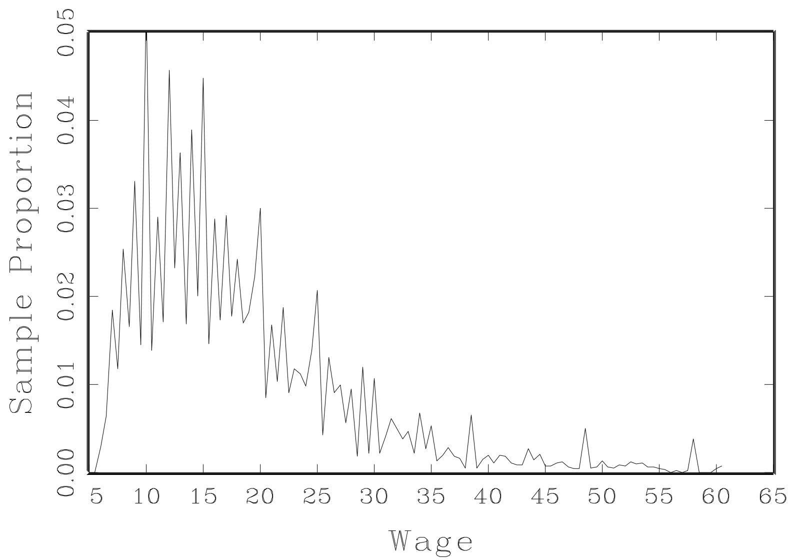

Figure 1.d
Difference in Conditional C.D.F.s
Low Schooling − High Schooling

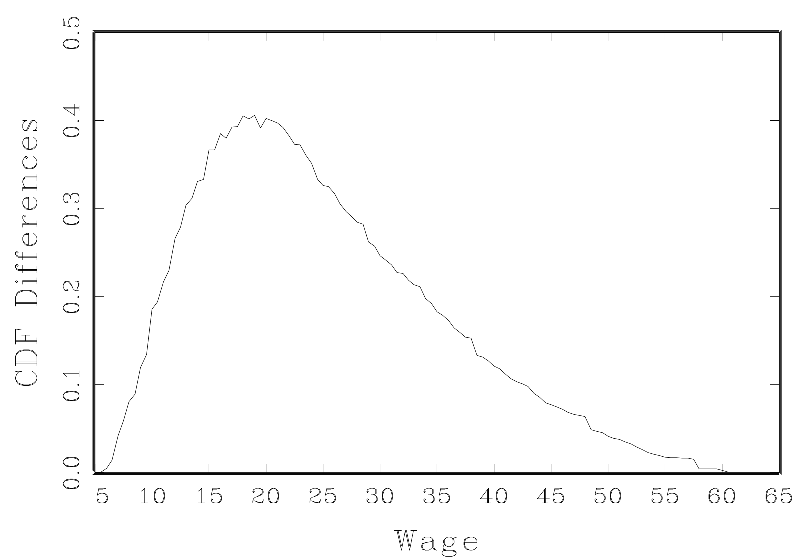

Figure 2.a
Actual and Predicted Non-College Wages
No Heterogeneity in a

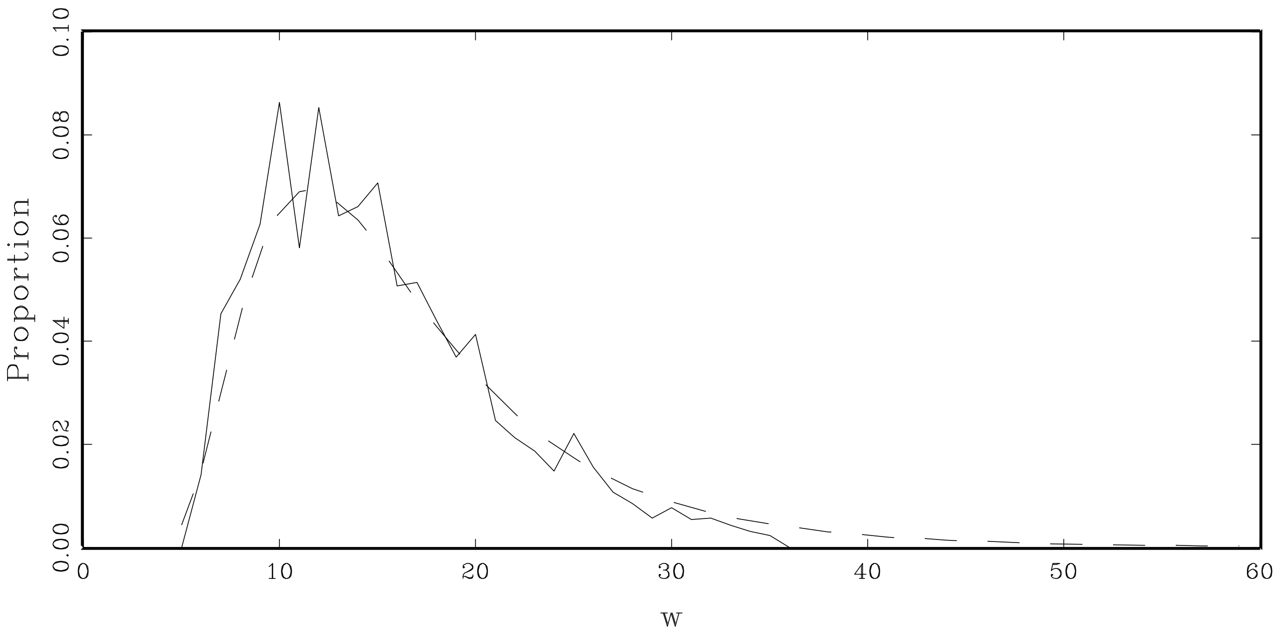

Figure 2.b
Actual and Predicted College Wages
No Heterogeneity in a

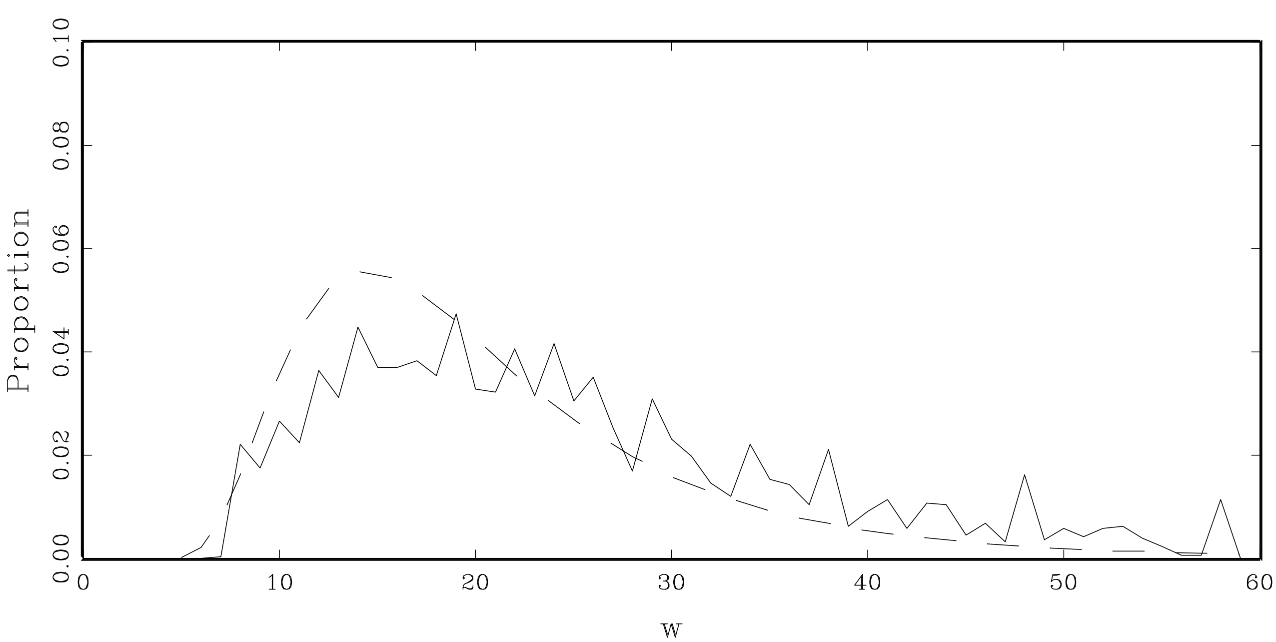

Figure 3
College Completion as Function of α