

Collegio Carlo Alberto



Labor Market Search and Schooling Investment

Christopher Flinn

Joseph Mullins

No. 295

March 2013

Carlo Alberto Notebooks

www.carloalberto.org/research/working-papers

Labor Market Search and Schooling Investment¹

Christopher Flinn
Department of Economics
New York University

Joseph Mullins
Department of Economics
New York University

Collegio Carlo Alberto
Moncalieri, Italy

First Draft: June 2010
This Draft: February 2013

Keywords: Labor market search; schooling choice; hold-up; Nash bargaining.

JEL Classifications: J24, J3, J64

¹The 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. We are 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. We are responsible for all errors, omissions, and interpretations.

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 in the partial equilibrium version of the model. However, large increases in bargaining power in the general equilibrium version of the model choke off firm vacancy creation and actually reduce the level of schooling investment.

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, and obtain estimates of both partial and general equilibrium versions of the model.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 (OTJ) search. We discuss some of the possible modifications of our findings that would occur if OTJ search was considered.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 (1996,1997) 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

¹A 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 (1995), Postel-Vinay and Robin (2002), Dey and Flinn (2005), Cahuc et al (2006), and Flinn (2006).

²There 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 (1999) 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.

³Adding 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.

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 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 framework. 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 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 in the partial equilibrium version of the model and that the amount of schooling is monotone increasing in bargaining power, as theory leads us to expect. However, in the general equilibrium version of the model,

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

large increases in the bargaining power of workers lead to decreases in schooling attainment and the welfare of workers as firms dramatically reduce the number of vacancies created. In the econometric section, we discuss why it is not possible to estimate the parameters of the Mortensen-Pissarides matching function given the data at our disposal. However, our results hold for a variety of “reasonable” values of the matching function parameters as reported in the Petrongolo and Pissarides (2001) survey of the empirical literature on the estimation of matching functions using aggregate data.

Our work is closest in spirit to that of Charlot and Decreuse (2005,2010), who perform a theoretical analysis of the hold-up problem in a search and matching framework that is a special case of the one considered here. In their model, productivity is solely a function of individual innate ability and their schooling level, whereas our model allows for match heterogeneity that we find to be empirically very important.⁵ The interesting theoretical point that they make is that the schooling decisions of individuals do not lead to efficient outcomes for the economy since individuals neglect the impact of their schooling decision on firms’ vacancy creation. As in our model, there is a critical ability level a^* such that all individuals with ability a less than a^* do not acquire higher education. When an individual near the margin goes to school, he lowers the quality level in each schooling sub-market, thereby reducing the incentives for firms to post vacancies in either market. This type of effect will also exist in our framework, though it will be muted due to the presence of match-specific heterogeneity, the variance of which dominates that of individual ability.

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 sub-markets to be characterized by different vectors of primitive parameters, such as contact and dissolution rates. In Section 4 we analyze the general equilibrium case in which contact rates are determined by the measure of unemployed searchers and the vacancy creation decisions of firms. In Section 5 we describe the sample used to estimate the model and discuss identification of model parameters and the estimator used. Section 6 presents model estimates and (empirical) comparative statics exercises. 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.

⁵Our model also differs in that we allow for different values of offer rates, dismissal rates, and bargaining power across the two schooling markets.

The flow (in continuous time) productivity of a match between worker i and firm j is assumed to be given by

$$y_{ij} = \tilde{a}_i \theta_{ij} \tilde{p}_j, \quad (1)$$

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 an important 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.⁶ 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 acquire a higher level of formal education. 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 $h_2 > h_1$. Similar possibilities exist on the

⁶There 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 our knowledge, no labor market search model with bargaining that includes pre-market investments has been estimated.

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. 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 throughout the paper we will assume that firms are homogeneous, i.e., $\tilde{p}_j = 1 \forall j$. In this case, the output at a match is given by

$$y = ah_s\theta,$$

where θ is i.i.d. with c.d.f. G , h_s is the individual’s (schooling) human capital level, and a is individual ability, which is a permanent draw from the distribution F_a with corresponding density f_a . Hereafter we drop the individual and firm subscripts - they are redundant since a and h refer to individuals 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.⁷

The analysis, both 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

$$b(a, s) = b_0 ah_s.$$

⁷This 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, 30 percent of our sample of 25-34 year old males fell into schooling class $s = 2$.

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

Since we primarily use data from the Current Population Survey, which is a point sample of the labor market process of household members, we assume no on-the-job (OTJ) search. This assumption is not innocuous for some of the comparative statics exercises we conduct below, and in the conclusion of the paper we discuss how allowing OTJ search could affect our results.

We begin by considering the case in which, across schooling “sub-markets,” 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) who is entering the labor market as an unemployed searcher, can be summarized solely in terms of the product $\tilde{a} \equiv ah_s$, and the value of unemployed search to such an individual is given by $V_U(\tilde{a})$. In terms of the Nash bargaining problem, the worker-firm pair solves

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

where

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

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 0.⁸ The solution to the Nash bargaining problem yields

$$w(\theta, \tilde{a}) = \alpha \theta \tilde{a} + (1 - \alpha) \rho V_U(\tilde{a}),$$

and since

$$\rho V_U(\tilde{a}) \equiv y^*(\tilde{a}) = \tilde{a} \theta^*(\tilde{a}),$$

we have

$$w = \tilde{a}(\alpha \theta + (1 - \alpha) \theta^*(\tilde{a})). \quad (2)$$

In terms of the value of unemployed search given \tilde{a} , we have

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

⁸This assumption is utilized below when we generalize the model to allow for endogenous vacancy creation by firms.

Since this last equation is independent of \tilde{a} , we have

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

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

$$y^*(a, s) = ah_s\theta^*. \quad (4)$$

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

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

We consider schooling level 1 as the baseline, that is, it is the required level of schooling for all individuals. To complete schooling level 2, instead, a total cost of c must be incurred, that is assumed to be the same for all individuals. Individuals base their decision to complete schooling level 2 on the comparison between the values of entering the labor market with human capital h_2 or h_1 . Then an individual with ability level a will attend school if

$$\begin{aligned} \rho^{-1}ah_2\theta^* - \rho^{-1}ah_1\theta^* &> c \\ \Rightarrow a &> \frac{\rho c}{(h_2 - h_1)\theta^*}. \end{aligned}$$

Then, given that $c > 0$ and $\theta^* > 0$, there exists a critical value a^* defined as

$$a^* = \frac{\rho c}{(h_2 - h_1)\theta^*},$$

with an individual of type $a \geq a^*$ acquiring schooling and a person $a < a^*$ not acquiring further schooling.⁹ In the empirical analysis conducted below, we assume that the support

⁹In an earlier version of the paper, we explicitly accounted for the difference in labor market entry dates under the two schooling options as part of the model. In such a case, with instantaneous cost \tilde{c} , the total discounted value of continuing in higher education is $\exp(-\rho\tau)\rho^{-1}ah_2\theta^*$, where τ is the length of time required to complete advanced schooling. The total cost absorbed over the schooling period is

$$c = \rho^{-1}\tilde{c}(1 - \exp(-\rho\tau)).$$

Then the individual acquires schooling level 2 if

$$\exp(-\rho\tau)\rho^{-1}ah_2\theta^* - \rho^{-1}ah_1\theta^* > c,$$

and the condition for schooling to be acquired by any a when $c > 0$ is that $\exp(-\rho\tau)h_2 - 1 > 0$. Given that ρ and τ are fixed in our analysis, this puts a strong restriction on the estimate of h_2 for their to be a critical value rule of the type we describe when explicitly considering discounting phenomena. We note that Cherlot and Decreuse (2005,2010) also ignore discounting when comparing the values of the two schooling choices in their theoretical analysis of schooling investment in the presence of search frictions and bargaining.

of the distribution of a is R_+ , so that the proportion of individuals in the population who acquire schooling is given by $1 - F_a(a^*) > 0$.

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. $\partial P_2 / \partial c < 0$. The proportion of the population attending college is decreasing in the direct costs of college attendance.
2. $\partial P_2 / \partial h_2 > 0$. This is perhaps the most intuitive result. The greater the impact on labor market productivity, the greater the number of individuals who complete college.
3. $\partial P_2 / \partial \theta^* > 0$. 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) $\partial P_2 / \partial \lambda > 0$. An increase in the arrival rate of offers increases θ^* , and hence increases the value of having a higher productivity distribution.
 - (b) $\partial P_2 / \partial \eta < 0$. An Increase in the (exogenous) separation rate decreases θ^* and hence decreases the value of becoming more productive when matched with an employer.

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 1 *Increases in bargaining power on the workers' side of the market result in increases in schooling level, or*

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

We should note that this result is only true for the case in which the contact rates between searchers and firms are held fixed, i.e., only in partial equilibrium. As we will see below, this result will not in general be true in the general equilibrium version of the model.

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, we look at 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 \tilde{a} is independent of \tilde{a} . This is due to the fact that the likelihood that any job is acceptable to an individual of type \tilde{a} is simply $\tilde{G}(\theta^*)$, which is independent of \tilde{a} . The proportion of time an individual of type \tilde{a} spends in unemployment, or the steady state probability that they will occupy the unemployment state, is simply

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

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 $[a^*, \infty)$. Now, from (2) we know that

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

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

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

and the corresponding conditional wage density is given by

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

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

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

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

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

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

$$f_{w|s}(w|s=1) = \frac{1}{\alpha \tilde{G}(\theta^*)} \int_{\underline{a}}^{a^*} a^{-1} g(\alpha^{-1}(\frac{w}{a} - (1-\alpha)\theta^*)) \frac{dF(a)}{F(a^*)}, \quad w \geq \underline{a}\theta^*, \quad (5)$$

where \underline{a} is the lowest value of a in the population, while the density of wages among the college-educated population is

$$f_{w|s}(w|s=2) = \frac{1}{\alpha h_2 \tilde{G}(\theta^*)} \int_{a^*}^{\bar{a}} a^{-1} g(\alpha^{-1}(\frac{w}{ah_2} - (1-\alpha)\theta^*)) \frac{dF(a)}{\tilde{F}(a^*)}, \quad w \geq a^* h_2 \theta^*. \quad (6)$$

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 2 *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 empirical 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 present some descriptive evidence regarding specific empirical implications. 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. We restrict the sample to males between the ages of 25 and 34, inclusive. Further details of the sample selection used are provided in section 5.1. 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.

Figure 1.a and 1.b contain the plots of the wage distributions by schooling group. 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 30 percent 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, $w_{(1)} < \dots < w_{(K)}$. First order stochastic dominance 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 2.16%, while among those with less schooling it is 4.67%. 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 $y(a, s, \theta) = ah_s\theta = \tilde{a}\theta$, so that the distribution of y is a function of the scalar \tilde{a} 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, \tilde{a} is no longer a sufficient statistic for the value of search of an individual; instead a minimal sufficient statistic is the pair (\tilde{a}, 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

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

The solution $\theta^*(\tilde{a}, s)$ now clearly is independent of \tilde{a} , as before, 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 $ah_s\theta^*(s) = \tilde{a}\theta^*(s)$, so that the value of unemployed search in this sub-market is given by $\rho^{-1}\tilde{a}\theta^*(s)$. Then the net value of college education to an individual of type a is

$$\rho^{-1}ah_2\theta^*(2) - c - \rho^{-1}a\theta^*(1),$$

so that the critical ability level a^* is given by

$$a^* = \frac{c\rho}{h_2\theta_2^* - \theta_1^*}.$$

For a^* to be positive, it is necessary that $h_2\theta_2^* - \theta_1^*$. This condition is satisfied in the empirical work conducted below.

3.1 Comparative Statics Results

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 sub-market. By simple extension of the homogeneous results above, the results regarding $\partial P_2/\partial c < 0$ 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 θ_1^* and θ_2^* . We note that

1. $\partial P_2/\partial \theta_1^* < 0$. As before, θ_1^* is not a primitive parameter, but the primitive parameters specific to sub-market 1 only affect the schooling decision through θ_1^* . Then
 - (a) $\partial P_2/\partial \lambda_1 < 0$. Increases in the arrival rate of offers in the low-schooling market increase θ_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) $\partial P_2/\partial \lambda_2 > 0$
 - (b) $\partial P_2/\partial \eta_2 < 0$
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

$$\frac{\partial P_2}{\partial \alpha_1} < 0.$$

$$\frac{\partial P_2}{\partial \alpha_2} > 0.$$

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 \tilde{a} is no longer sufficient for describing an individual's probability of labor market events, in general, except when \tilde{a} 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 \tilde{a} 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

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

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

$$f_{w|a,1}(w|a, 1) = \frac{1}{\alpha_1 a} \frac{g(\alpha_1^{-1}(\frac{w}{a} - (1 - \alpha_1)\theta_1^*))}{\tilde{G}(\theta_1^*)}, \quad w \geq \underline{a}\theta_1^*,$$

while the wage density for the high schooling group is

$$f_{w|a,2}(w|a, 2) = \frac{1}{\alpha_2 h_2 a} \frac{g(\alpha_2^{-1}(\frac{w}{ah_2} - (1 - \alpha_2)\theta_2^*))}{\tilde{G}(\theta_2^*)}, \quad w \geq a^*h_2\theta_2^*.$$

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

$$f_{w|s}(w|s = 1) = \frac{1}{\alpha_1 \tilde{G}(\theta_1^*)} \int_{\underline{a}}^{a^*} a^{-1} g(\alpha_1^{-1}(\frac{w}{a} - (1 - \alpha_1)\theta_1^*)) \frac{dF(a)}{F(a^*)}, \quad w \geq \underline{a}\theta_1^*,$$

and

$$f_{w|s}(w|s = 2) = \frac{1}{\alpha_2 h_2 \tilde{G}(\theta_2^*)} \int_{a^*}^{\infty} a^{-1} g(\alpha_2^{-1}(\frac{w}{ah_2} - (1 - \alpha_2)\theta_2^*)) \frac{dF(a)}{\tilde{F}(a^*)}, \quad w \geq a^*h_2\theta_2^*.$$

4 Endogenous Contact Rates

Up until this point, we have considered that the contact rates in the model, the common λ in the homogeneous market version and the pair (λ_1, λ_2) in the heterogeneous case, as constants. On theoretical grounds, this assumption has important implications in terms of market efficiency. If there exists a planner with the ability to determine the proportion of the surplus given to worker and the firm, then setting $\alpha = 1$ is efficient. This is due to our assumption that only the workers make pre-market entry investments that impact

post-market entry productivity realizations. With fixed contact rates, there is no reason, from the viewpoint of efficiency, to choose any value other than 1. In this section we look at the reasons why $\alpha = 1$ will not be the efficient surplus division parameter in the general equilibrium case.

4.1 The Homogeneous Case

In this section we consider the vacancy creation decision of firms when all agents on the supply side face the same set of environment parameters, which are: G, λ, η . If we are to have one contact rate facing all individuals independent of their schooling level, this means that employers are not able to direct their vacancies to particular schooling markets.

Under the standard matching function paradigm of Mortensen and Pissarides, we think of the contact rate as being determined as a constant returns to scale (CRS) production technology

$$\begin{aligned} M &= M(u, v) \\ &= vm(k), \end{aligned}$$

where $k = u/v$ is the measure of labor market tightness, v is the measure of vacancies, and m is a concave function of k . In the homogenous markets case, firms create vacancies that can generate a match with any type of individual, \tilde{a} , with the likelihood of the firm meeting a type \tilde{a} proportional to the population density of type \tilde{a} .

We will assume that M is Cobb-Douglas, as is commonly done, with $M(u, \tilde{a}) = Au^\delta v^{(1-\delta)}$, $\delta \in (0, 1)$. For purposes of estimation, in what follows we set $A = 1$. Then $m(k) = k^\delta$, and the contact rate from the point of view of workers, is

$$\begin{aligned} \lambda &= \frac{M(u, v)}{u} = \frac{\tilde{a}k^\delta}{u} \\ &= k^{\delta-1}. \end{aligned}$$

From the point of view of firms with vacancies, the contact rate is

$$\begin{aligned} \lambda_F &= \frac{M(u, \tilde{a})}{v} \\ &= k^\delta. \end{aligned}$$

Under the standard free entry condition that results in the expected value of a vacancy being equal to 0, the vacancy rate is determined by

$$0 = -\psi + k^\delta \tilde{G}(\theta^*) J,$$

where J is the expected value of a filled vacancy. With no on-the-job search, the value of a filled vacancy with a type \tilde{a} individual and a match draw of θ is

$$\frac{(\tilde{a}\theta - w) + \eta V_v}{\rho + \eta} = \frac{\tilde{a}\theta - w}{\rho + \eta}$$

under the free entry condition ($V_v = 0$). Since the wage is

$$w = \tilde{a}(\alpha\theta + (1 - \alpha)\theta^*),$$

then

$$\begin{aligned} \tilde{a}\theta - w &= \tilde{a}\theta - (\tilde{a}(\alpha\theta + (1 - \alpha)\theta^*)) \\ &= \tilde{a}(1 - \alpha)(\theta - \theta^*), \end{aligned}$$

so that

$$J = (\rho + \eta)^{-1} \{(1 - \alpha)E\tilde{a} \int_{\theta^*} (\theta - \theta^*) \frac{dG(\theta)}{\tilde{G}(\theta^*)}\},$$

since \tilde{a} and θ are independently distributed. Then the free entry condition implies

$$0 = -\psi + k^\delta(1 - \alpha) \frac{E\tilde{a}}{\rho + \eta} \int_{\theta^*} (\theta - \theta^*) dG(\theta).^{10}$$

4.2 Heterogeneous Markets

In considering endogenous contacts in heterogeneous markets, where in the partial equilibrium case we have assumed that $\lambda_1 \neq \lambda_2$, it follows that we allow firms to post vacancies in specific schooling sub-markets. This is eminently reasonable, since firms posting jobs routinely specify educational requirements for the position. Because any firm is free to target either schooling type, the free entry condition applies to vacancy postings of both types, so that $V_v^1 = V_v^2 = 0$.

By an obvious generalization of the homogeneous markets case, we will assume that matches in schooling market i are determined by

$$M_i = u_i^{\delta_i} v_i^{1-\delta_i}, \quad i = 1, 2,$$

so that the contact rates in the two markets are

$$\begin{aligned} \lambda_i &= k_i^{\delta_i - 1}, \\ \lambda_{F,i} &= k_i^{\delta_i}, \\ k_i &= \frac{u_i}{v_i}, \end{aligned}$$

¹⁰It is interesting to note that when firms are not able to direct vacancies to a schooling sub-market, the inefficient over-education result of Cherlot and Decreuse (2005) disappears. Then the tax on education that they impose to attain the efficient schooling level might be replaced with a prohibition on firms mentioning education credentials when they post vacancies. Of course, both “solutions” depend on the very restrictive assumptions underlying these models.

for $i = 1, 2$. The unemployment rate is defined for each schooling sub-market, consistent with the schooling-specific vacancy postings. With the cost of a holding a vacancy in schooling market i given by ψ_i , the free entry conditions are given by

$$0 = -\psi_1 + k_1^{\delta_1}(1 - \alpha_1)(\rho + \eta_1)^{-1} \int^{a^*} a \frac{dG(a)}{G(a^*)} \int_{\theta_1^*} (\theta - \theta_1^*) dG(\theta) \quad (7)$$

$$0 = -\psi_2 + k_2^{\delta_2}(1 - \alpha_2)(\rho + \eta_2)^{-1} h_2 \int_{a^*} a \frac{dG(a)}{G(a^*)} \int_{\theta_2^*} (\theta - \theta_2^*) dG(\theta). \quad (8)$$

Using a similar framework,¹¹ Charlot and Decreuse (2005) make the point that a decentralized equilibrium will generally result in inefficient outcomes (for any given common value of $\alpha = \alpha_1 = \alpha_2$). Individuals on the margin look at their gains from acquiring schooling without considering the impact of their decision on the contact rates that other individuals in both schooling groups will experience. Those individuals who take up schooling are the more able, thus as individuals near the margin leave the low-schooling group to enter the high-schooling group, these transitions decrease the average quality of both schooling groups. This decreases the incentives of firms to create vacancies in both of the sub-labor markets. The authors propose a tax on higher education to solve this problem, since the inefficiency results from too many (of the lower ability) individuals entering the high-schooling market.

In the case of one sharing parameter for both markets, we know that the sharing parameter α has three impacts on the behavior of the agents. Increases in α are associated with the following:

1. Increases in the return to schooling investment since the severity of the hold-up problem is reduced.
2. Decreases in the firms' incentives to create vacancies through the reduced share of the surplus they receive.
3. Decreases in the incentives of firms in each sector to produce vacancies since a^* will decrease, thus lowering the expected output in both sectors, further reducing the incentive to create vacancies.

All of these impacts are present when trying to determine an "optimal" single sharing parameter α .

5 Econometric Issues

We begin this section by discussing the data utilized to estimate the model(s). We then provide a discussion of the estimation method and identification of model parameters given

¹¹Charlot and Decreuse do not include match-specific heterogeneity in productivity nor do they allow different rent-sharing rules in the two sectors.

the nature of the data available. Model estimates and comparative statics exercises are presented in the following section.

5.1 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 2005.¹² ORGs include households in their 4th or 8th survey month, and in these months detailed earnings and employment information is ascertained. We selected all males between the ages of 25 and 34, inclusive.

To be included in the final sample, an individual had to either be employed or unemployed, and had to have valid schooling completion information. In addition, we drop individuals from the sample who did not attend school beyond the 8th grade. We impute wage information from data on usual weekly hours and usual weekly earnings. Although some individuals have missing information on earnings or unemployment duration, we include them in our calculation of the college completion rate and the employment rate, while excluding them, by necessity, from any moments involving wages or unemployment duration. Our final sample consists of 21,311 individuals. Of these, 17,494 individuals have valid wage information and 794 have valid information on unemployment duration.

After experimenting with various schooling classifications systems, we selected 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 the high schooling 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.

Table 1 contains the descriptive statistics from the final estimation sample. We see that 30 percent of individuals 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 in the entire sample was quite low, at 3.73 percent, with a significant difference between the unemployment rates of the high school group (2.16 percent) and of the low schooling group (4.67 percent), upon which we have already commented. On average, an unemployed sample member had been searching for work for 20.62 weeks.

The average hourly wage in the sample is \$17.94 (in 2005 dollars), with a substantial difference between those with less than college completion (\$15.34) and those who have

¹²We selected the year 2005 instead of a more recent year so as not to capture the turmoil in the labor market following the Great Recession beginning in 2008.

completed college (\$23.40). There is also a substantial degree of dispersion in these distributions, with much more dispersion seen in the college wage distribution. The presence of some outliers in both distributions has a large impact on the means and variances. Since we will use deciles of the wage distribution to derive model estimates, the effect of wage outliers on our estimates should not be large.

5.2 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 our sample, the value of 0.67 seems to be in the right range of values.

Estimating a model in which deconvolution is attempted is always challenging, as will be evident below. To gain further identifying information regarding the individual distributions of a and θ , we exploit a very limited amount of data taken from the Survey of Income and Program Participation (SIPP). We use longitudinally linked data from waves 5 and 6 of the 2004 SIPP. Each wave of the SIPP is administered every 4 months, with each wave containing labor force participation and earnings information for the previous 4 months. Individuals in their fourth (and final) longitudinal month of Waves 5 and 6 include the months of May through December, 2005. We make the same age restrictions that were imposed for the main sample. We keep only individuals who (1) were employed in their final month of both waves¹³; (2) report being at a different employer in wave 6; and (3) have valid wage information at both sample points. These restrictions leave us with 77 workers from which to compute the moments of the change in wages when job switching.

Let w_t be the wage in the job held a time t . Then

$$w_t = \tilde{a}(\alpha\theta_t + (1 - \alpha)\theta^*),$$

so that the differences in the log-wage rates is

$$\begin{aligned} \Delta \log(w) &\equiv \log(w_2) - \log(w_1) \\ &= \log(\theta_2 + (1 - \alpha)\theta^*) - \log(\alpha\theta_1 + (1 - \alpha)\theta^*). \end{aligned}$$

¹³The design of the SIPP is such that the final wave month will not correspond to the same calendar month for all individuals.

Under our assumption of i.i.d. sampling, the expected value of $\Delta \log(w)$ is 0, and the variance is

$$2\mathbb{E}[\log(\alpha\theta + (1 - \alpha)\theta^*)^2 | \theta \geq \theta^*]. \quad (9)$$

This result is used below in forming our estimator of the model parameters. It is helpful since it is only a function of $G(\theta)$, θ^* , and α . In particular, the distribution of ability and h_2 do not appear.

5.3 Identification

The primitive parameters in the homogeneous markets case are $\rho, b, \lambda, \eta, F_a, G, h_2$, and c . Much of the identification analysis can be conducted using results from Flinn and Heckman (1982), hereafter referred to as FH, after noting which of the parameters determine labor market outcomes explicitly once we condition on the observed value of schooling, s . Their analysis was likelihood-based, and for reasons to be discussed below, we employ a method of moments estimator. However, the arguments that they make carry over to the case of our estimator given that we select “appropriate” moments to match. The advantage of our estimator is that it remains well-defined for samples in which some individual observations are zero-probability events under the model. While less asymptotically efficient than m.l. when the model specification is “correct,” as long as the method of moments estimator includes functions of the data that are similar to those comprising the minimal sufficient statistics of the data that characterize the likelihood function, identification and “good” asymptotic properties should apply to the moments-based estimator that we utilize. The reasonableness of the parameter estimates and the small boot-strapped standard errors for most of the parameters provide some prima facie evidence in support of this claim.

As in the m.l. estimator used by FH, we treat the decision rules a^* and θ^* as constants. There is no loss in efficiency in doing so, since the simple search model they investigated and the one estimated here are both fundamentally underidentified. The model they investigate is a special case of the one analyzed here. In particular, it is the special case of our model for which $\alpha = 1$, $a = 1$, and $h = 1$ for all agents. With no schooling decision, the sole decision rule of the model is θ^* . They show that parametric distributional assumptions are required to estimate the model. Given the parametric assumption, all model parameters are identified except for the pair (b, ρ) . Fixing one of these at some predetermined value allows the other to be consistently estimated given the functional equation defining the critical value. Typically, the value of ρ is fixed and the value of b is then imputed. In this manner, consistent estimates of λ, η, G , and b are obtained given a value of ρ .

The extension to the model we estimate allows α to be a free parameter and introduces a nondegenerate distribution of a , given by F_a , a human capital level h_2 (which is presumably greater than 1, the normalized value of h_1), and a cost of schooling, c . As we showed above, the critical schooling value a^* is a function of all of the primitive parameters of the model through θ^*, h_2 , and c (and assuming a value ρ). Since c only appears in the schooling choice

problem, an estimated value of a^* , in conjunction with consistent estimates of F_a , h_2 , is used to form a consistent estimate of c .

As we showed above, conditional on s , variability in schooling decisions and wage outcomes (across individuals and over time) is generated by the two independent random variables, a and θ . As we have shown, under our model assumptions the critical match value θ^* is independent of the schooling decision 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

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

We will assume that F_a belongs to a parametric family, which will include an assumption regarding the support of the distribution of a . If we were to assume that $\underline{a} = 1$, for example, then $\underline{w}_1 = \theta^*$. In this case, using the order statistic estimator of FH, 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.

Taking this argument one step further, it is possible to reduce the number of parameters to be estimated by one by using the first order statistic of the distribution of wages in the high-schooled group, since $\underline{w}_2/\underline{w}_1 = a^* h_2$ under the assumption that $\underline{a} = 1$. In this case we can write

$$h_2 = \frac{1}{a^*} \frac{\underline{w}_2}{\underline{w}_1},$$

and a superconsistent estimator of the function $h_2(a^*)$ is given by

$$\hat{h}_2(a^*) = \frac{1}{a^*} \frac{\min\{w_i\}_{i \in S_2}}{\min\{w_i\}_{i \in S_1}}. \quad (10)$$

By using this function and treating a^* as a free “parameter,” we eliminate the need to estimate h_2 as a free parameter.

The wage equations for the two schooling groups are convolutions of truncated random variables a and θ . It is well known that solving the deconvolution problem is challenging even under tight functional form assumptions and a large number of i.i.d. observations. The wage densities for the two groups are exhibited in (5) and (6), where we see that they depend on the distributions G and F_a , the decision rules θ^* and a^* , and the scalar parameters h_2 and α . The parameter α is largely identified from the labor share equation

we use in the estimation, and since the estimate of a^* is effectively obtained from the proportion of those who are in the high schooling class, i.e., $P(s = 2) = \tilde{F}_a(a^*)$, the sample proportion of those in the high schooling class in conjunction with a consistent estimate of F_a serves to determine a^* , which taken together with (10) serves to identify h_2 .

Even with all of the structure that the model provides on the form of the convolution, strict functional form assumptions are required to identify the parameters characterizing G and F_a . From FH we know that functional form assumptions are required to identify G when rejected job offers are not available (as they virtually never are). Because the entire distribution of a is reflected in the wage distributions, theoretically it may be possible to identify F_a nonparametrically under certain conditions. However, as shown by Heckman and Singer (1984) in the case of mixture distributions, poor estimates of F_a would likely emerge even with an enormous number of observations. For these reasons, we assume a parametric form for F_a .

In order to make the distributions of θ and a as symmetric as possible, we assume that both are lognormal, with parameters $(\mu_\theta, \sigma_\theta)$ and (μ_a, σ_a) , respectively. This setup permits a normalization of $\mu_\theta = 0$. Notice that the lognormal assumption regarding F_a implies that the support of F_a is R_+ , which is at odds with some of our previous arguments that relied on the assumption that the lower support point of F_a was 1. While we could have utilized a truncated (from below at 1) lognormal distribution, we encountered few problems in estimating the model with the estimator described below under the assumption that the lower support point of F_a was 0.

An extremely important piece of information that was utilized in sorting out the contributions of θ and a to the wage distributions was described in the previous subsection. From the SIPP data, we differenced the wages earned by a respondent in two different jobs (separated by an unemployment spell). Under the model assumptions (stationarity and no OTJ search), the mean of this difference should be 0; we found that there was no reason to reject this hypothesis. Under the model, the variance of the difference is a function of α and $V(\theta|\theta > \theta^*)$. Since the mean of θ is normalized to 0, this is a function solely of σ_θ . Thus, given a credible estimator for α , this moment contains a large amount of information regarding the distribution of θ .

Our credible estimator of α comes from imposing the restriction that the labor share from the model match the labor share in the aggregate economy, which we argued above could be taken to be two-thirds. The constraint was imposed in a manner similar to Flinn (2006), except for the difference in the form of the estimator used. When estimating the heterogeneous markets case, the aggregate labor share measure was defined as the ratio of the weighted average of wages in the high- and low-schooling markets divided by the weighted average of output in the two sectors, where the weights are just the proportions of individuals in the low- and high-schooling markets. We use this procedure because we do not have access to labor share by schooling class.

Identification of the rate parameters, λ and η , is straightforward with the data to which we have access, particularly under the no OTJ search assumption. The identification

arguments for these parameters can be found in Flinn (2006), for example.

We have discussed identification in the homogeneous markets case. There are no new conceptual issues that arise in the identification of the heterogeneous markets model. The estimation of the additional rate parameters (λ_i and η_i for sectors $i = 1, 2$) is straightforward since we observe steady state unemployment and the duration distribution of on-going unemployment spells for both schooling groups. The most challenging issue is the identification of α_1 and α_2 . This is accomplished through the use of the labor share information and restrictions on the relationship between θ_1^* and θ_2^* . Under the assumption that b_0 is common to all individuals, using all other identified parameters, we have $\theta_1^* = \theta_1^*(\alpha_1, b_0, \omega_1)$ and $\theta_2^* = \theta_2^*(\alpha_2, b_0, \omega_2)$, where ω_i is the vector of all the other parameters that characterize the employment decision rule in market i . With our estimates of θ_1^* and θ_2^* , ω_1 and ω_2 , we can use these relationships along with the labor share information to identify α_1, α_2 , and b_0 .

5.4 Estimator

Although our identification discussion was likelihood-based, for a variety of reasons we utilize a method of moments estimator to estimate the model. Under the data generating process of the model, there are a number of sharp restrictions on the support of the wage distributions and the relationship between the two that aid in identification but that are not generally consistent with the empirical distributions observed. In such a case, measurement error in wage observations is often added to the model. This is not really a feasible alternative here given that we are already trying to estimate a convolution, so that the addition of another random variable to the wage process can only exacerbate the identification problems that we face.

We instead choose to use a moment-based estimator which employs a large amount of information characterizing the wage distributions by schooling class, but which does not impose such a large penalty on the estimator for violating some of the implications of the data generating process (DGP) of the model. Of course, there is some loss of efficiency (assuming that the model is correctly specified) in general, although we find that we are able to precisely estimate the majority of the parameters characterizing the model in any event.

The information from the sample that is used to defined the estimator is given by M_N , where there are N sample observations. Under the DGP of the model, the analogous characteristics are given by $\tilde{M}(\omega)$, where ω is the vector of all identified parameters (which are all parameters except ρ). Then the estimator is given by

$$\hat{\omega}_{N, W_N} = \arg \min_{\omega \in \Omega} (M_N - \tilde{M}(\omega))' W_N (M_N - \tilde{M}(\omega)),$$

where W_N is a symmetric, positive-definite weighting matrix and Ω is the parameter space. The weighting matrix, W_N , is a diagonal matrix with most elements equal to the inverse

of the variance of the corresponding element of M_N , which is estimated by resampling the data.¹⁴ The exceptions to this rule are the weights attached to the average squared difference in ln wage observations across an unemployment spell, which is computed using SIPP data, the relationships involving the labor share in the homogeneous markets version of the model, and the relationships involving the labor share and the θ_1^* and θ_2^* in the heterogeneous markets case. In a likelihood-based estimator, the constraints involving everything other than the moment from the SIPP data would be imposed as an exact equality (as in Flinn, 2006). The analog here is to attach a very high weight to discrepancies of the fitted characteristics under the model to the corresponding data (as in labor share information) or functions of other model parameters (in the case of θ_1^* and θ_2^* in the heterogeneous markets case). The very high weight given to the sample moment from the SIPP data was due to its outsized value in helping to solve the deconvolution problem that we face.

Under our random sampling assumption, it follows that $M_N \xrightarrow{P} M$, the population value of the sample characteristics used in estimation. Since W_N is a positive-definite matrix by construction, our moment-based estimator is consistent since $\hat{\omega}_{N,Q} \xrightarrow{P} \omega$ for any positive-definite matrix Q . We compute bootstrap standard errors using 200 replications.

The sample characteristics we use vary slightly across the homogenous and heterogeneous markets specification. In both cases we use the location of deciles 1 (at the tenth percentile) through 9 (at the 90 percentile) in the conditional (on schooling) wage distributions. This is then 18 sample characteristics. For the homogeneous case, we supplement these with the sample proportion in the unemployment state and the mean duration of unemployed search in the entire sample. As mentioned above, we also include the squared differences in ln wages from the SIPP data and the labor share value. In all then, there are 22 pieces of sample information utilized.

In the heterogeneous markets case, we distinguish the unemployed proportions and the average duration of unemployment by schooling class. We also add the restrictions relating the critical match values θ_1^* and θ_2^* and the underlying b , α_1 , and α_2 . In all then, we utilized 26 pieces of information to estimate the heterogeneous markets model.

5.5 Estimation of Parameters in the GE Version of the Model

We now consider the estimation of the parameters that appear in the GE version of the model. Because it is reasonable to assume that vacancies can be targeted to a particular

¹⁴We computed the M_N^g vector for each of Q resamples of the original N data points, and the covariance matrix of M_N is given by

$$W_N = \left(Q^{-1} \sum_{g=1}^G (M_N^g - M_N)(M_N^g - M_N)' \right)^{-1}.$$

The number of draws, Q , was set at 500.

schooling submarket, we only perform the empirical analysis for the heterogeneous markets case.

As stated above, any firm can post a vacancy in either schooling market, and it is assumed that free entry drives the expected value of a vacancy to zero in both markets. One of the principal problems with estimating the parameters that appear in the GE version of the model is the lack of credible information on vacancies. Even with this information, we can easily see that the unrestricted GE specification is not identified. We can rewrite (7) and (8) more succinctly as

$$0 = -\psi_i + k_i^{\delta_i} A_i,$$

where A_i is the expected value of a filled vacancy in schooling market i . From our first stage estimates of the partial equilibrium model, we can consistently estimate A_i . Now $k_i = u_i/v_i$; a consistent estimate of the steady state unemployment rate in market i , u_i , is available, and for the moment, assume that v_i is as well. Then the FEC produces two equations in four unknowns, and thus we cannot hope to attain consistent estimates of these parameters without further restrictions.

We consider the case in which v_i is unobserved in each sector, so that it must be estimated along with the primitive parameters λ_i and δ_i , $i = 1, 2$. The only case in which it is possible to estimate these parameters is under the strong restrictions:

$$\begin{aligned} \psi_1 &= \psi_2 \equiv \psi \\ \delta_1 &= \delta_2 \equiv \delta. \end{aligned} \tag{11}$$

We first must determine vacancies in each sector. We have a consistent estimator of the arrival rate of job contacts from the point of view of workers in sector i , λ_i , as well as the SS unemployment rate in the sector, u_i . Under (11), we have

$$\begin{aligned} \lambda_i &= k_i^{1-\delta} \\ \Rightarrow v_i &= \lambda_i^{-\frac{1}{\delta-1}} u_i \end{aligned}$$

so that

$$\begin{aligned} k_i^\delta &= \left(\frac{u_i}{u_i \lambda_i^{-\frac{1}{\delta-1}}} \right)^\delta \\ &= \lambda_i^{\frac{\delta}{\delta-1}}. \end{aligned}$$

Then we can write the two equation system as

$$\begin{aligned} 0 &= -\psi + \lambda_1^{\frac{\delta}{\delta-1}} A_1 \\ 0 &= -\psi + \lambda_2^{\frac{\delta}{\delta-1}} A_2. \end{aligned}$$

With consistent estimates of λ_1 , λ_2 , A_1 , and A_2 , we can solve this two equation system in two unknowns for unique values of δ and ψ . Since these two functions are continuous functions of consistent estimators, the solutions to the system are consistent estimates of ψ and δ .

While we can consistently estimate ψ and δ under the assumption that they are common in the two labor markets, this restriction may not even approximately hold. As we will see in the empirical results discussed below, this solution method yields negative estimates of δ , which is outside of the permissible set of $(0, 1)$. We interpret this as a strong rejection of the homogeneity assumption. In this case, we must allow heterogeneity in at least one of these parameters, and in our counterfactual experiments we allow heterogeneity in the flow costs of holding vacancies, ψ_1 and ψ_2 , while we continue to insist on a common value of δ . In this case, the parameters are underidentified, and we are only able to “back out” estimates of ψ_1 and ψ_2 after conditioning on specific values of δ . We utilize typical estimates of this parameter from the macroeconomics empirical literature (see the survey by Petrongolo and Pissarides, 2001). We perform our exercises for values of δ in the set $\{0.4, 0.6\}$ as well as at the model estimate of bargaining power $\hat{\alpha}$.

6 Model Estimates and Comparative Statics Exercises

6.1 Homogenous Model

Our estimates of the homogenous model appear in Table 3. Our method of moments procedure produces point estimates of

$$(\lambda, \eta, \mu_a, \sigma_a, \sigma_\theta, \alpha, h_2, \theta^*).$$

From the observed college completion rate we can then back out the implied values of a^* and c . In addition, setting $\rho = 0.004$, we can back out an estimate of b_0 from the reservation wage equation.

Our estimates imply a 8.7% return in productivity to school investment, which is in the range of previous estimates in the literature. In addition, the standard deviation of match values is estimated to be approximately 2.5 times the standard deviation of ability in the population. Although distorted by their respective truncation points, this will correspond roughly to the relative contributions of match and worker heterogeneity to the overall wage variance. In previous work, Postel-Vinay and Robin (2002) found that such a decomposition varies by the occupational category of the worker, but given their results¹⁵, we take this number as being within a range of reasonable aggregate values. We should bear in mind two things, however, in making this comparison. First, Postel-Vinay and Robin are estimating the variance of \tilde{a} , which is the product of individual ability and human capital produced by schooling, or $\tilde{a} = ah_{s(a)}$, which is not a primitive distribution in the case examined here.

¹⁵See Table VII of Postel-Vinary & Robin (2002).

Secondly, the other component of the match value was the firm’s productivity. In our case, the other component determining match productivity is a match-specific draw θ .

We estimate worker bargaining power in the model to be 0.55, while our estimates imply that the probability of accepting a match is nearly 1. The estimated arrival rate of offers (0.20) implies that the average wait for a contact is about 5 months, while the job destruction rate (0.0075) implies that jobs last on average about 11 years. Needless to say, this estimate, in particular, would change if we were using longitudinal information and if our model included OTJ search.

Table 4 presents the sample characteristics from the data used to estimate the model, along with the model’s implied characteristics evaluated at the point estimates. There is a reasonably good fit throughout the deciles of the wage distributions, though there is some systematic under- or over-prediction across deciles in the two distributions. This is largely due to the restrictiveness of the modeling assumptions and the large weight given to the information from the SIPP data which was essential to solve the deconvolution problem.

6.2 Heterogeneous Model

We now turn to our estimates from the heterogeneous model, in which the constraint that all workers face the same arrival rate λ and separation rate η in each schooling market is relaxed. In this case, we seek to minimise the criterion function with choices of

$$(\lambda_1, \lambda_2, \eta_1, \eta_2, \mu_a, \sigma_a, \sigma_\theta, \alpha_1, \alpha_2, h_2, \theta_1^*, \theta_2^*, b_0),$$

which contrasts with the homogenous case in that the value of unemployment b_0 can no longer be freely backed out from the parameter estimates. Rather, we require a choice of b_0 that satisfies the reservation wage equation for both high and low schooling groups, given values of θ_1^* and θ_2^* .

We first estimate the model allowing differences in all labor market parameters. These results are presented in Table 5. However, we find in this case that the estimated bargaining power parameters (α_1, α_2) are very close. This is largely due to the fact that the key piece of identifying information for the bargaining power parameters is the labor share. Since we do not have access to separate labor share measures by schooling group, the aggregated measure we have access to tends to restrict the difference between them.¹⁶ Motivated by this finding, we estimate the model under the restriction that $\alpha_1 = \alpha_2$, while the parameters $(\lambda_s, \eta_s, \theta_s^*)$ remain free to differ across markets. We focus our attention on the results of this model, which are presented in Table 6. The common α assumption actually is useful in performing the comparative statics exercises we consider next.

The human capital return to schooling is now estimated at 7.57%, slightly lower than in the homogenous case, but still in a reasonable range. We find that both the rate of

¹⁶Of course, we don’t know that the actual values of α_1 and α_2 are not similar, or even equal. The point is only that this estimator will tend to reduce our ability to estimate true differences if in fact there are any.

arrival of offers and the hazard rate out of unemployment are slightly higher for the high schooling type. In both markets, the implied probability of accepting a match is close to 1, as was the case for the homogenous model. The high schooling types are estimated to face a slightly higher rate of job destruction, although the difference is not so large as to undo the finding in the data that high schooling individuals have lower unemployment rates.

Table 7 presents the moments from the data used to estimate the model, along with the model implied moments at our parameters of best fit. We see that the addition of the new parameters leads to some degradation in the ability of the model to fit all of the sample information utilized. In the end, adding heterogeneity in some basic primitive parameters does not lead to a dramatic difference in our inferences. This is in part due to the limited amount of information available to estimate some key model parameters.

6.3 Comparative Statics Exercises

We now turn to the question of whether a shift in the bargaining power towards the worker can induce improvements in welfare and remediate the holdup problem. As mentioned above, we only consider the heterogenous markets case in which employers can post vacancies in both schooling sub-markets.

6.3.1 Partial Equilibrium

We first consider the effect of changes in the worker’s bargaining power in partial equilibrium. That is, we do not consider the effect of varying α on vacancy creation, only the effect on workers’ schooling decisions and reservation match value policies. To do this, we use our estimates of the segmented market model under the restriction that $\alpha_1 = \alpha_2$. This provides the cleanest case for examination, since there is no clear benchmark when $\alpha_1 \neq \alpha_2$. The results of this experiment are presented in Figure 2.

Panel (a) depicts changes in the proportion of agents who elect to go to college, which is determined by changes in a^* in partial equilibrium. Panel (b) shows the unemployment rate in steady state, which is a weighted average of unemployment rates in each market where the weights are simply the proportion of low and high educated agents. Panel (c) shows worker welfare in steady state; a weighted average of worker’s value functions in all possible employment and ability states across schooling markets. The choice of weights corresponds directly to the steady state mass of workers in each particular state. Finally, panel (d) shows productivity in steady state, taken simply as total output normalized by the employment rate, or equivalently as average output in the economy.

We see that the completion rate is quite sensitive to bargaining power in partial equilibrium. In fact, when bargaining power drops below 0.35, no one completes college. On the other hand, as α increases to one, virtually all individuals in the population complete college.¹⁷ To obtain a completion rate of 50 percent requires an increase in bargaining power

¹⁷It is important to keep in mind that in conducting these exercises, we keep the cost of schooling constant.

in the model to about 0.64, compared to its estimated value of 0.55. From Panel (b) we see that the unemployment rate increases as individuals become more selective concerning the idiosyncratic match value, due to them receiving ever larger proportions of the surplus. In Panels (c) and (d), we see that the steady state welfare on the supply side of the market and the average steady state productive are both monotonically increasing in the worker's bargaining power. This is predictable from the theory alone; these estimates simply give us the quantitative magnitudes, and they are large.

Since these comparative static exercises have been performed in partial equilibrium, however, it is reasonable to expect the returns to such a shift in bargaining power are overstated. An immediate question, to which we now turn, is the size and direction of these effects when firms' job creation decisions are endogenously determined.

6.3.2 General Equilibrium

Before we evaluate the impact of changes in bargaining power, we make note of the fact that our segregated market model admits multiple equilibria. To see this, consider the following. Conditional on a marginal worker, a^* , who is indifferent between high and low schooling, the firm's expected productivity of a worker found in each market is determined. Given these expectations, market tightness (k_1, k_2) and reservation policies (θ_1^*, θ_2^*) can be determined using the reservation wage equations and the free entry conditions for each market. This defines respective values of unemployment in each market and, hence, an implied marginal worker a^{**} . Call this mapping from a^* to a^{**} , φ . General equilibrium then in this model requires the existence of a marginal worker a^* such that $a^* = \varphi(a^*)$. Figure 3 shows that the mapping φ has multiple intersections with the 45° line, implying multiple equilibria.

Thankfully, our estimates in this case make an equilibrium selection, since the model estimates imply an a^* . Figure 3 shows that our estimates imply the equilibria with the higher level of schooling investment. Therefore, in our experiments below, we solve always for the high equilibrium and do not consider transition dynamics away from this point.

Recall from our previous discussion, in Section 5.5, that the vacancy posting costs ψ_1 and ψ_2 are identified subject to a choice of the workers' Cobb-Douglas share in the matching function, δ . It is well known, however, that in the standard search and matching model, welfare outcomes depend critically on the choice of this parameter. In particular, welfare is maximized when the workers' bargaining power is equal to their share in the matching function, $\delta = \alpha$ (Hosios 1990). In light of this, we choose to evaluate labor market effects at three points; $\delta = 0.4$, $\delta = \hat{\alpha}$ and $\delta = 0.6$. This encompasses a range of values considered reasonable in the literature (Petrongolo and Pissarides, 2001).

The results of our general equilibrium experiment when $\delta = 0.4$, $\delta = \hat{\alpha}$ and $\delta = 0.6$ are displayed in Figures 4, 5 and 6, respectively. The panels in each figure correspond exactly

It is unreasonably to think that this would be the case if the size of the college-educated population doubled or tripled.

to those described in Section 6.3.1 for Figure 2. The most striking result is that the general equilibrium effects strongly temper the results found in the partial equilibrium analysis. In particular, when $\delta = 0.4$, both the completion rate and steady state welfare are in fact decreasing when α increases beyond our baseline estimate. This pattern can be identified as the search market effects (since decreasing δ moves us closer to the Hosios condition) dominating the holdup effects in general equilibrium.

The effects on unemployment in each case are strong, as the shift in bargaining power crowds out vacancy posting by firms, decreasing the flow rate out of unemployment. General equilibrium effects in the labor market as α increases dampen the incentive to obtain schooling and so we see much weaker effects on the college completion rate and worker productivity than in the partial equilibrium case. These variables seem to vary almost one-to-one, since productivity changes largely come from the number of workers who undertake the productivity enhancing schooling decision (the effect of changes in θ_1^* and θ_2^* appear to be second order).

Finally, we can observe that marginal increases in worker bargaining power are welfare improving for the $\delta = \hat{\alpha}$ and $\delta = 0.6$ cases. This is most revealing for the case in which $\delta = \hat{\alpha}$ and delivers the essence of the hold up problem; the introduction of a first stage investment decision breaks the constrained pareto optimality that was previously guaranteed by the Hosios condition. In this case we see that while a small change in α from our estimated value can improve welfare and productivity, large increases in α worsen productivity (through the employment effect) and the welfare of agents on the supply side of the market.

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 focus of our analysis is 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 college completion. In the partial equilibrium case, we found dramatic effects. At values of the bargaining power parameter less than 0.35, we found that no one completed college, while as $\alpha \rightarrow 1$ virtually all individuals completed college. In the general equilibrium case, results were strongly different. Beginning from our estimated value of $\hat{\alpha} = 0.55$, movements far above this level led to decreases in college enrolment, welfare, productivity, and employment as employers stopped creating vacancies. The fact that α has a role in both incentivizing investment on one side of the market and in incentivizing job creation on the other lead to a modification of the standard Hosios condition that only considers the latter problem. The optimal level of α from the point of view of steady state productivity was a value slightly greater than

our estimated value of 0.55.

There are several ways in which our approach can be generalized. Most obvious is allowing for firm heterogeneity and capital investment, as was discussed in Section 2.1. This is clearly a very interesting direction to explore, though from an empirical standpoint this extension seems to be a difficult one. As we have discussed above, even without firm heterogeneity solving the deconvolution problem is extremely challenging. Adding another source of heterogeneity in the wage distribution will only exacerbate the problem.

On-the-Job (OTJ) search, on the other hand, does seem to be a tractable extension and one which may significantly alter some of the conclusions we have reached. The extent to which it does will be impacted by the assumptions made regarding firm behavior when an employee receives an alternative offer and the rate at which these contacts occur.

Using the Bertrand competition assumption utilized in Postel-Vinay and Robin (2002), Dey and Flinn (2005), and Cahuc et al. (2006), the effect of the surplus division parameter α on schooling investment may be considerably weakened. To see why, we will first quickly review the assumption in the Nash-Bargaining case (Dey and Flinn (2005) and Cahuc et al. (2006)). Under our assumption on the structure of productivity, $y = \tilde{a}\theta$, the individual's flow productivity at two potential employers is $y' = \tilde{a}\theta'$ and $y'' = \tilde{a}\theta''$. In the papers mentioned, the firms engage in a bidding war for the employee's services up to the point where one firm drops out, which in this case will be the firm with the smallest match value θ . Say that $y'' < y'$. The firm where productivity is y'' exhausts its surplus when it offers to set the wage equal to the individual's productivity, so that $w'' = \tilde{a}\theta''$, with an associated value of unemployment at that firm given by $V_E(\tilde{a}\theta'', \tilde{a}, \theta'')$, where the first argument is the individual's wage at the firm. This value serves as the outside option in the surplus division exercise with the other firm. When a currently unemployed individual bargains with a firm, his outside option is simply V_U , the value of unemployed search; in this case, analogous to the situation we have analyzed in this paper, the bargaining power parameter is a key component of the determination of the rate of return to schooling investment. In the case of OTJ search, particularly when contacts with alternative employers are relatively frequent, the role of α becomes much less important. This is reflected in the estimates of α that researchers have obtained under the assumption of OTJ search with Bertrand competition between employers. Almost uniformly, estimated values of α decline significantly in this case, since much of the dispersion in wage distributions can be generated through employer competition rather than bargaining power (in the form of α) per se. With a lower (anticipated) estimate of α , but firm competition, it would be of interest to revisit the hold-up problem.

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] Charlot, O., and B. Decreuse (2005) “Self-Selection in Education with Matching Frictions,” *Labour Economics* 12: 251–267.
- [8] Charlot, O. and B. Decreuse (2010) “Over-Education for the Rich, Under-Education for the Poor: A Search-Theoretic Microfoundation,” *Labour Economics* 17: 886–896.
- [9] Dey, M. and C. Flinn (2005) “An Equilibrium Model of Health Insurance Provision and Wage Determination,” *Econometrica* 73: 571–627.
- [10] 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.
- [11] Flinn, C. (2006) “Minimum Wage Effects on Labor Market Outcomes under Search, Bargaining, and Endogenous Contact Rates.” *Econometrica* 74: 1013–1062.
- [12] Flinn, C., and J. Heckman (1982) “New Methods for Analyzing Structural Models of Labor Force Dynamics.” *Journal of Econometrics* 18: 115–68.
- [13] 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.

- [14] 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.
- [15] Hosios, A. (1990) “On the Efficiency of Matching and Related Models of Search and Unemployment,” *Review of Economic Studies* 57: 279-298.
- [16] Keane, M and K. Wolpin (1997) “The Career Decisions of Young Men,” *Journal of Political Economy* 105: 473-522.
- [17] Krueger, A. (1999) “Measuring Labor’s Share,” *American Economic Review* 89(2): 45-51.
- [18] Lee, D. (2005) “An Estimable Dynamic General Equilibrium Model of Work, Schooling and Occupational Choice,” *International Economic Review*, 46: 1-34.
- [19] Lee, D and K. Wolpin (2006) “Intersectoral Labor Mobility and the Growth of Service Sector,” *Econometrica* 47: 1-46.
- [20] Malcomson, J. (1997) “Contracts, Hold-Up, and Labor Markets,” *Journal of Economic Literature* 35: 1916-1957.
- [21] 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.
- [22] Pissarides, C. (2000) *Equilibrium Unemployment Theory*, 2nd edition. Cambridge, MA: MIT Press.
- [23] Postel-Vinay, F., and J-M Robin (2002) “Equilibrium Wage Dispersion with Heterogeneous Workers and Firms.” *Econometrica* 70: 2295-1350.
- [24] Shimer, R., R. Rogerson and R. Wright (2005) “Search Theoretic Models of the Labor Market: A Survey,” *Journal of Economic Literature*, 43: 959-988.
- [25] Sullivan, P. (2010) “A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search,” *International Economic Review* 51: 289-317.

Table 1: Descriptive Statistics: CPS
 Schooling Level

	All	H	C
$P(s = C)$	0.3003	0	1
$P(U)$	0.0373	0.0467	0.0216
$\mathbb{E}[t_U U]$	20.62	20.3858	17.3510
$\mathbb{V}[t_U U]$	25.30	25.50	22.16
$\mathbb{E}[w E]$	17.94	15.34	23.40
$\mathbb{V}[w E]$	108.15	62.60	160.00

Table 2: Descriptive Statistics: SIPP

$\mathbb{E}[\Delta \log(w)^2]$	0.3224
$\mathbb{V}[\Delta \log(w)^2]$	1.7865

Parameter	Estimate
μ_a	3.0602 (0.0029546)
σ_a	0.20383 (0.0043648)
σ_θ	0.56521 (0.0059142)
α	0.55307 (0.0023914)
λ	0.19776 (0.0063379)
f	0.19388 (0.0061649)
η	0.0074907 (0.00056976)
b	-7.8959 (0.86023)
h_2	1.0872 (0.0019078)
θ^*	0.31197 (0.003111)
a^*	3.162 (0.090517)
c	160.5618 (3.4964)

Table 3: Homogenous Model Estimates

Moment	Data	Model
q_{10}^1	2.0794	2.0655
q_{20}^1	2.2513	2.2355
q_{30}^1	2.3843	2.3678
q_{40}^1	2.4865	2.4857
q_{50}^1	2.621	2.6008
q_{60}^1	2.7081	2.7207
q_{70}^1	2.8472	2.8498
q_{80}^1	2.9957	3.0093
q_{90}^1	3.2189	3.2339
q_{10}^2	2.3922	2.492
q_{20}^2	2.628	2.6591
q_{30}^2	2.7788	2.7882
q_{40}^2	2.9065	2.9066
q_{50}^2	3.0288	3.02
q_{60}^2	3.157	3.1414
q_{70}^2	3.277	3.266
q_{80}^2	3.442	3.4259
q_{90}^2	3.6743	3.6505
$\mathbb{E}[t_U]$	20.6159	20.6318
$P[U]$	0.037258	0.037199
$\mathbb{E}[\Delta \log(w)_i^2]$	0.32242	0.36817

Table 4: Homogenous Model Fit of Moments

Note: Let q_p^s denote the p th percentile of the wage distribution for schooling group s .

Parameter	Estimate
μ_a	3.0597 (0.0032551)
σ_a	0.19978 (0.00030664)
σ_θ	0.54987 (0.00026628)
α_1	0.55345 (0.0002546)
α_2	0.54826 (0.00010708)
λ_1	0.1876 (0.00016853)
λ_2	0.21086 (0.00041446)
η_1	0.0068219 (8.7729e-06)
η_2	0.0080187 (2.1549e-05)
b	-7.8666 (0.0015153)
h_2	1.0793 (1.4115e-05)
θ_0^*	0.31166 (0.00059555)
θ_1^*	0.31369 (0.00028504)

Table 5: Heterogeneous Model Estimates - Different Bargaining Powers

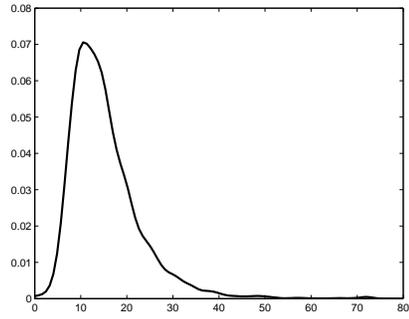
Parameter	Estimate
μ_a	3.066 (0.0026925)
σ_a	0.16519 (0.00031046)
σ_θ	0.58339 (0.00079219)
α_1	0.5547 (0.00029747)
α_2	0.5547 (0.00029747)
λ_1	0.22674 (0.00030753)
λ_2	0.25002 (0.00019077)
f_1	0.22282 (0.00033436)
f_2	0.24293 (0.00029233)
η_1	0.0076237 (1.8526e-06)
η_2	0.0082426 (5.4371e-06)
b	-9.3934 (1.0508e-07)
h_2	1.0757 (7.3133e-05)
θ_0^*	0.29133 (4.8033e-05)
θ_1^*	0.32899 (0.0008078)
a^*	3.1525 (0.062576)
c	366.028 (4.0624)

Table 6: Heterogeneous Model Estimates - Same Bargaining Power

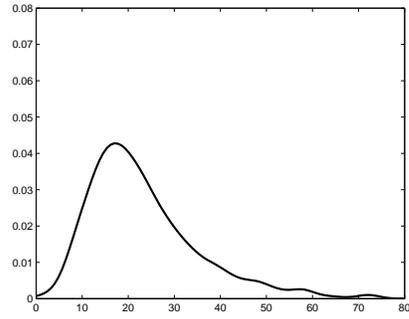
Moment	Data	Model
q_{10}^1	2.0794	2.0623
q_{20}^1	2.2513	2.2382
q_{30}^1	2.3843	2.3757
q_{40}^1	2.4865	2.4959
q_{50}^1	2.621	2.6156
q_{60}^1	2.7081	2.7396
q_{70}^1	2.8472	2.8698
q_{80}^1	2.9957	3.0318
q_{90}^1	3.2189	3.267
q_{10}^2	2.3922	2.48
q_{20}^2	2.628	2.6447
q_{30}^2	2.7788	2.7725
q_{40}^2	2.9065	2.8927
q_{50}^2	3.0288	3.0018
q_{60}^2	3.157	3.1202
q_{70}^2	3.277	3.2498
q_{80}^2	3.442	3.4039
q_{90}^2	3.6743	3.6217
$\mathbb{E}[t_U \mid s = 1]$	20.3858	17.9515
$\mathbb{E}[t_U \mid s = 2]$	17.351	16.4655
$\mathbb{E}[U \mid s = 1]$	0.0467	0.033082
$\mathbb{E}[U \mid s = 2]$	0.021565	0.032816
$\mathbb{E}[\Delta \log(w)_i^2]$	0.32242	0.35573

Table 7: Heterogeneous Model Fit of Moments

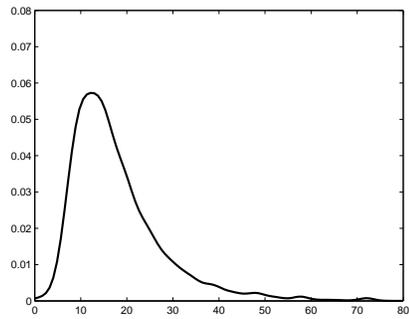
Note: Let q_p^s denote the p th percentile of the wage distribution for schooling group s .



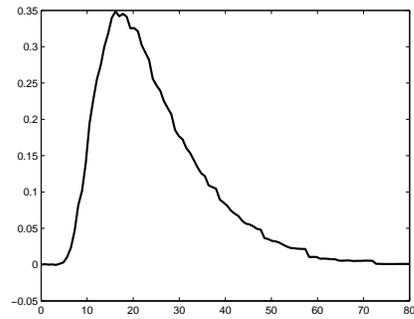
(a) Wage Distribution Low Schooling



(b) Wage Distribution High Schooling



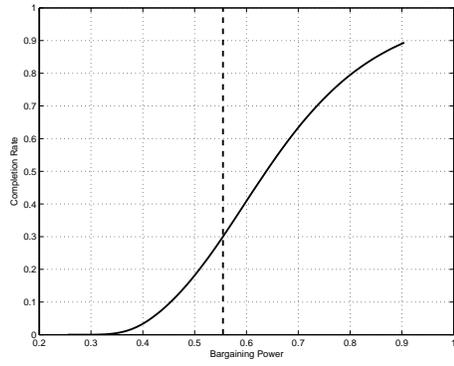
(c) Wage Distribution Total Sample



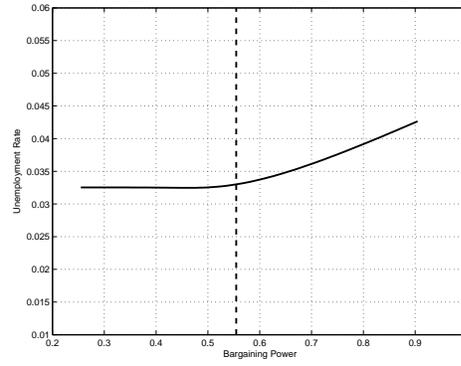
(d) Difference in Conditional CDFs: $F(w | s = 1) - F(w | s = 2)$.

Figure 1: Wage Distributions

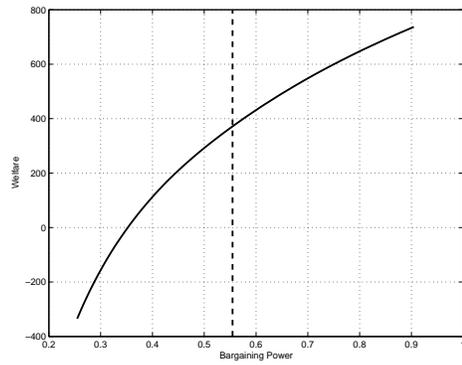
Note: Densities are estimated using a Gaussian kernel with bandwidth set using Silverman's rule of thumb.



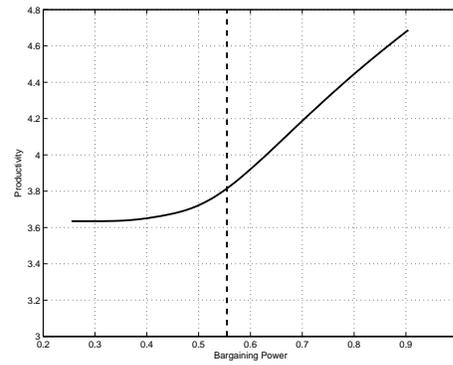
(a) College Completion Rate



(b) Unemployment



(c) Steady State Welfare



(d) Productivity

Figure 2: Effects in Partial Equilibrium of Changing Bargaining Power

Note: Dashed line indicates the estimated bargaining power for the model

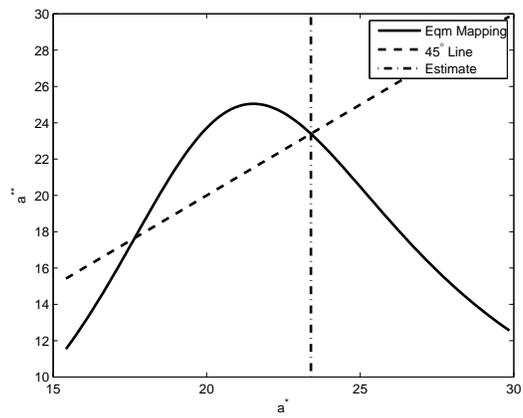
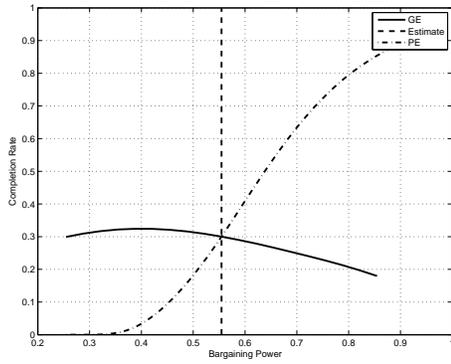
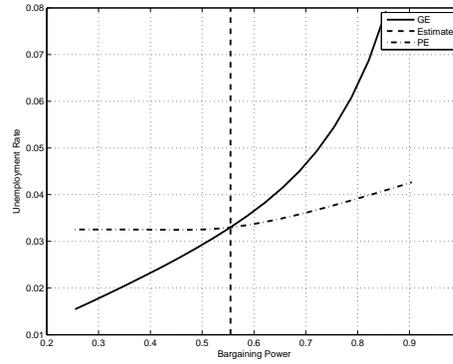


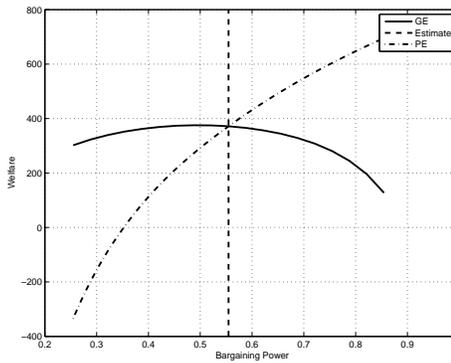
Figure 3: Multiple Equilibria in the Heterogeneous Model



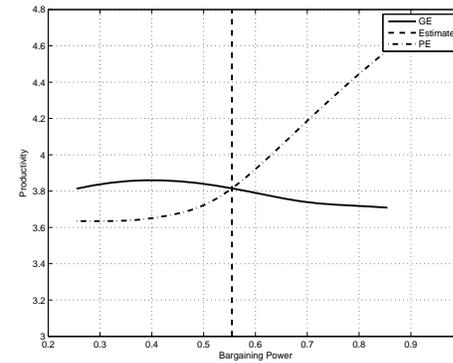
(a) College Completion Rate



(b) Unemployment



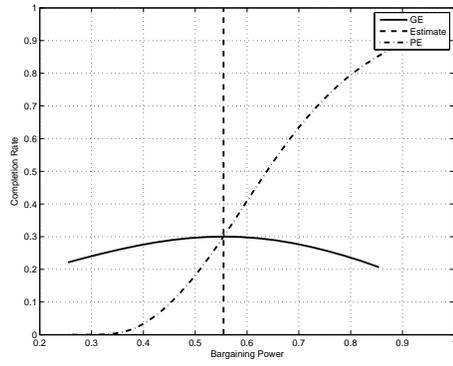
(c) Steady State Welfare



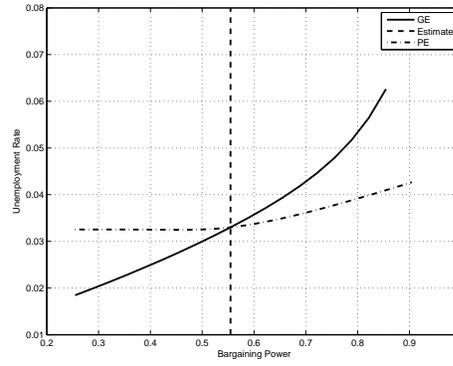
(d) Productivity

Figure 4: Effects in General Equilibrium of Changing Bargaining Power: $\delta = 0.4$

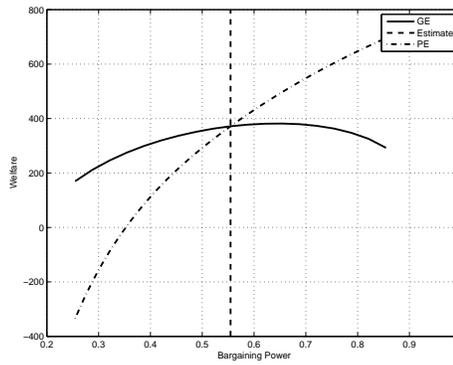
Note: Dashed line indicates the estimated bargaining power for the model



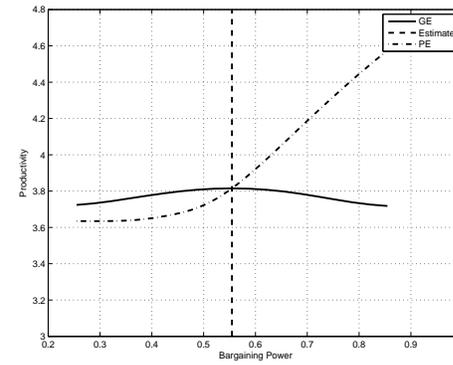
(a) College Completion Rate



(b) Unemployment



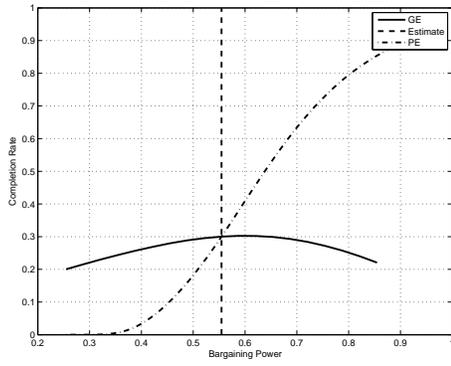
(c) Steady State Welfare



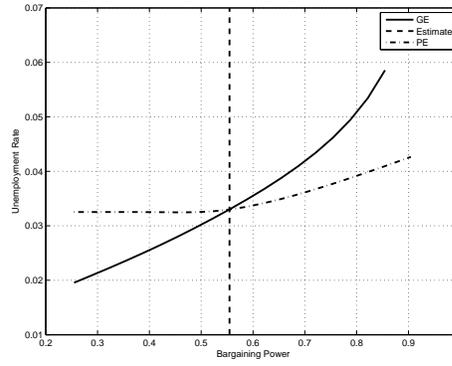
(d) Productivity

Figure 5: Effects in General Equilibrium of Changing Bargaining Power: $\delta = \hat{\alpha} = 0.55$

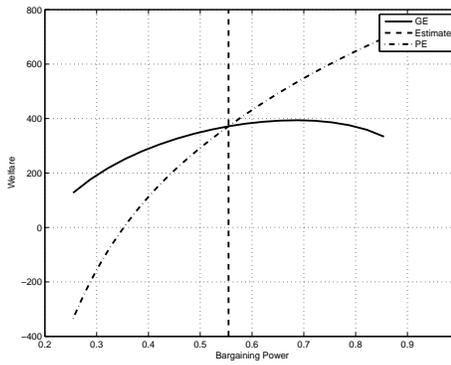
Note: Dashed line indicates the estimated bargaining power for the model



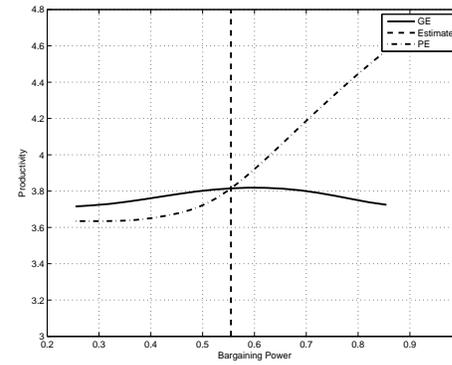
(a) College Completion Rate



(b) Unemployment



(c) Steady State Welfare



(d) Productivity

Figure 6: Effects in General Equilibrium of Changing Bargaining Power: $\delta = 0.6$

Note: Dashed line indicates the estimated bargaining power for the model