

LABOR MARKET SEARCH AND SCHOOLING INVESTMENT*

BY CHRISTOPHER FLINN AND JOSEPH MULLINS¹

New York University, U.S.A., and Collegio Carlo Alberto, Moncalieri, Italy; New York University, U.S.A.

We generalize a search, matching, and bargaining model to allow individuals to acquire productivity-enhancing schooling prior to labor market entry. In general, search frictions and weakness in bargaining position contribute to underinvestment in schooling from an efficiency perspective. Using estimates of a general equilibrium version of the model in which firm vacancy creation decisions are included, we find that minimum wages and schooling subsidies improve aggregate welfare, but have very different welfare impacts across the ability distribution. In particular, policies that maximize the average welfare of workers have strongly negative effects on the welfare of the least able.

1. INTRODUCTION

A large number of papers, both theoretical and applied, have examined labor market phenomena using the search and matching framework, with some embedded in a simple general equilibrium setting.² 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 article, we extend the standard search and matching framework to allow for endogenous schooling decisions.³ This is an important generalization of the model, since the large changes in the labor market over the past several decades can be expected to have had major impacts on schooling decisions.⁴ The college enrollment rate of high school graduates was 30% in 1973, whereas it was 48% in 2009. Coincidental with this change was an increase in the college premium and inequality in the earnings distribution. Although our model is set in a stationary environment, we are able to look at the manner in which labor market conditions influence and are influenced by schooling decisions.

*Manuscript received March 2013; revised April 2014.

¹ The C.V. Starr Center for Applied Economics at New York University has partially funded this research. James Mabli greatly aided in the development of the modeling strategy. We have received very helpful comments from two anonymous referees as well as from Cristian Bartolucci, Fabien Postel-Vinay, participants in seminars at Collegio Carlo Alberto, CMU, Washington University, ASU, and participants in the SOLE Meetings in London, June 2010 and 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. Please address correspondence to: Christopher Flinn, Department of Economics, New York University, 19 W. 4th Street, New York, NY 10012. Phone: 212-998-8925. Fax: 212-995-3932. E-mail: christopher.flinn@nyu.edu.

² A large number of macroeconomic labor applications are cited in Pissarides (2000) and the recent survey by Shimer et al. (2005). In terms of econometric implementations of the model, examples are Flinn and Heckman (1982), Eckstein and Wolpin (1990), Postel-Vinay and Robin (2002), Dey and Flinn (2005), Cahuc et al. (2006), and Flinn (2006).

³ There are a number of ambitious empirical papers that 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 holdup problem since they are based on a competitive labor market assumption. Eckstein and Wolpin (1995) 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 decision.

⁴ For example, a recent article by Gemici and Wiswall (2014) shows that a significant factor in the “college premium” over the past several decades is the changing mix of majors over this period.

Recently, there has been an upsurge of political interest in increasing the minimum wage (and even of indexing it to the consumer price level). The impact of a minimum wage has been investigated using a search framework in models of wage posting with firm heterogeneity (van den Berg and Ridder, 1998) and models with bargaining or surplus division with heterogeneity in match productivity (e.g., Flinn, 2006; Flinn and Mabili, 2009). Although the general equilibrium framework of Flinn (2006) can be used to investigate the impact of minimum wages on labor market participation rates, he assumes that individuals are ex ante identical and make no human capital investment decisions prior to entering the labor market.⁵ Although there have been a number of empirical papers on the impact of minimum wages on schooling, all of this research has been conducted outside of an explicit model of labor market behavior (see, e.g., Cunningham, 1981; Ehrenberg and Marcus, 1982; Neumark and Wascher, 1995; and Neumark and Nizalova, 2007). These studies have come to somewhat different conclusions regarding the linkage between the minimum wage and enrollment decisions, in part due to differences in data sets, time periods analyzed, and econometric methodology. In general, however, increases in the minimum wage have been found to be related to decreases in the enrollment rate. As Neumark and Wascher (2003) conclude, “the question of how and for whom minimum wages affect human capital formation merits further scrutiny,” and this is one of the goals of our analysis.

Over the last several decades, the cost of acquiring higher education has risen precipitously when compared with inflation as measured by the generic CPI-U. For example, from 1978 through 2009, inflation-adjusted college tuition has tripled. The impact of these price increases on labor market outcomes, such as unemployment and earnings inequality, can only be considered within a behavioral model such as the one developed and estimated below. We are able to examine the impact of changes in the price of higher education, a parameter that is estimated in our model, on labor market outcomes.

In order to examine the effects of minimum wages and college costs on labor market outcomes, we build 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. The individual's productivity at a firm is determined by her initial ability, her schooling level, and an idiosyncratic match productivity value drawn from a distribution G . As is standard, we utilize a surplus division rule to determine wages and the vacancy creation decisions of firms.

The fact that individuals make schooling decisions, which increase their subsequent labor market productivity prior to entering the labor market, leads to a classic holdup problem. There is a long-standing literature examining the essence of the holdup problem and the role contracts play to reduce, or altogether avoid, holdup (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, as a result, greater market frictions generally lead to more serious holdup problems. Acemoglu and Shimer (1999) examine the potential for holdup problems in frictional markets and investigate the manner in which markets can internalize the resulting externalities. Their focus is on identifying ways in which holdup 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. In our view, the commitment requirements for these types of contracts are a serious impediment to their implementation. This leads us to examine the holdup problem in a quasi-random search and matching framework.

⁵ The Flinn (2006) framework assumes i.i.d. match draws and ex ante homogeneous searchers. The “marriage” setup used by Postel-Vinay and Robin (2002) and Cahuc et al. (2006) assumes that the flow productivity of a match between worker i and firm j is given by $a_i b_j$, with distributions F_a and F_b of these types in the population. The distribution of types is considered a primitive, not impacted by the structure of the labor market or labor market policies, such as the minimum wage.

The generalized Nash bargaining power parameter has a direct impact on the extent of the holdup problem the worker faces vis-a-vis pre-market schooling investment decisions. Although 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, the nature of the renegotiation process, and 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 Mably (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, a particular bargaining protocol. 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. Although the holdup 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). Our estimates of the surplus division parameter are consistent with those reported in the small number of studies that attempt to estimate this class of models.

As our model structure makes clear, simply estimating separate behavioral models of the labor market for different schooling classes is at a minimum inefficient, and, more seriously, may lead to misinterpretations of labor market structure. For this reason, whenever possible, potentially endogenous individual characteristics acquired before or after entry into the labor market should be incorporated into the structure of the search, matching, and bargaining model. In order to do so in a tractable manner requires stringent assumptions regarding the productivity process, bargaining, etc., as is evident in what follows. Using our simple and reasonably tractable model, we are able to make some preliminary judgments regarding the impact of holdup on schooling investment. We find that policies that attempt to redistribute the surplus between firms and workers, such as minimum wages and schooling subsidies (to individuals), tend to promote more schooling investment and actually lead to welfare improvements for workers and efficiency gains for the economy. This is due to the fact that these policies reduce vacancy creation by firms, which actually leads to preferable equilibrium outcomes given our model estimates.

This article is organized as follows. In Section 2, we develop a bargaining model in a partial equilibrium framework, with education decisions made prior to entering the labor market. Section 3 extends the basic model to allow schooling submarkets to be characterized by different vectors of primitive parameters, such as contact and dissolution rates. In Section 4, we generalize the model to include on-the-job search. Section 5 considers the impact of a minimum wage within a partial equilibrium model of the labor market. Section 6 puts the model into a general equilibrium setting using the Mortensen-Pissarides matching function adapted to the two labor market case. In Section 7 we discuss the data used in the estimation, identification of model parameters, and define the estimator. In Section 8 we discuss the parameter estimates, and Section 9 contains the results of our policy experiments with minimum wages and schooling subsidies. Section 10 concludes.

2. HOMOGENEOUS SCHOOLING MARKETS

We assume that flow match productivity between a worker i and a firm j is given by

$$y_{ij} = v_i \theta_{ij},$$

where v_i is worker i 's productivity type and θ_{ij} is the match productivity value between the worker and the firm, which is assumed to be independently and identically distributed according to the distribution function G .⁶ The extension of our article is to allow the v distribution to be endogenous, with its form determined by the (exogenous) distribution of initial productivity endowments, F , a distribution of match-specific productivities, G , and all other parameters characterizing the labor market environment, Ω . The mapping from F into the distribution of v is reasonably straightforward. Schooling decisions are binary, with an individual i with initial ability a_i either deciding to go to school ($s_i = 1$) or not ($s_i = 0$). Her ability level after schooling is given by $v_i = h_s a_i$, where $h_0 = 1$ and $h_1 > 1$. When there is no danger of confusion, we will simply define $h \equiv h_1$; then $h - 1 > 0$ measures the growth in any individual's ability by completing additional schooling. We will show that the decision to complete higher education has a critical value property, where an individual with ability level a_i completes higher education if and only if $a_i \geq a^*(c, G, \Omega)$, where c denotes the common cost of completing higher education. The distribution of v is then determined by F , G , and Ω . Changes in any of these arguments, particularly in G or Ω , will lead to changes in the ability distribution of agents entering the market.

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 throughout the remainder of the article.)
2. The flow value of unemployment to a type a individual with schooling level s is given by

$$b(a, s) = bah_s.$$

This last assumption is similar to that made in Postel-Vinay and Robin (2002) and in Bartolucci (2013).

To fix ideas, we begin by briefly describing the model with no on-the-job search, since in this case we are able to derive simple and transparent results. 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 $v \equiv ah_s$, and the value of unemployed search to such an individual is given by $V_U(v)$. In terms of the surplus division problem, the worker–firm pair solves

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

where

$$V_E(w, v) = \frac{w + \eta V_U(v)}{\rho + \eta}$$

⁶ The analyses of Postel-Vinay and Robin (2002) and Cahuc et al. (2006) assume that flow productivity is given by $y_{ij} = a_i b_j$, where a and b denote the worker's and firm's productivity type, respectively. Dey and Flinn (2005) and Flinn (2006) assume that $y_{ij} = \theta_{ij}$. These authors adopt different specifications of flow productivity in large part due to the nature of the data with which they work. In the case of the first two papers, matched worker–firm information is available, enabling the authors to identify distributions of worker and firm types nonparametrically. The other papers rely on supply side data only and do not attempt to estimate the distribution of firm types. We only use “supply side” data but with the repeated measurements we have available argue that we can separately identify the distributions of match and individual-specific productivities.

$$V_F(w, \theta, v) = \frac{\theta v - w}{\rho + \eta}.$$

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 (FEC) that drives the value of an unfilled vacancy to 0.⁷ The solution to the surplus division problem yields

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

and since

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

we have

$$(1) \quad w = v(\alpha\theta + (1 - \alpha)\theta^*(v)).$$

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

$$(2) \quad \begin{aligned} \rho V_U(v) &= bv + \lambda \int_{\theta^*(v)} (V_E(v, \theta) - V_U(v)) dG(\theta) \\ \Rightarrow v\theta^*(v) &= bv + \frac{\lambda\alpha v}{\rho + \eta} \int_{\theta^*(v)} (\theta - \theta^*(v)) dG(\theta). \end{aligned}$$

Since this last equation is independent of v , we have

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

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

$$(3) \quad y^*(a, s) = ah_s\theta^*.$$

This result makes the analysis 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 0 as the baseline, that is, it is a required level of schooling for all individuals. To complete schooling level 1 instead, a total cost of c must be incurred, which we assume to be the same for all individuals. Individuals base their decision to complete schooling level 1 on the comparison between the values of entering the labor market with human capital h_0 or h_1 . Then an individual with ability level a will attend school if

$$\begin{aligned} \rho^{-1}ah_1\theta^* - \rho^{-1}ah_0\theta^* &\geq c \\ \Rightarrow a &\geq \frac{\rho c}{(h_1 - h_0)\theta^*}. \end{aligned}$$

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

$$a^* = \frac{\rho c}{(h_1 - h_0)\theta^*},$$

⁷ This assumption is utilized below when we generalize the model to allow for endogenous vacancy creation by firms.

with an individual of type $a \geq a^*$ acquiring schooling and a person $a < a^*$ not acquiring further schooling.⁸ In the empirical analysis conducted below, we assume that the support of the distribution of a is R_+ , so that the proportion of individuals in the population who acquire schooling is given by $1 - F(a^*) > 0$.

2.1. *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 the reason we go through this modeling specification.

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

$$P_1 \equiv P(s = 1) = \tilde{F}(a^*),$$

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

1. $\partial P_1 / \partial c < 0$. The proportion of the population attending college is decreasing in the direct costs of college attendance.
2. $\partial P_1 / \partial h > 0$. This is perhaps the most intuitive result. The greater the impact on labor market productivity, the greater the measure of individuals who complete college.
3. $\partial P_1 / \partial \theta^* > 0$. 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 most of various labor market parameters on the schooling decision.
 - (a) $\partial P_1 / \partial \lambda > 0$. An increase in the arrival rate of offers increases θ^* and hence increases the value of having a higher productivity distribution.
 - (b) $\partial P_1 / \partial \eta < 0$. An increase in the (exogenous) separation rate decreases θ^* and hence decreases the value of becoming more productive when matched with an employer.
 - (c) $\partial P_1 / \partial b > 0$. An increase in the “baseline” flow value of occupying the unemployment state increases the value of that state and the value of going to college.
4. $\partial P_1 / \partial \alpha > 0$. An increase in the worker’s share increases investment.

The last result is one of the main focuses of our attention, and it is intuitive. It is well known that the efficient level of productivity is obtained by giving the investing agents all of the surplus from their investment, which is the case as $\alpha \rightarrow 1$. With endogenous vacancy creation, a value of $\alpha = 1$ would effectively shut down the labor market and could never be an optimal choice with respect to worker welfare or efficiency. We return to this issue below.

⁸ In an earlier version of the model, we explicitly accounted for the difference in labor market entry dates under the two schooling options. In such a case, with instantaneous cost \tilde{c} , the total discounted value of continuing in higher education is $\exp(-\rho\tau)\rho^{-1}ah_1\theta^*$, 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 1 if

$$\exp(-\rho\tau)\rho^{-1}ah_1\theta^* - \rho^{-1}ah_0\theta^* \geq c,$$

and the condition for schooling to be acquired by any a when $c > 0$ is that $\exp(-\rho\tau)h_1 - 1 > 0$. Given that ρ and τ are fixed in our analysis, this puts a strong restriction on the estimate of h_1 for there 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 with search frictions.

2.2. *Empirical Implications.* This simple model produces some counterfactual empirical implications. First, in the case of unemployment experiences, we see that the likelihood of finding a type ν individual unemployed is

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

Thus, the assumption that the primitive parameters are identical across schooling groups produces the implication that there is no difference in unemployment experiences across schooling groups. In the data, we know that individuals with less schooling are more likely to be found in the unemployment state.

In terms of the wage distributions by schooling level, we do find systematic differences, of course. 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 (1) we know that

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

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

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

and the corresponding conditional wage density is given by

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

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

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

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

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

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 in the low-schooling group is

$$(4) \quad f(w|s=0) = \frac{1}{\alpha \tilde{G}(\theta^*)} \int_{\underline{a}}^{a^*} a^{-1} g\left(\alpha^{-1}\left(\frac{w}{a} - (1 - \alpha)\theta^*\right)\right) \frac{dF(a)}{F(a^*)}, \quad w \geq \underline{a}\theta^*,$$

where \underline{a} is the lowest value of a in the population, while the density of wages in the high-schooling population is

$$(5) \quad f(w|s=1) = \frac{1}{\alpha \bar{a} \tilde{G}(\theta^*)} \int_{a^*}^{\bar{a}} a^{-1} g\left(\alpha^{-1}\left(\frac{w}{ah} - (1 - \alpha)\theta^*\right)\right) \frac{dF(a)}{\bar{F}(a^*)}, \quad w \geq a^* h \theta^*,$$

where \bar{a} is the greatest value of a in the population.

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 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 $a^*h\theta^*$ for those with college. Since $a^*h > \underline{a}$, the lower support of the distribution of the college wage distribution lies strictly to the right of the high school wage distribution. It is straightforward to establish that $F(w|s = 0) \geq F(w|s = 1)$ for all $w \geq \underline{a}\theta^*$, so that the wage distribution of the college educated first-order stochastically dominates those who are not college educated. This implication is generally consistent with cross sectional wage distributions drawn from the Current Population Survey, for example.

3. SEPARATE SCHOOLING SUBMARKETS

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 submarkets are often found to be markedly different (e.g., Flinn, 2002). In particular, it is often noted that the unemployment rate differs across schooling groups, with those with lower completed schooling 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 separate 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 . 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 $y(a, s, \theta) = ah_s\theta = v\theta$, so that the distribution of y is a function of the distribution of the scalar v 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 submarkets. Because primitive parameters differ across markets, v is no longer a sufficient statistic for the value of search of an individual; instead, a minimal sufficient statistic is the pair (v, s) . This is clear if we reconsider the functional equation determining the value of search in the homogeneous submarkets case, which was given in (2), adapted to the heterogeneous case. We now have

$$v\theta_s^*(v) = bv + \frac{\lambda_s\alpha_s v}{\rho + \eta_s} \int_{\theta_s^*(v)} (\theta - \theta_s^*(v))dG(\theta).$$

The solution $\theta_s^*(v)$ now clearly is independent of v , as before, but is not independent of s . Thus, there is a common critical value θ_s^* shared by all individuals with schooling choice s , which is independent of their ability a (conditional on s).

The critical match value for an individual of type a in schooling market s is given by $ah_s\theta_s^* = v\theta_s^*$, so that the value of unemployed search in this submarket is given by $\rho^{-1}v\theta_s^*$. The net value of college education to an individual of type a is

$$\rho^{-1}ah\theta_1^* - c - \rho^{-1}a\theta_0^*,$$

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

$$a^* = \frac{c\rho}{h\theta_1^* - \theta_0^*}.$$

For a^* to be positive, it is necessary that $h\theta_1^* - \theta_0^* > 0$.

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 submarket. By simple extension of the homogeneous results above, the results regarding $\partial P_1/\partial c < 0$ remain the same since the cost structure of acquiring schooling is identical in the two cases. It is also clearly the case that $\partial P_1/\partial h > 0$. The main departure from the previous case regards the presence of θ_0^* and θ_1^* . We note that

1. $\partial P_1/\partial \theta_0^* < 0$. As before, θ_0^* is not a primitive parameter, but the primitive parameters specific to submarket 0 only affect the schooling decision through θ_0^* . Then
 - (a) $\partial P_1/\partial \lambda_0 < 0$. Increases in the arrival rate of offers in the low-schooling market increase θ_0^* , and increase the relative value of the low-schooling option.
 - (b) $\partial P_1/\partial \eta_0 > 0$. Such an increase decreases the value of a low-schooling option.
2. $\partial P_1/\partial \theta_1^* > 0$.
 - (a) $\partial P_1/\partial \lambda_1 > 0$.
 - (b) $\partial P_1/\partial \eta_1 < 0$.
3. Perhaps most interesting is the impact of market-specific bargaining powers α_s on the schooling decision. When there is one bargaining power parameter that holds throughout all educational labor markets, the meaning of holdup is relatively unambiguous. When there are market-specific bargaining power parameters, a relative notion of holdup is more appropriate. Clearly we have

$$\frac{\partial P_1}{\partial \alpha_0} < 0.$$

$$\frac{\partial P_1}{\partial \alpha_1} > 0.$$

It is important to note that α_1 could be quite low, and yet a substantial proportion of agents may choose the high schooling level if α_0 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. The unemployment rate is now

$$P(U|v, s) = \frac{\eta_s}{\eta_s + \lambda_s \tilde{G}(\theta_s^*)} = P(U|s), \quad s = 0, 1.$$

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 levels.

In terms of schooling-specific wage distributions, the lower bound on the support of the wage distribution associated with schooling type s is now given by $\underline{w}(0) = \underline{a}\theta_0^*$ for the low-schooling group and by $\underline{w}(1) = a^*h\theta_1^*$ for the high-schooling group. The conditional density of wages for the low-schooling group is given by

$$f(w|a, s = 0) = \frac{1}{\alpha_0 a} \frac{g(\alpha_0^{-1}(\frac{w}{a} - (1 - \alpha_0)\theta_0^*))}{\tilde{G}(\theta_0^*)}, \quad w \geq \underline{a}\theta_0^*,$$

while the wage density for the high-schooling group is

$$f(w|a, s = 1) = \frac{1}{\alpha_1 h a} \frac{g(\alpha_1^{-1}(\frac{w}{ah} - (1 - \alpha_1)\theta_1^*))}{\tilde{G}(\theta_1^*)}, \quad w \geq a^*h\theta_1^*.$$

Since the model with heterogeneous schooling submarkets continues to imply that those who continue to schooling level $s = 1$ form a connected set $[a^*, \infty)$, the unconditional (on a) wage densities have the simple forms

$$f(w|s = 0) = \frac{1}{\alpha_0 \tilde{G}(\theta_0^*)} \int_{\underline{a}}^{a^*} a^{-1} g\left(\alpha_0^{-1} \left(\frac{w}{a} - (1 - \alpha_0)\theta_0^*\right)\right) \frac{dF(a)}{F(a^*)}, \quad w \geq \underline{a}\theta_0^*,$$

and

$$f(w|s = 1) = \frac{1}{\alpha_1 h \tilde{G}(\theta_1^*)} \int_{a^*} a^{-1} g\left(\alpha_1^{-1} \left(\frac{w}{ah} - (1 - \alpha_1)\theta_1^*\right)\right) \frac{dF(a)}{\tilde{F}(a^*)}, \quad w \geq a^* h \theta_1^*.$$

With the lack of restrictions on the orderings of the primitive parameters across the two markets, it is not possible to obtain the implication that $F(w|s = 1)$ first order stochastically dominates $F(w|s = 0)$, which was true in the homogenous markets case.

4. ON-THE-JOB SEARCH

We have gone through the previous two specifications of the model without OTJ search in an attempt to build some intuition regarding the relationship between schooling outcomes and the primitive parameters characterizing the search environment. However, we know that there are a large number of job-to-job transitions that are not consistent with the assumption that no offers are received while currently employed. This generalization is also critical for any analysis of the holdup problem, since as Postel-Vinay and Robin (2002), Dey and Flinn (2005), and Cahuc et al. (2006) have shown, under certain assumptions regarding bargaining or surplus division protocols, bidding between employers competing for the services of a worker potentially allows the individual to obtain a large share of the surplus even when her notional bargaining power (α) is low. Even when firms do not respond to outside offers, having the option to receive offers when employed increases the value of unemployed search for each schooling group except when workers are already receiving all of the surplus of the match. For simplicity, and due to identification problems when taking the model to the data, we will henceforth assume that both schooling markets share a common notional surplus division parameter, $\alpha = \alpha_0 = \alpha_1$.

We now consider two specifications of the bargaining problem, one in which there is competition between two employers for the services of a single individual (which occurs at a measure-0 of times during the individual's labor market career). This is what we refer to as the renegotiation case. We then describe the situation in which firms do not respond to offers from other firms, either because it is a strategic decision not to do so or because there is no credible way to convey such offers. This is referred to as the no-renegotiation case. Although the payoffs to workers and firms will differ in the two cases, both imply efficient mobility decisions, that is, the firm at which the individual is more productive will always obtain the services of the individual.

In discussing both bargaining environments, we allow the primitive parameters to vary across schooling markets. The contact rate for an unemployed searcher with schooling s will be denoted $\lambda_{U,s}$, the contact rate for an employed searcher is given by $\lambda_{E,s}$, and the exogenous destruction rate remains η_s .

4.1. Renegotiation. In the renegotiation case, firms periodically compete over a given worker's services. Of course, the individual's ability ν is the same at both firms, but the match productivity will differ with probability 1 if θ has an absolutely continuous distribution. Due to the free entry condition discussed below, all firms have an outside option value of zero. When two firms compete for the same worker, their positions are symmetric conditional on the match productivity draws, that is, the current employer has no advantage or disadvantage in obtaining the services of its current employee with respect to the other firm. Let θ' and θ denote the two productivity draws at the firms, and let $\theta' > \theta$, in which case we call θ' the dominant value and θ

the dominated match value. The firms engage in Bertrand competition in terms of wage setting until the firm with the dominated match value θ drops out after offering a wage that leaves it with no surplus, which is $w = v\theta$. The winning firm (at which productivity is $v\theta'$) and worker divide the surplus using the value of working at the firm with match value θ at the wage $v\theta$ as the worker's outside option and zero as the firm's outside option and with labor receiving the proportion α of the surplus.

We denote the value of employment to a worker with match values θ' and θ with ability level a in schooling market s as $V_{E,s}(\theta', \theta, v)$. For the case in which the individual is coming from the state of unemployment, the "dominated" match value is simply given by the reservation match value for this individual, which for now we write as $\theta_s^*(v)$. First, consider the value of an individual's problem given the state variables (θ', θ, v, s) and the wage offer w , which we can write as

$$V_{E,s}(\theta', \theta, v; w) = \frac{w + \eta_s V_{U,s}(v) + \lambda_{E,s} \left\{ \int_{\theta}^{\theta'} V_{E,s}(\theta', \tilde{\theta}, v) dG(\tilde{\theta}) + \int_{\theta'} V_{E,s}(\tilde{\theta}, \theta', v) dG(\tilde{\theta}) \right\}}{\rho + \eta_s + \lambda_{E,s} \tilde{G}(\theta)},$$

where the first integral in the numerator is the gain in welfare attributable to an increase in the outside option, and the second term reflects welfare gains when a better match is discovered, in which case the current θ' becomes the "dominated" match value. In this case, there will be (efficient) turnover. The firm's value (for the ultimate employer of the worker, which is the firm at which the match value is θ') is

$$V_{F,s}(\theta', \theta, v; w) = \frac{v\theta' - w + \lambda_{E,s} \int_{\theta}^{\theta'} V_{F,s}(\theta', \tilde{\theta}, v) dG(\tilde{\theta})}{\rho + \eta_s + \lambda_{E,s} \tilde{G}(\theta)}.$$

In the case of Bertrand competition, the wage offer at the firm with the dominated match value will be $v\theta$, so that the value of remaining at the losing firm after all of the surplus has been extracted by the worker is

$$V_{E,s}(\theta, \theta, v) = \frac{v\theta + \eta_s V_{U,s}(v) + \lambda_{E,s} \int_{\theta} V_{E,s}(\tilde{\theta}, \theta, v) dG(\tilde{\theta})}{\rho + \eta_s + \lambda_{E,s} \tilde{G}(\theta)}.$$

Then we can write

$$V_{E,s}(\theta', \theta, v) = \max_w (V_{E,s}(\theta', \theta, v; w) - V_{E,s}(\theta, \theta, v))^\alpha V_{F,s}(\theta', \theta, v)^{1-\alpha},$$

and this leads us to the following expression for the wage:

$$w_s(\theta', \theta, v) = v(\alpha\theta' - (1 - \alpha)\theta) + \alpha\lambda_{E,s} \int_{\theta}^{\theta'} V_{F,s}(\theta', \tilde{\theta}, v) dG(\tilde{\theta}) - (1 - \alpha)\lambda_{E,s} \left[\int_{\theta}^{\theta'} V_{E,s}(\theta', \tilde{\theta}, v) dG(\tilde{\theta}) + \int_{\theta'} V_{E,s}(\tilde{\theta}, \theta', v) dG(\tilde{\theta}) - \int_{\theta} V_{E,s}(\tilde{\theta}, \theta, v) dG(\tilde{\theta}) \right].$$

Now, as in the case of no OTJ search, posit that $V_{E,s}(\theta', \theta, ah_s) = ah_s \bar{V}_{E,s}(\theta', \theta)$, $V_{F,s}(\theta', \theta, ah_s) = ah_s \bar{V}_{F,s}(\theta', \theta)$, and $V_{U,s}(ah_s) = ah_s \bar{V}_{U,s}$, in which case the wage equation is

$$\begin{aligned} w_s(\theta', \theta, v) &= v(\alpha\theta' - (1 - \alpha)\theta) + v\alpha\lambda_{E,s} \int_{\theta}^{\theta'} \bar{V}_{F,s}(\theta', \tilde{\theta}) dG(\tilde{\theta}) \\ &\quad - v(1 - \alpha)\lambda_{E,s} \left[\int_{\theta}^{\theta'} \bar{V}_{E,s}(\theta', \tilde{\theta}) dG(\tilde{\theta}) + \int_{\theta'} \bar{V}_{E,s}(\tilde{\theta}, \theta') dG(\tilde{\theta}) - \int_{\theta} \bar{V}_{E,s}(\tilde{\theta}, \theta) dG(\tilde{\theta}) \right] \\ &= vq_s(\theta', \theta), \end{aligned}$$

and we have

$$v\bar{V}_{E,s}(\theta', \theta) = \frac{vq_s(\theta', \theta) + \eta_s v\bar{V}_U + v\lambda_{E,s} \left[\int_{\theta}^{\theta'} \bar{V}_{E,s}(\theta', \tilde{\theta}) dG(\tilde{\theta}) + \int_{\theta'} \bar{V}_{E,s}(\tilde{\theta}, \theta') dG(\tilde{\theta}) \right]}{\rho + \eta_s + \lambda_{E,s} \tilde{G}(\theta)}$$

Then

$$\bar{V}_{E,s}(\theta', \theta) = \frac{q_s(\theta', \theta) + \eta_s \bar{V}_{U,s} + \lambda_{E,s} \left[\int_{\theta}^{\theta'} \bar{V}_{E,s}(\theta', \tilde{\theta}) dG(\tilde{\theta}) + \int_{\theta'} \bar{V}_{E,s}(\tilde{\theta}, \theta') dG(\tilde{\theta}) \right]}{\rho + \eta_s + \lambda_{E,s} \tilde{G}(\theta)}$$

$$\bar{V}_{F,s}(\theta', \theta) = \frac{\theta' - q_s(\theta', \theta) + \lambda_{E,s} \int_{\theta}^{\theta'} \bar{V}_{F,s}(\theta', \tilde{\theta}) d\tilde{G}(\tilde{\theta})}{\rho + \eta_s + \lambda_{E,s} G(\theta)}$$

and

$$\bar{V}_{U,s} = \frac{b + \lambda_{U,s} \int_{\theta_s^*} \bar{V}_{E,s}(\tilde{\theta}, \theta_s^*) dG(\tilde{\theta})}{\rho + \lambda_{U,s} \tilde{G}(\theta_s^*)}$$

In assessing the relationship between holdup and investment, allowing for OTJ search is a realistic and important generalization. When $\lambda_{E,s} = 0$, the share of the surplus received by the agent is α . When $\lambda_{E,s} > 0$, and with Bertrand competition between employers, the average share of the surplus received by the worker will exceed α , often substantially so.⁹

4.2. *No Renegotiation.* As we will see below, the Bertrand competition case implies parameter values that on the face of it seem unreasonable. As in Flinn and Mabli (2009), we also estimate the model under the assumption that no firms renegotiate with their employees. In this case, all wage bargaining on the job uses as an outside option the value of unemployed search for an individual of type v , which is an option that is always attainable by any employee whether or not contracts can be enforced.

In this case, the value of employment at a match value of θ is only a function of θ and v . Then

$$V_{E,s}(\theta, v; w) = \frac{w + \eta_s V_{U,s}(v) + \lambda_{E,s} \int_{\theta} V_{E,s}(\tilde{\theta}, v) dG(\tilde{\theta})}{\rho + \eta_s + \lambda_{E,s} \tilde{G}(\theta)}$$

and

$$(6) \quad V_{F,s}(\theta, v; w) = \frac{v\theta - w}{\rho + \eta_s + \lambda_{E,s} \tilde{G}(\theta)}$$

Then

$$V_{E,s}(\theta, v) = \max_w (V_{E,s}(\theta, v; w) - V_{U,s}(v))^\alpha V_{F,s}(\theta, v; w)^{1-\alpha},$$

which leads to the wage equation

$$(7) \quad w_s(\theta, v) = \alpha v\theta + (1 - \alpha) \left(\rho V_{U,s}(v) - \lambda_{E,s} \int_{\theta} [V_{E,s}(\tilde{\theta}, v) - V_{U,s}(v)] dG(\tilde{\theta}) \right),$$

⁹ In the limit, as $\lambda_U \rightarrow \infty$ and $\lambda_E \rightarrow \infty$, the labor model becomes competitive in that each worker will be located at the best match value $\tilde{\theta}$, the finite upper support of G , and will be paid $v\theta$. They will receive all of the surplus from this match as long as θ is continuously distributed on $[\theta, \tilde{\theta}]$.

where we have used the fact that there will be efficient mobility in the sense that a worker of type ν with a match productivity value of $\nu\theta$ will join a new firm where productivity is $\nu\theta'$ if and only if $\theta' > \theta$. Once again, posit that $V_{E,s}(\theta, \nu) = \nu\bar{V}_{E,s}(\theta)$, $V_{F,s}(\theta, \nu) = \nu\bar{V}_{F,s}(\theta)$, and $V_{U,s}(\nu) = \nu\bar{V}_{U,s}$, so that $\theta_s^*(\nu) = \nu\theta_s^*$. Then

$$\begin{aligned} w_s(\theta, \nu) &= \nu \left[\alpha\theta + (1 - \alpha) \left(\rho\bar{V}_{U,s} - \lambda_{E,s} \int_{\theta} [\bar{V}_{E,s}(\tilde{\theta}) - \bar{V}_{U,s}] dG(\tilde{\theta}) \right) \right] \\ &= \nu q_s(\theta), \end{aligned}$$

and the model is closed with

$$\bar{V}_{U,s} = \frac{b + \lambda_{U,s} \int_{\theta_s^*} \bar{V}_{E,s}(\theta) dG(\theta)}{\rho + \lambda_{U,s} \tilde{G}(\theta_s^*)}.$$

For either bargaining protocol, the schooling decision has the same form as in the case without OTJ search. In particular, we have that an individual will obtain higher education when $a \geq a^* = c/(h\bar{V}_{U,1} - \bar{V}_{U,0})$. For the same primitive parameters, the critical value a^* will vary between the renegotiation and no-renegotiation cases, since the values $\bar{V}_{U,1}$ and $\bar{V}_{U,0}$ will differ.

5. MINIMUM WAGES

We have developed the model with no constraints on the compensation decision of the worker and firm, although at some skill (ν) and match productivity (θ) levels statutory minimum wage constraints may bind. In this section we examine the manner in which a minimum wage, m , will impact the labor market within our partial equilibrium setting. For the most part, we follow Flinn and Mabl (2009) in the development of this model. Since we will use the estimates from the no-renegotiation OTJ search case in the sequel, we limit our discussion to the consideration of this bargaining environment.

As in Flinn (2006), the minimum wage is viewed solely as a constraint on the surplus division problem. A minimum wage m generally will impact the valuation of all labor market states, at least for a subset of the population. As before, the acceptance decision is characterized by a reservation value, with all match draws $\theta \geq \theta_s^*(\nu, m)$ being, in principle, acceptable. The minimum wage constraint can bind for an individual of type ν if and only if $\nu \times \theta_s^*(\nu; m) < m$, or $\theta_s^*(\nu; m) < m/\nu$. Ignoring the minimum wage constraint for the moment, we can write the wage equation (7) as

$$\tilde{w}_s(\theta, \nu; m) = \alpha\nu\theta + (1 - \alpha) \left[\rho V_{U,s}(\nu; m) - \lambda_{E,s} \int_{\theta} [V_{E,s}(\tilde{\theta}, \nu; m) - V_{U,s}(\nu; m)] dG(\tilde{\theta}) \right].$$

Now the wage equation $\tilde{w}_s(\theta, \nu; m)$ is increasing in θ , and for any individual with $\theta_s^*(\nu; m) < m/\nu$ there exists a unique $\hat{\theta}_s(\nu; m)$ defined by

$$m = \tilde{w}_s(\hat{\theta}, \nu; m).$$

For all values of $\theta \in [m/\nu, \hat{\theta}_s(\nu; m)]$ individuals with $\theta_s^*(\nu; m) < m/\nu$ will be paid the minimum wage. For notational convenience, for individuals not bound by the minimum wage constraint, that is, those for whom $\theta_s^*(\nu; m) \geq m/\nu$, we define $\hat{\theta}_s(\nu; m) = \theta_s^*(\nu; m)$. Then for a type ν individual the wage at a match value θ in minimum wage regime m is

$$w_s(\theta, \nu; m) = \begin{cases} m & \text{if } \theta_s^*(\nu; m) < m/\nu \text{ and } \theta \in [m/\nu, \hat{\theta}_s(\nu; m)] \\ \tilde{w}_s(\theta, \nu; m) & \text{if } \theta \geq \hat{\theta}_s(\nu; m). \end{cases}$$

In the Flinn and Mabli (2009) treatment of the OTJ search case, there is a degenerate ability distribution F and no schooling decision, so that match productivity is simply given by θ . Because individuals in their model are ex ante homogeneous, a binding minimum wage applies to all members of the population. In the case considered here, if v is unbounded, there will be a set of individuals for whom the minimum wage will not be binding, although as the minimum wage becomes indefinitely large eventually it will become binding for all individuals as defined by v . The minimum wage in this model has real distributional effects within the population of workers.

We will be particularly interested in the manner in which the minimum wage affects the schooling decision of individuals. By increasing one’s ability through the completion of higher education, the individual can partially escape the negative employment effects of the minimum wage by increasing the distribution of flow productivities she faces. More formally, we can write the value of unemployment for an individual of type a without schooling as

$$\rho V_{U,0}(a; m) = ba + \lambda_{U,0} \int_{\max\{\theta^*(a;m), m/a\}} [V_{E,0}(\theta, a; m) - V_{U,0}(a; m)] dG(\theta),$$

whereas if she obtains higher education

$$\rho V_{U,1}(ah; m) = bah + \lambda_{U,1} \int_{\max\{\theta^*(ah;m), m/(ah)\}} [V_{E,1}(\theta, ah; m) - V_{U,1}(ah; m)] dG(\theta).$$

We note that with the minimum wage “distortion,” it is no longer the case that the job acceptance decision is a function only of θ given the schooling level. That is why we cannot write $V_{U,s}(v; m)$ as $v\bar{V}_{U,s}(m)$. This makes the analysis of how the minimum wage affects schooling decisions more complex. An individual of type a will complete college under minimum wage m iff

$$V_{U,1}(ah; m) - V_{U,0}(a; m) \geq c.$$

For individuals with a large enough value of a , the minimum wage m will not be binding with or without schooling, so the decision rule simplifies to the ones considered previously. For those with low values of a , there will be values of the minimum wage that will be binding whether they complete college or not. Finally, there is a set of individual types for whom a minimum wage m will not be binding if they complete college, but will be if they do not. It is not possible to characterize these sets analytically, but we will conduct policy exercises in which we will be able to quantitatively assess the degree to which the imposition of a minimum wage impacts the schooling decisions of population members differentiated by their ability endowments.

6. ENDOGENOUS CONTACT RATES

In this section, we consider how the contact rates $(\lambda_{U,s}, \lambda_{E,s})$, $s = 0, 1$, are determined endogenously in a general equilibrium framework. We assume random search within each schooling market, but vacancies are created for certain schooling types and are available only to the particular schooling group for which they are posted. We adapt the standard Mortensen and Pissarides (1994) framework to this two market environment and allow firms to post vacancies v_s in schooling market s at a constant marginal cost ψ_s . Firms and workers are then matched with each other according to a constant returns to scale matching technology. Let U_s be the mass of unemployed workers of schooling type s and $E_s = 1 - U_s$ be the mass of employed workers in this schooling group. The flow of matches created, M_s , is given by

$$M_s = (U_s + \phi_s E_s)^{\delta_s} v_s^{1-\delta_s},$$

where δ_s is the Cobb–Douglas parameter in market s , $\delta_s \in (0, 1)$. The parameter ϕ_s , $0 < \phi_s \leq 1$, reflects the lower search efficiency of individuals who are currently employed relative to the unemployed. The rate of contacts per firm searching in market s is

$$q_s(k) = k_s^{\delta_s},$$

where $k_s = (U_s + \phi_s E_s)/v_s$, and k_s is a measure of market tightness. The proportion of searchers of schooling level s who are employed is given by $(\phi_s E_s)/(U_s + \phi_s E_s)$, so that the mass of matches in market s that involve an employed worker is simply $(\phi_s E_s)/(U_s + \phi_s E_s) \times M_s$, which means that the flow rate of contacts for the employed is

$$\begin{aligned}\lambda_{E,s} &= \frac{(\phi_s E_s)}{(U_s + \phi_s E_s)} \frac{(U_s + \phi_s E_s)^{\delta_s} v_s^{1-\delta_s}}{E_s} \\ &= \phi_s k_s^{\delta_s - 1}.\end{aligned}$$

A similar argument is used to find the contact rate for unemployed searchers in group s ,

$$\lambda_{U,s} = k_s^{\delta_s - 1}.$$

A fact that will be utilized in the estimation of demand side parameters below is that $\phi_s = \lambda_{E,s}/\lambda_{U,s}$.

The final components of the general equilibrium framework are the flow costs associated with vacancy creation, which are given by ψ_s , $s = 0, 1$. Firms can create vacancies in either market, and entry continues until the expected value of creating a vacancy is driven to zero in both markets. We will consider the form of this expression for the no-renegotiation case both because it is simpler and because we find empirical support for this specification. The value of the firm of being matched with an individual of type ah_s with a match draw of θ is $V_{F,s}(\theta, ah_s)$, which is given in (6). In the high-schooling market, we know that the minimum value of v is equal to a^*h , whereas in the low-schooling market the maximum value of v is a^* . Let the steady state distribution of match values among the employed in market s be given by $H_{E,s}(\theta)$, which has support $[\theta_s^*, \infty)$. Then the expected value of a vacancy in the high-schooling market is

$$(8) \quad -\psi_1 + \frac{k_1^{\delta_1}}{U_1 + \phi_1 E_1} \times \left(U_1 \int_{a^*} \int_{\theta_1^*} V_{F,1}(\theta, ah) dG(\theta) dF(a) \right. \\ \left. + \phi_1 E_1 \int_{a^*} \int_{\theta_1^*} \int_{\theta'} V_{F,1}(\theta, ah) dG(\theta) dH_{E,1}(\theta') dF(a) \right),$$

and the expected value of a vacancy in the low-schooling market is

$$(9) \quad -\psi_0 + \frac{k_0^{\delta_0}}{U_0 + \phi_0 E_0} \times \left(U_0 \int_{a^*} \int_{\theta_0^*} V_{F,0}(\theta, a) dG(\theta) dF(a) \right. \\ \left. + \phi_0 E_0 \int_{a^*} \int_{\theta_0^*} \int_{\theta'} V_{F,0}(\theta, a) dG(\theta) dH_{E,0}(\theta') dF(a) \right).$$

Under the free entry condition (FEC), both of these terms are equal to zero.

In the presence of a binding minimum wage, these expressions are altered in important ways. The expected value of posting a vacancy in the high-schooling market becomes

$$-\psi_1 + \frac{k_1^{\delta_1}}{U_1 + \phi_1 E_1} \times \left(U_1 \int_{a^*(m)} \int_{\max\{\theta_1^*(ah;m), m/(ah)\}} V_{F,1}(\theta, ah; m) dG(\theta) dF_{U,1}(a|m) \right. \\ \left. + \phi_1 E_1 \int_{a^*(m)} \int_{\max\{\theta_1^*(ah;m), m/(ah)\}} \int_{\theta'} V_{F,1}(\theta, ah; m) dG(\theta) dH_{E,1}(\theta' | ah, m) dF_{E,1}(a|m) \right),$$

where $H_{E,1}(\theta|ah, m)$ is the steady-state distribution of match values in the high-schooling market for individuals of type ah under the minimum wage m , and $F_{l,s}(a|m)$ is the steady-state distribution of a in labor market state l and schooling group s under the minimum wage m . The corresponding expression for the low-schooling market is

$$-\psi_0 + \frac{k_0^{\delta_0}}{U_0 + \phi_0 E_0} \times \left(U_0 \int \int_{\max\{\theta_0^*(a;m), m/a\}}^{a^*(m)} V_{F,0}(\theta, a; m) dG(\theta) dF_{U,0}(a|m) + \phi_0 E_0 \int \int_{\max\{\theta_0^*(a;m), m/a\}}^{a^*} \int_{\theta} V_{F,0}(\theta, a; m) dG(\theta) dH_{E,0}(\theta|a, m) dF_{E,0}(a|m) \right),$$

where $H_{E,0}(\theta|a, m)$ is the steady-state distribution of match values in the low-schooling market for individuals of type a under the minimum wage m . Under the FEC, vacancies will be determined at the points at which both these expressions equal zero, which clearly is a function of the minimum wage m .

By examining these expressions, we see that minimum wages affect the vacancies, and hence equilibrium contact rates, in a variety of ways. Most obviously, as in partial equilibrium, they restrict the set of match values that will result in an employment contract. Secondly, they alter the surplus division between the worker and firm, thus impacting $V_{F,s}(\theta, ah_s; m)$. Thirdly, they affect the composition of the low- and high-schooling populations through their impact on $a^*(m)$. We will see all of these forces at work in the quantitative exercises conducted below.

7. ECONOMETRIC ISSUES

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

7.1. Data Description. Our primary data source is the Survey of Income and Program Participation (SIPP). The SIPP is a nationally representative, household-based survey consisting of several longitudinal panels. Each panel lasts four years in total. The survey is administered in four-month “waves,” at which point information is collected retrospectively for the previous four months. Therefore, each panel contains 12 waves of the survey. Our analysis is based on information collected over a 12-month period from Waves 4, 5, and 6 of the 2004 panel. Since the first longitudinal month is not the same for all sample members, we observe individuals for the first time between October 2004 and January 2005 and we observe their final month between September 2005 and December 2005.

The sample characteristics we use in defining the estimator involve labor market states at the beginning (first month) and end (last month) of this 12-month period. In what follows, we index the beginning of the period as $t = 0$ and the end as $t = 1$. In addition to this comparison, we link individuals across the intermediate wave (Wave 5) in order to track their employment history. We use this to infer whether any given individual has experienced a spell of unemployment or a change in employer between $t = 0$ and $t = 1$.

Below we describe how information in the SIPP is coded, the restrictions we impose on the final sample, and further details on the construction of our sample characteristics from this data set. We exploit exclusively the available sample information on labor force status, labor earnings and hours, demographics, and educational attainment. Since the model requires us to distinguish individuals based on categories of “high” and “low” schooling, we must take a stand on this definition. In what follows, we assign all those with educational attainment equal to or exceeding a four-year bachelor’s degree to the high education 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 four years of college.

Although our model allows for a variety of forms of individual heterogeneity, we thought that we should apply it to a relatively homogeneous subpopulation. For this reason, we look exclusively at males from 25 to 34 years of age, inclusive. The vast majority of these individuals would have completed their formal schooling. In addition, since stationary models are weak when it comes to capturing life-cycle events, we used a fairly narrow age range. We also exclude all members of the sample who appear “out of the labor force” in the first month of the sample, the formal definition of which is given below. Finally, to be included in the sample, an individual must appear in and have valid labor force information for each of the waves of the survey that we utilize. Our final sample consists of 4,000 individuals. Of these, 1,072 report having a four-year bachelor’s degree or higher. This produces a schooling completion rate of 26.8%.

A critical source of identifying information for our model comes from the wage and employment histories of individuals in the sample. In the SIPP, weekly labor force status information is collected retrospectively for the four months since the last wave. Categories of weekly labor force status are as follows: (i) with a job/business and working; (ii) with a job/business and not on layoff, but absent without pay; (iii) with a job/business, but on layoff and absent without pay; (iv) no job/business and looking for work or on layoff; and (v) no job/business, not looking and not on layoff. In our application, we define an individual to be employed in any given week if he reports any of the first three categories. Similarly, we record an unemployment spell if an individual responds within category (iv) or (v) for any given week. Finally, we designate an individual as being “out of the labor force” if he reports category (v) for every week in the first month of the sample. That is, they are classified as out of the labor force if individuals report having no job and having spent no time looking for employment.

In addition to information on employment status, the survey tracks the employer history of each sample member. In each wave, an index number denotes the firm to which the retrospective earnings and hours information applies. This information is augmented by a variable that indicates whether the individual is still working for the employer for whom the information is given. This information can be linked across waves to determine whether an individual is at the same employer at two consecutive points in time. We let m_t denote the worker’s employer at time t ; if $m_0 = m_1$, the worker is employed at the same firm at both points in time. Since workers may hold more than one job or fail to report a job that they currently still hold, we only assign a new index number if it can be confirmed that the individual is no longer working for the employer observed in the first sample month when we observe him later in the survey. If an individual is employed at both points in time, has $m_0 \neq m_1$, and if he has not experienced an unemployment spell during the intervening period, we can infer that a job-to-job transition has occurred.

Table 1 provides some descriptive statistics for these variables. We see that in the high-schooling group, the unemployment rate at the beginning of the sample is 2.1%. Of those employed at time 0, 8.1% experience an unemployment spell over the 12 month period, whereas 18.0% have made a job-to-job transition with no intervening unemployment spell.¹⁰ The unemployment rate in the low-schooling group is 5.8%. Of those employed at time 0, 14.1% experience an unemployment spell over the 12-month period, whereas 17.9% make a job-to-job transition.

Wages and hours are recorded on a monthly basis in the survey. Although many sample members report an hourly wage, this rate is not always directly available. For those who do not report an hourly rate, we impute the wage by dividing reported labor earnings by reported hours. To limit measurement error, we make this imputation only for those who report greater

¹⁰ This number is a lower bound for the total number of job-to-job transitions, since individuals may have undergone such a transition and then experienced an unemployment spell, which removes them from this particular conditioning set. Our use of a simulated moments estimator allows us to incorporate any type of sample path in our estimation procedure.

TABLE 1
DESCRIPTIVE STATISTICS

	Combined	Low schooling	High schooling
Observations	4000	2928	1072
Employment transitions			
Unemployed	0.048	0.058	0.021
Employed	0.952	0.942	0.979
Employed at beginning and at least one unemployment spell	0.125	0.141	0.081
New employer, no unemployment spell	0.180	0.179	0.180
Wages			
$E[w]$	16.288	14.529	21.417
$V[w]$	59.382	42.322	73.866
Number of valid observations	3140	2338	802
Number of imputed observations	1268	666	602

than 30 hours of work per week. In order to reduce the impact of measurement error further, we trim wages by discarding the highest and lowest 2.5% of wage observations. The wages utilized in estimation are those that survive these selection criteria.

Table 1 presents some descriptive statistics for the wage distribution at the beginning of the sample. The mean hourly wage is 16.29, with a significant difference between those in the low-schooling category (14.53) and the high-schooling category (21.42). There is an appreciable degree of dispersion in these distributions, with more dispersion in the college-educated group, as is to be expected.

Among the valid wage observations, 666 individuals in the low-schooling group have an imputed wage, whereas 602 observations in the high-schooling group are imputed. That a lower portion of the high-schooling workers receive an hourly salary is unsurprising, given the typical occupations and work arrangements of these two groups.

Due to the challenges of identifying worker bargaining power that are discussed below, we utilize information on the labor share of the surplus. We will exploit 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 surplus division parameter, α . The discussion in Krueger (1999) led us to believe that 0.67 was a reasonable value to use for the labor share for this group of labor market participants.¹¹

The federal minimum wage was \$5.15 during the sample period, although a number of state minimum wages were higher, often substantially so. Since we are using data for individuals in the age group 25–34, and given the wage distributions discussed above, we assume that there is no binding minimum wage when estimating the model, which considerably simplifies the estimation problem. Given the model estimates that we obtain under this assumption, we perform policy experiments in which binding minimum wages are introduced into the labor market. In performing the policy experiments at the point estimates of the primitive parameters, we find that a minimum wage of 5.15 is not binding, so that the estimates are consistent with our assumption that minimum wages are not binding at such low levels for the set of individuals in our sample.

7.2. Identification. The specification of the model we estimate is the most general one that we have considered, in which we allow each schooling class to inhabit its own market, which is characterized by its own contact rates when individuals are in the unemployed and employment

¹¹ In Flinn's (2006) study of minimum wage effects on labor market outcomes, the labor share was computed from the Consolidated Income Statement of McDonald's corporation for 1996 and was found to be about 53%. This was deemed reasonable since the data used in estimation were for workers between the ages of 16 to 24, inclusive. For the older workers in our sample, who are all full-time labor market participants and most of whom have completed schooling, the value of 0.67 seems to be appropriate.

states and their own exogenous dissolution rates when employed. Additionally, the markets share certain parameters, including the matching distribution, G , the discount rate, ρ , and the coefficient of ah_s associated with flow utility in the unemployment state, b . We have also assumed, simply due to identification issues, that both markets share the same value of the surplus division parameter, α . Of course, the ability endowment distribution, F , the human capital augmentation parameter for those with a college education, h , and the lump sum cost of schooling, c , are common to individuals making the schooling decision before entering the labor market. Thus, the primitive parameters of the model are $\rho, b, \lambda_{U,s}, \lambda_{E,s}, \eta_s, F, G, h, c$, and α , for the two schooling classes, $s = 0, 1$.

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 (explicitly) determine labor market outcomes once we condition on the observed schooling level, s . The FH analysis was likelihood based, and for reasons to be discussed below, we employ method of (simulated) moments (MSM) estimators. However, the arguments that they make carry over to the case of our estimator given that we select appropriate sample characteristics 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; in this case the maximum likelihood estimator is undefined, and to handle such cases typically measurement error must be explicitly introduced into the model.¹² Although there is some loss in asymptotic efficiency, as long as the MSM estimator includes functions of the data that are similar to those comprising the minimal sufficient statistics that characterize the likelihood function, identification and “good” asymptotic properties carry over to the MSM estimator that we utilize. The reasonableness of the parameter estimates and the small bootstrapped 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_j^* as constants. There is no loss in efficiency in doing so, since the simple search model considered in FH and the one estimated here are both fundamentally underidentified. The model that they investigate is a (very) special case of the one analyzed here. In particular, it is the special case of our model for which $\alpha = 1, a = 1, h = 1$, and $\lambda_E = 0$ for all agents. With no schooling decision, the sole decision rule of the model is θ^* . They show that a certain (broad) class of 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 by solving the functional equation that determines θ^* evaluated at consistent estimates of all of the other parameters. In this manner, consistent estimates of λ_U, η, G , and b are obtained given an assumed value of ρ .

The extension to the model we estimate allows α to be a free parameter along with a nondegenerate distribution of a , given by F , introducing a human capital level h (which is presumably greater than 1) and a cost of college completion, c . As we showed above, the critical schooling value a^* is a function of all of the primitive parameters of the model through $\theta_1^*, \theta_2^*, h$, 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 and h , is used to form a consistent estimator 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 . Under our model assumptions (and assuming no binding minimum wage), the critical match value θ_j^* is independent of ability, conditional on the schooling level s , but does vary by schooling level in the heterogeneous markets case. We begin by considering the simple case in which $\lambda_{E,j} = 0, j = 0, 1$. If the model parameters are identified in this case, it is fairly straightforward

¹² See, for example, Flinn (2002) and Dey and Flinn (2005) for discussions regarding the introduction of measurement error into likelihood-based estimators to eliminate this problem.

to establish that the case of OTJ search only strengthens identification arguments, even if the OTJ search case introduces two new free parameters, $\lambda_{E,j}$, $j = 0, 1$. Recall from (7) that the wages for individuals in the two schooling classes will be given by

$$w_0 = a(\alpha\theta + (1 - \alpha)\theta_0^*), \quad a < a^*, \quad \theta \geq \theta_0^*$$

$$w_1 = ah(\alpha\theta + (1 - \alpha)\theta_1^*), \quad a \geq a^*, \quad \theta \geq \theta_1^*$$

so that the distribution of log wages in each schooling group is a convolution of functions of the distributions of a and θ ,

$$\ln w_0 = \ln a + \ln(\alpha\theta + (1 - \alpha)\theta_0^*), \quad a < a^*, \quad \theta \geq \theta_0^*$$

$$\ln w_1 = \ln a + \ln h + \ln(\alpha\theta + (1 - \alpha)\theta_1^*), \quad a \geq a^*, \quad \theta \geq \theta_1^*$$

The distribution of the log wage in each group is a convolution of a truncated normal random variable, $\ln a$, and a nonlinear function of a truncated lognormal random variable, θ . Then we can write

$$\ln w_0 \sim R_0(\ln w_0; \mu_a, \sigma_a, \mu_\theta, \sigma_\theta, \alpha, a^*, \theta_0^*)$$

$$\ln w_1 \sim R_1(\ln w_1; \mu_a, \sigma_a, \mu_\theta, \sigma_\theta, h, \alpha, a^*, \theta_1^*).$$

It is straightforward to verify that, with access to only a cross section of the SIPP data, wage and school completion information is sufficient to identify all of the parameters characterizing the distributions R_0 and R_1 under these distributional assumptions using a maximum likelihood estimator. However, there exist additional support conditions that must be satisfied for the m.l. estimator to be defined. In particular, under our lognormality assumptions on a and θ , the only one that is potentially binding is that $w_1 \geq a^*\theta_1^*$. Following FH, we can use this constraint as a way to eliminate the estimation of the parameter θ_1^* by substituting into the likelihood $\hat{\theta}_1^* = \min\{w_i\}_{i:s_i=1}/a^*$, since $\min\{w_i\}_{i:s_i=1}$ is a superconsistent estimator of $a^*\theta_1^*$.

Although the log-likelihood function of wages can be used to show that all of the parameters in the distribution functions R_0 and R_1 are identified in theory, including the bargaining surplus parameter α , Monte Carlo sampling experiments conducted and discussed in Flinn (2006) indicate that an extremely large number of observations are required to obtain well-behaved sampling distributions of the m.l. estimator, and his case is a special case of the one examined here. It is for this reason that we follow Flinn (2006) in utilizing information from outside the sample on firms' share of the surplus to aid in the identification of α , in particular.

We now consider the addition of OTJ search. This makes the cross-sectional distribution of wages in the two schooling groups considerably more complex to describe, in addition to adding the two parameters to be estimated. However, with access to limited amounts of event-history data, the wage changes associated with job-to-job transitions aid considerably in identification. Let w' be the new wage and w the former wage in a job-to-job transition. In our OTJ search model without wage renegotiation, it must be the case that $w' > w \geq \theta_j^*v$. Then the distribution of

$$\ln w' - \ln w = \ln \left(\alpha\theta' + (1 - \alpha) \left(\rho\bar{V}_{U,s} - \int_{\theta'} [\bar{V}_{E,s}(\tilde{\theta}) - \bar{V}_{U,s}] dG(\tilde{\theta}) \right) \right)$$

$$- \ln \left(\alpha\theta + (1 - \alpha) \left(\rho\bar{V}_{U,s} - \int_{\theta} [\bar{V}_{E,s}(\tilde{\theta}) - \bar{V}_{U,s}] dG(\tilde{\theta}) \right) \right),$$

in schooling market s is not a function of μ_a and σ_a . Using these types of wage change data in conjunction with cross-sectional wage information greatly aids in solving the difficult deconvolution problem we face.

Although the decision rule a^* is theoretically identified from the distributions of the wage data for the two schooling classes, there is additional sample information that greatly increases the

precision of the estimate. Since $P(s = 1) = \tilde{F}(a^*)$, $a^* = \tilde{F}^{-1}(P(s = 1))$. Given a consistent estimator of $P(s = 1)$, which is simply $\hat{P}(s = 1) = \sum_{i=1}^N s_i/N$, we have $\hat{a}^* = \tilde{F}^{-1}(\hat{P}(s = 1); \mu_a, \sigma_a)$, and this quantity is substituted into the log-likelihood function. In the case of no OTJ search, to estimate the cost of schooling we note that

$$c = \frac{a^*}{\rho} (\theta_1^* h - \theta_0^*),$$

so that with an assumed value of ρ and consistent estimators for all of the other parameters and decision rules on the right-hand side, a consistent estimator of c is obtained. In the case of OTJ search, this expression must be altered slightly and becomes

$$c = a^* (h\bar{V}_{U,1} - \bar{V}_{U,0}).$$

It remains to consider the rate parameters, the estimation of which is straightforward. Under our model structure, if uncensored durations of unemployment and completed employment spells¹³ are available, then all of these spells follow (conditional) exponential distributions, the constant hazard rates of which are given by $h_{U,j} = \lambda_{U,j} \tilde{G}(\theta_j^*)$ and $h_{E,j} = \eta_j$. Thus, consistent estimators of these hazard rates, along with consistent estimators of G and θ_j^* , are used to recover $\lambda_{U,j}$ and η_j , $j = 0, 1$. Given an observed wage w , we know that $w = ah_s(\alpha\theta + (1 - \alpha)(\rho\bar{V}_{U,s} - \int_{\theta} [\bar{V}_{E,s}(\tilde{\theta}) - \bar{V}_{U,s}] dG(\tilde{\theta})))$, and it is straightforward to show that $w(\theta)$ is strictly increasing so that there exists an inverse function $\theta = \tau(w)$ that is monotone increasing as well. The rate of job-to-job transitions conditional on the current wage w is given by $\lambda_{E,j} \tilde{G}(\tau(w))$. With access to the durations of job spells that ended in a job-to-job transition, the maximum likelihood estimator of $\lambda_{E,j}$ is given by

$$\hat{\lambda}_{E,j} = \frac{N_{J,j}}{\sum_{i: s_i=j} \tilde{G}(\hat{\tau}(w); \hat{\mu}_\theta, \hat{\sigma}_\theta) t_i},$$

where $N_{J,j}$ is the number of individuals with job-to-job transitions in schooling group j , t_i is the length of the job spell that ends in a job-to-job transition, and the “ $\hat{\tau}$ ” denotes a consistent estimator of the parameter or decision rule. Allowing for right-censoring in the duration data is accomplished in the obvious manner.

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 model, 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.

7.3. Estimator. Although our identification discussion was likelihood based, for a variety of reasons we utilize a method of simulated moments estimator to estimate the model. Under the data-generating process (DGP) of the model, there are a number of sharp restrictions on the support of the wage distributions by schooling class that are generally not 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

¹³ A complete employment spell is a sequence of job spells between two unemployment spells, and the duration of a spell is the sum of the durations of the jobs that comprise it.

trying to estimate a convolution, so that the addition of another random variable to the wage process can only exacerbate the difficulty of separately identifying the F and G distributions.

We chose to use a moment-based estimator that employs a large amount of information characterizing the wage distributions by schooling class, but that does not impose such a large penalty on the estimator for violating some of the implications of the 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.

The information from the sample that is used to define 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 and decision rules 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 elements equal to the inverse of the variance of the corresponding element of M_N . The exception to this rule is the weight attached to the labor share, which is treated as known (i.e., it is given an extremely large weight).

Under our random sampling assumption, we have that $\text{plim}_{N \rightarrow \infty} M_N = M$, 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 $\text{plim}_{N \rightarrow \infty} \hat{\omega}_{N,Q} = \omega$ for any positive-definite matrix Q . We compute bootstrap standard errors using 100 replications.

The sample characteristics we use to estimate both versions of the heterogeneous markets model, with and without renegotiation, are the same and are listed in Table 3, below. All together there are 16 sample characteristics for each schooling group and the proportion of the sample in the high-schooling group, for a total of 33 sample characteristics. The labor share value is also included as a constraint on the estimator.

The first characteristics listed refer to labor market states at the two points in time. Some of the sample characteristics are conditional and some are not. For example, the first characteristic is the proportion of the sample employed at time 0, which is taken to correspond to the steady-state employment rate. The second characteristic is the proportion of the sample who are employed at time 1 given that they were employed at time 0. The fourth characteristic is the unconditional probability that an individual is employed at the same employer at time 0 and time 1 (that is, this must be 0 for people employed at different employers at time 0 and 1 as well as for all other individuals who were not employed at one or both of these dates). The characteristic N_U is the number of unemployment spells in the one year period of time, and F_U is the proportion of the one-year period spent in unemployment. The last seven characteristics relate to the first and second moments of wages both unconditionally (at time 0) and conditional on events that occurred over the one-year observation period.

7.4. Estimation of Parameters in the GE Version of the Model. 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 schooling 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 (8) and (9) more succinctly as

$$0 = -\psi_s + k_s^{\delta_s} A_s,$$

where A_s is the expected value of a filled vacancy in schooling market s , and $k_s = (U_s + \phi_s E_s)/v_s$, with ϕ_s the effective search units associated with employed workers and v_s denoting the measure

of vacancies in sector s . From our first-stage estimates of the partial equilibrium model, we can consistently estimate A_s , along with U_s , $E_s (= 1 - U_s)$, and ϕ_s . If vacancies are observed, then we can consistently estimate k_s , so that the FEC produces two equations in four unknowns, ψ_0 , ψ_1 , δ_0 , δ_1 , and thus we cannot hope to obtain consistent estimates of these parameters without further restrictions.

We consider the (more realistic) case in which v_i is unobserved in each sector, so that it must be estimated along with the other primitive parameters ψ_s and δ_s , $s = 1, 2$. Since the model is not identified even when v_s is known, it clearly is not when v_s must be determined as well. As noted in Flinn (2006), in such a case it is necessary to restrict the matching function either to a class of functions with no unknown parameters or by assuming values of the parameters required to characterize a parametric function. Under our Cobb–Douglas assumption with the constant returns to scale restriction, by assuming that TFP is equal to 1, only one parameter is required to characterize the matching function in each market. We use typical values of this parameter from the macroeconomics literature as surveyed by Petrongolo and Pissarides (2001) and perform our quantitative exercises using values of δ_s taken from the set $\{0.4, 0.5, 0.6\}$.

Given values of δ_s , we can obtain a consistent estimate of the vacancy rate from the matching function and the consistent estimates of the contact rates $\lambda_{U,s}$ and $\lambda_{E,s}$. Given the vacancy rate we can consistently estimate k_s , and hence with an assumed value of δ_s and consistent estimates of A_s and k_s , we define a consistent estimator of ψ_s . Consistent estimates of ψ_s are required to perform the policy experiments reported below within a general equilibrium framework.

8. MODEL ESTIMATES

We estimated the full (i.e., heterogeneous labor markets) model for both the renegotiation and the no-renegotiation cases. In terms of functional form assumptions, recall that we assume that individual ability, a , has a lognormal distribution, with $\ln a$ having mean μ_a and standard deviation σ_a . Match productivity, which is independently distributed with respect to a , also follows a lognormal distribution, with $\ln \theta$ having mean μ_θ and standard deviation σ_θ . Since the flow value of unemployment is νb and flow output is $\nu \theta$, it is difficult to separately identify the distributions of a and θ . For this reason, we have fixed the mean of $\ln \theta$, μ_θ , to zero. We are not able to argue that this is an inconsequential normalization.

We begin by discussing the MSM estimates of the model that allows contract renegotiation between workers and firms. The estimates are presented in the first column of Table 2. Before discussing the results in detail, it is necessary to provide some intuition. As noted in Section 4.1, under contract renegotiation the amount of the surplus obtained by the worker is impacted both by the surplus division parameter α and by the rate of contacts with potential employers while employed, $\lambda_{E,s}$. It is tautologically true that when $\alpha = 1$ individuals capture all of the surplus from the match, but it is also noteworthy that as $\lambda_{E,s} \rightarrow \infty$, the worker also gains all of the surplus from the match, attains a wage equal to $\nu \times \bar{\theta}$, where $\bar{\theta}$ is the upper bound on the support of θ , and the market exhibits no unemployment (i.e., the market is in competitive equilibrium).¹⁴ Thus, $\lambda_{E,s}$ is an important determinant both of the steady-state distribution of match values and the share appropriated by workers.

The first thing to note is the extremely low point estimates of $\lambda_{E,0}$, $\lambda_{E,1}$, and α . Whereas the contact rate among the unemployed is estimated to be 0.522 for the low-schooling group and 0.203 for the high-schooling group, the estimates of the employed contact rates are 0.016 and 0.012, respectively, over an order of magnitude less for both groups. These estimates imply that contacts with other firms occur, on average, only every five or six years. The estimated exogenous job dissolution rates are approximately 50% larger than the contact rates in the employment state for both schooling groups. The estimated value of α is only 0.091. One of the focuses of interest is the amount of variation in the ability and match productivity distributions. We see that they are not very different, with the estimated standard deviations of ability and

¹⁴ Recall that $\lambda_{E,s} = \phi_s \lambda_{U,s}$, so $\lambda_{E,s} \rightarrow \infty \Rightarrow \lambda_{U,s} \rightarrow \infty$.

TABLE 2
PARAMETER ESTIMATES: STANDARD ERRORS CALCULATED USING 100 BOOTSTRAP SAMPLES

Parameter	No renegotiation	Renegotiation
$\lambda_{U,0}$	0.342 (0.003)	0.520 (0.004)
$\lambda_{E,0}$	0.048 (0.001)	0.016 (0.000)
η_0	0.017 (0.001)	0.028 (0.001)
θ_0^*	0.680 (0.006)	0.140 (0.005)
$\lambda_{U,1}$	0.365 (0.005)	0.207 (0.009)
$\lambda_{E,1}$	0.066 (0.006)	0.011 (0.000)
η_1	0.008 (0.001)	0.010 (0.000)
θ_1^*	0.680 (0.006)	0.070 (0.002)
α	0.252 (0.002)	0.090 (0.002)
h	1.128 (0.002)	1.063 (0.001)
μ_a	3.102 (0.024)	3.015 (0.031)
σ_a	0.142 (0.005)	0.131 (0.006)
σ_θ	0.129 (0.003)	0.151 (0.004)
b	-0.371 (0.004)	-137.360 (5.423)
c	1254.015 (29.750)	1378.577 (98.652)
\bar{c}	28.825 (0.684)	31.688 (2.268)

match productivity being 0.128 and 0.152, respectively. Finishing college is estimated to increase an individual's ability by 6.27% ($\hat{h} = 1.0627$). The total cost of schooling, c , is estimated to be 1155.086. We can provide a more useful interpretation of this number if we consider it to be

$$c = \int_0^{48} \bar{c} \exp(-\rho t) dt,$$

where \bar{c} is the flow cost of college, which is paid for over 48 months. Using $\rho = 0.05/12$, the flow cost of college is estimated to be 54.140, which is approximately 2.5 times the average wage paid to college graduates.

Under the no-renegotiation specification, the estimates of the contact rate parameters $\lambda_{E,0}$ and $\lambda_{E,1}$ remain low in comparison with the contact rates when unemployed, but they are significantly higher than in the renegotiation case. The estimated rates for the low- and high-schooling groups are 0.048 and 0.066, which imply average arrival rates of approximately 21 and 15 months, respectively. For both schooling groups, these rates are substantially larger than the exogenous dismissal rates (0.017 and 0.008 for the low- and high-schooling groups). The surplus division parameter estimate is 0.252 in this specification, which is almost three times larger than the estimate under the renegotiation assumption. This is to be expected, of course, since the path to obtaining larger shares of the surplus through firm competition has been shut down. Under this specification, the standard deviations of native ability and match

productivity are of similar size to the estimates obtained under the renegotiation assumption, although the estimated standard deviation of match heterogeneity, 0.129, is now less than the estimated standard deviation of native ability, 0.142. In this specification, the estimate of the percentage increase in ability due to college completion rises to 12.8, approximately twice the value obtained under the renegotiation assumption. The cost of schooling is estimated to be 1254.015, which yields an estimated flow cost of schooling (\bar{c}) of 58.777.

For completeness, we also report here the estimated values of ψ_0 and ψ_1 , which are used when conducting the policy experiments below in general equilibrium. Given a choice of the elasticity δ_s and TFP^{15} of the matching function, ψ_s can be estimated using the free entry condition for $s = 0, 1$, as described in Section 7.4. For our preferred model with no renegotiation, we estimate flow costs of 317.42 in the low-schooling market and 311.65 in the high-schooling market. Since this cost can be arbitrarily scaled by different choices of the TFP parameter, it is not possible to give any meaningful interpretation of these numbers, but it is interesting that the costs are similar under these symmetric normalizations.

In the first column of Table 3, we present the values of the sample characteristics used in defining the MSM estimator, and the second column contains the fitted values from the renegotiation specification. One can see that the renegotiation model fits several sample characteristics quite poorly, the most notable ones being the probability that an individual who was unemployed at baseline was employed one year later (0.682 [data] versus 0.957 [model] for the low-schooling group and 0.739 [data] versus 0.871 [model] for the high-schooling group) and the probability that an individual who was at a different employer at times 0 and 1 had no intervening unemployment spell (0.704 [data] versus 0.199 [model] for the low-schooling group and 0.791 [data] versus 0.408 [model] for the high-schooling group). The no-renegotiation model (column 3), while also fitting some moments relatively poorly, performs much better overall. In terms of sample characteristics discussed above, the estimated probability that an individual who was unemployed at time 0 is employed one year later is 0.919 for the low-schooling group and 0.892 for the high-schooling group, whereas the estimated probability that an individual working at a different firm at time 1 had any unemployment in the intervening period is 0.459 for the low-schooling group and 0.625 for the high-schooling group. Both specifications of the model clearly have some difficulty generating state occupancy probabilities and state transition probabilities simultaneously, which is not unexpected. The stationarity assumption is clearly at odds with reality, and the simplicity of the model structure further curtails our ability to fit both types of probabilities. For example, if job acceptance decisions were functions of an individual's ability, the length of time spent in unemployment search would not be (unconditionally) distributed as an exponential random variable. However, given the model structure that leads to job acceptance decisions that are independent of ability, all individuals of the same schooling class will have a constant exit rate from unemployed search.

On a more positive note, both specifications of the bargaining model perform relatively well in terms of fitting moments of the wage distribution and wage changes between times 0 and 1, conditional on mobility status. In general, the no-renegotiation model seems to fit second-order moments better, and the renegotiation specification does a better job fitting first-order moments. In the end, the no-renegotiation specification fits this set of characteristics better than the renegotiation specification, with the distance measures being 2.050 and 4.840, respectively. This result is consistent with what is found in Flinn and Mablí (2009), where bootstrap-based hypothesis tests were used to select the no-renegotiation specification as the one most consistent with the data. Given that we also find the point estimates associated with the no-renegotiation specification to be more reasonable, the remainder of our quantitative analysis uses these estimates.

¹⁵ In our application, we normalize this value to 1.

TABLE 3
MOMENTS AND MODEL FIT

Moment	Data	No renegotiation	Renegotiation
Low-schooling moments			
$\mathbb{P}[E_0 = 0]$	0.058	0.046	0.049
$\mathbb{P}[E_1 = 1 \mid E_0 = 1]$	0.943	0.959	0.942
$\mathbb{P}[E_1 = 1 \mid E_0 = 0]$	0.682	0.919	0.937
$\mathbb{P}[m_0 = m_1]$	0.691	0.710	0.654
$\mathbb{P}[m_0 \neq m_1]$	0.252	0.259	0.305
$\mathbb{P}[m_0 \neq m_1 \mid wks_u = 0]$	0.704	0.458	0.201
$\mathbb{E}[N_u \mid E_0 = 1, E_1 = 1]$	0.104	0.147	0.271
$\mathbb{E}[F_u \mid E_0 = 1]$	0.036	0.035	0.044
$\mathbb{E}[F_u \mid E_0 = 0]$	0.499	0.287	0.208
$\mathbb{E}[w_0 \mid m_0 = m_1]$	15.029	15.388	12.949
$\mathbb{E}[w_0^2 \mid m_0 = m_1]$	267.966	240.644	262.133
$\mathbb{E}[w_0 \mid m_0 \neq m_1, wks_u = 0]$	13.057	14.121	12.257
$\mathbb{E}[w_1 \mid m_0 \neq m_1, wks_u = 0]$	14.474	15.456	18.175
$\mathbb{E}[w_0 \times w_1 \mid m_0 \neq m_1, wks_u = 0]$	214.094	221.070	227.518
$\mathbb{E}[w_0^2 \mid m_0 \neq m_1, wks_u = 0]$	206.890	202.510	239.987
$\mathbb{E}[w_1^2 \mid m_0 \neq m_1, wks_u = 0]$	261.817	242.494	339.861
High-schooling moments			
$\mathbb{P}[E_0 = 0]$	0.021	0.023	0.050
$\mathbb{P}[E_1 = 1 \mid E_0 = 1]$	0.974	0.977	0.953
$\mathbb{P}[E_1 = 1 \mid E_0 = 0]$	0.739	0.892	0.871
$\mathbb{P}[m_0 = m_1]$	0.748	0.795	0.834
$\mathbb{P}[m_0 \neq m_1]$	0.227	0.186	0.125
$\mathbb{P}[m_0 \neq m_1 \mid wks_u = 0]$	0.791	0.625	0.409
$\mathbb{E}[N_u \mid E_0 = 1, E_1 = 1]$	0.061	0.071	0.075
$\mathbb{E}[F_u \mid E_0 = 1]$	0.017	0.017	0.031
$\mathbb{E}[F_u \mid E_0 = 0]$	0.461	0.288	0.388
$\mathbb{E}[w_0 \mid m_0 = m_1]$	21.951	24.785	21.292
$\mathbb{E}[w_0^2 \mid m_0 = m_1]$	554.670	621.291	616.738
$\mathbb{E}[w_0 \mid m_0 \neq m_1, wks_u = 0]$	19.506	22.363	19.620
$\mathbb{E}[w_1 \mid m_0 \neq m_1, wks_u = 0]$	23.422	24.340	24.731
$\mathbb{E}[w_0 \times w_1 \mid m_0 \neq m_1, wks_u = 0]$	520.038	549.974	485.229
$\mathbb{E}[w_0^2 \mid m_0 \neq m_1, wks_u = 0]$	463.099	507.211	576.119
$\mathbb{E}[w_1^2 \mid m_0 \neq m_1, wks_u = 0]$	645.396	599.172	630.291
Combined moments			
Labor share	0.670	0.669	0.674
$\mathbb{P}[s = 1]$	0.268	0.268	0.268
Model fit			
$(M_N - \tilde{M}(\omega))' W_n (M_N - \tilde{M}(\omega))$		2.050	4.840

9. POLICY EXPERIMENTS

In this section, we use our estimates to examine the welfare effect of two potentially relevant policy interventions: the imposition of a minimum wage and a subsidy to the cost of schooling. There are (at least) two criteria by which one might reasonably evaluate these policies. Without redistribution, a natural welfare candidate is just an appropriately weighted sum of (ex ante) worker welfare. We define

$$(10) \quad W_0 = \int^{a^*} V_{U,0}(a) dF(a) + \int_{a^*} (V_{U,1}(ah) - c) dF(a).$$

This is the population expectation of ex ante values to workers whose education decisions are characterized by a cutoff type, a^* . Since the minimum wage is typically thought of as a parsimonious policy instrument for improving outcomes for workers, this is the primary criterion to evaluate the merits of such a policy. However, as in Flinn (2006) and Flinn and

Mabli (2009), minimum wages can also have beneficial efficiency effects as well, and this is even more likely to be the case when schooling decisions are endogenous.

We will also define a measure of aggregate resources in the steady state, and this will be of particular interest when considering a government policy to affect the cost of schooling. We define the flow of total resources in market s in the steady state as

$$(11) \quad Y_s = \underbrace{(1 - U_s) \int \int ah_s \theta dH_{E,s}(\theta|a, m) dF_{E,s}(a|m) + U_s \int bah_s dH_{U,s}(a|m)}_{\text{Output}} - \underbrace{\frac{\phi_s(1 - U_s) + U_s}{k_s}}_{\text{Vacancy Cost}=VC_s} \psi_s.$$

Now, let $P_1 = \tilde{F}(a^*)$ be the schooling rate. In the steady state, the planner's welfare criterion is

$$(12) \quad W_1 = (1 - P_1) \cdot Y_0/\rho + P_1 \cdot (Y_1/\rho - c).$$

To understand the efficiency properties of this model, it is important to recognize the different effects at play. First, our bargaining assumptions imply that the returns to education in aggregate resources are not aligned with workers' private returns; this is the standard "holdup" problem. Second, firms' incentives to post vacancies are not aligned with the planner's, which is once again due to the bargaining assumptions of the model and assumptions on the match technology. This issue, commonly referred to as a "congestion externality," is also at play here. Policies that push firms closer to the planner's choice of vacancy posting will produce efficiency gains, ceteris paribus. Finally, since the schooling decisions of workers affect the firm's conditional expectation of worker ability in each market, there is an additional externality here that creates market inefficiencies. This was the effect highlighted by Charlot and Decreuse (2005, 2010).

To enrich our understanding of how these effects exhibit themselves in our model, we perform a series of experiments in Appendix A to determine their relative degrees of importance.

9.1. The Minimum Wage. In this section, we examine the labor market consequences of imposing a minimum wage, m . In order to better understand the effects at play, we first perform this experiment in partial equilibrium, with fixed contact rates, before allowing the contact rates to be determined endogenously through the free entry condition. As we discuss in Section 7.4, this requires us to take a stand on the elasticity, δ_s , of the matching function. Here we assume that $\delta_0 = \delta_1 = 0.5$; however, in Appendix A.4 we briefly consider the sensitivity of these results to different choices of δ_0 and δ_1 . In addition, we examine welfare changes both when workers' education decisions are fixed at the level found in our estimates and when workers are allowed to make their investment decisions to maximize their welfare given the value of the minimum wage, m .

We begin by keeping contact rates fixed (the partial equilibrium case). We consider the welfare effect of setting the minimum hourly wage in the range between \$5 and \$30. In this case, the only welfare consideration is the trade-off between making low matches more profitable to the worker and increasing the probability that a given match is unacceptable to the firm. By increasing all wages bargained at less than m , surplus is transferred from the firm to the worker. However, for any worker of type v , the minimum wage will render all previously acceptable matches untenable when

$$(13) \quad \theta < \frac{m}{v}$$

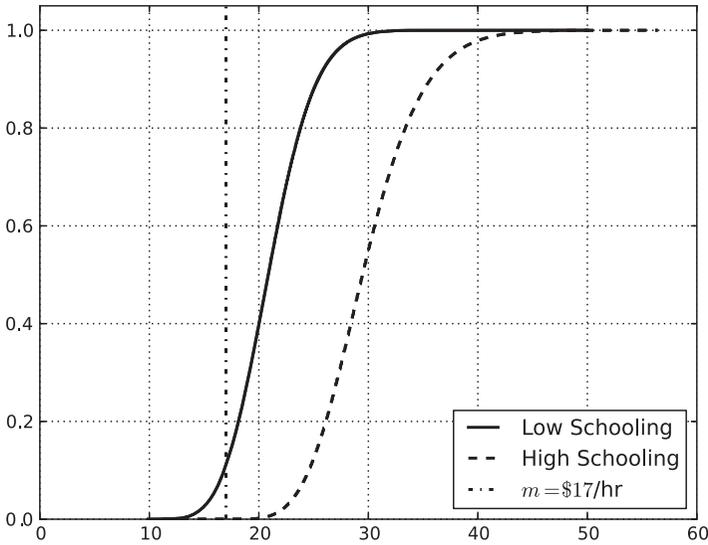


FIGURE 1

THE UNCONDITIONAL DISTRIBUTION OF POTENTIAL OUTPUTS IN EACH MARKET

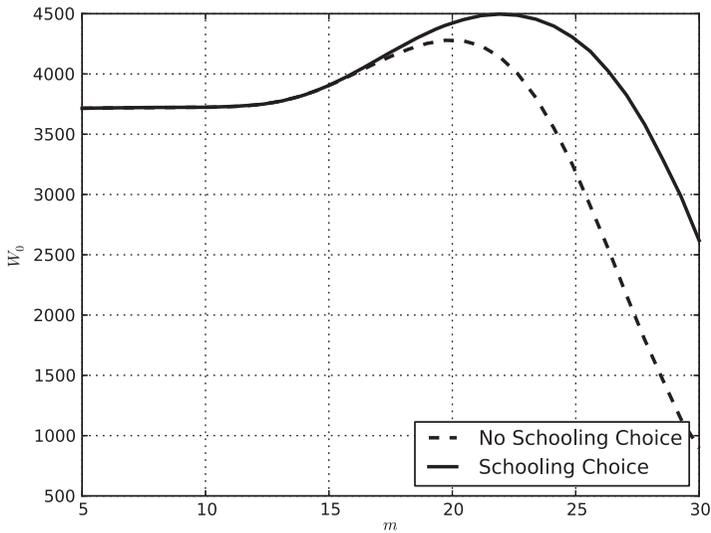


FIGURE 2

TOTAL WORKER WELFARE IN PARTIAL EQUILIBRIUM (FIXED k) WITH AND WITHOUT ENDOGENOUS SCHOOLING

To examine the potential effect of (13), we consider the density across potential outputs in worker–match pairs. This is achieved by integrating across combinations of $ah_s\theta$ using the unconditional density of worker types in each market ($f(a | a < a^*)$ and $f(a | a \geq a^*)$) and the density g of potential match draws. We take a^* from our estimates. Figure 1 shows the cumulative distribution across potential outputs in each market. To facilitate analysis, we consider a minimum wage of \$17 an hour. At this choice, we see that approximately 11% of potential matches are no longer viable in the low-schooling market, whereas virtually no matches in the high-schooling market bind against this constraint.

Figure 2 contains the results of our first exercise. The hump shape in worker welfare shows these countervailing effects at work for both cases (fixed versus endogenous schooling).

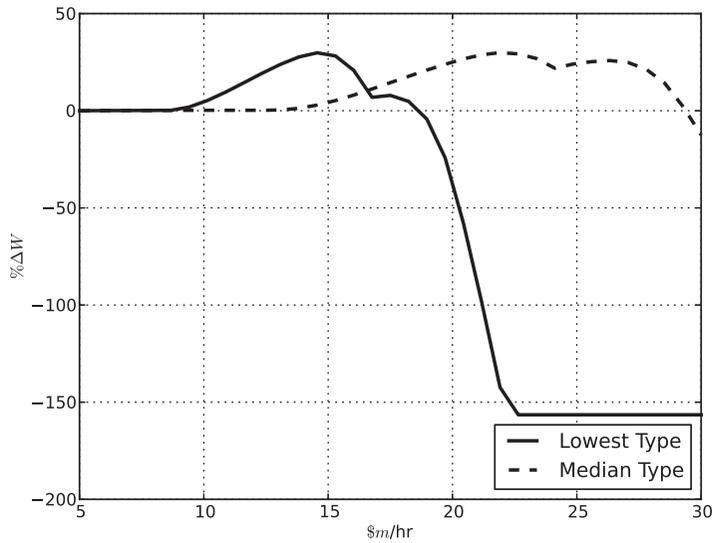


FIGURE 3

PERCENTAGE CHANGE IN WORKER WELFARE IN PARTIAL EQUILIBRIUM FOR THE LOWEST AND MEDIAN WORKER TYPES

Although raising the minimum wage increases worker welfare by increasing the return to low matches, increasingly more combinations of a and θ are made unacceptable to the firm, the effect of which begins to dominate. Allowing workers to adjust their schooling decision yields extra gains in welfare, as some workers are able to increase their productivity and hence increase the range of viable matches.

Another relevant consideration when selecting the minimum wage concerns the distributional effects of such a policy. The results in Figure 2 suggest an “optimal” minimum wage between \$20 and \$22 an hour. However, it is not clear at this level what the consequences are for the welfare of people at the bottom end of the skill distribution. Knowing the consequences of the minimum wage across the spectrum of heterogeneity is an important policy input and was not considered in Flinn (2006) since individuals were assumed to be identical ex ante.

To investigate this, we evaluate welfare changes for workers of different ability types, a . Let $V_{U,s}(ah_s; m)$ indicate the value in unemployment to a worker of type a when the minimum wage is set at m . We calculate

$$\% \Delta W_0(ah_s; m) \equiv 100 \left(\frac{\bar{V}(a; m) - \bar{V}(a; 0)}{\bar{V}(a; 0)} \right),$$

$$\bar{V}(a; m) = \mathbf{1}_{\{a \geq a^*\}}(V_{U,1}(ah_s; m) - c) + \mathbf{1}_{\{a < a^*\}}V_{U,0}(a; m).$$

Figure 3 plots this percentage welfare change for two types: the lowest ability and the median ability type. As we can see in Figure 3, the distributional consequences of the minimum wage are extreme. We see that low skill types are severely adversely affected by higher minimum wages. As was argued previously, this is due entirely to the fact that the minimum wage rules out previously acceptable matches, resigning low-skill workers to long periods of unemployment. Using a Rawlsian criterion to evaluate the minimum wage policy intervention significantly recasts our analysis of the “optimal” minimum wage. In particular, looking at minimum wages in the range of \$20–22 per hour, we see that low ability types are significantly less well off than in the case with no minimum wage.

We now turn to the welfare effect of minimum wages when contact rates are determined endogenously by firms’ vacancy posting decisions. In this setting the minimum wage transfers surplus from firms to workers, which induces firms, through the free entry condition, to post

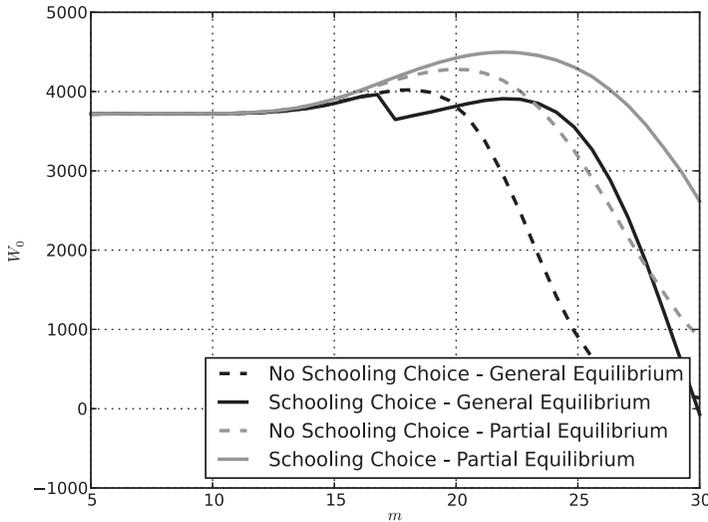


FIGURE 4

TOTAL WORKER WELFARE IN GENERAL EQUILIBRIUM (ENDOGENOUS k), WITH AND WITHOUT ENDOGENOUS SCHOOLING

fewer vacancies. Hence, the minimum wage forces the contact rate down in both schooling markets. This will serve to moderate the (average) welfare gains seen in the partial equilibrium case.

It is worth noting that the magnitudes of these results rest on the responsiveness of firms' vacancy posting to changes in the surplus they receive from a match. As is illustrated in Hagedorn and Manovskii (2008), this elasticity is tightly linked to the size of firms' profits. Within the model there are two key parameters that may effect the size of profits: the flow value of unemployment, b , and the worker's bargaining share, α . In particular, Hornstein et al. (2011) note that to match frictional wage dispersion, structural estimates of search models often produce small or even negative estimates of b . Aside from philosophical differences on what constitutes "reasonable" estimates of b , we are unconcerned by this for two additional reasons. First, this analysis applies only to models without on-the-job search; hence it is not clear if this critique applies in our case. Second, since we target labor's share of output, a small estimate of b will not affect the mean size of firms' profits, since other parameters (such as bargaining power) will adjust to keep the mean share of payments to labor at their empirical level. Looking outside the model, it is reasonable to expect that including payments to capital, as in Hagedorn and Manovskii (2008), will increase the responsiveness of market tightness to the policy changes considered here. We apply this caveat to our analysis below.

An additional effect to consider in the case of endogenous schooling choices is the selection effect. As the schooling rate increases, the conditional expectation of ability in both markets decreases, which drives down contact rates even further. We can see both of these effects at play in Figure 4. Welfare returns to the minimum wage are dampened in both cases (with and without a schooling choice), which implies a smaller "optimal" minimum wage.

The visible kink in welfare for the general equilibrium case with endogenous schooling is due to a discontinuity in schooling choice; workers at this point pool into the high-schooling market. The sudden change in the composition of abilities forces a discontinuous decrease in the contact rate, which causes the drop in welfare that we see. This case highlights the importance of selection effects in our model. We further discuss this effect when considering the efficiency properties of the minimum wage.

To complete our analysis of worker welfare, in Figure 5 we replicate Figure 3 to see the effects of minimum wage changes across different types. The general patterns found are very similar

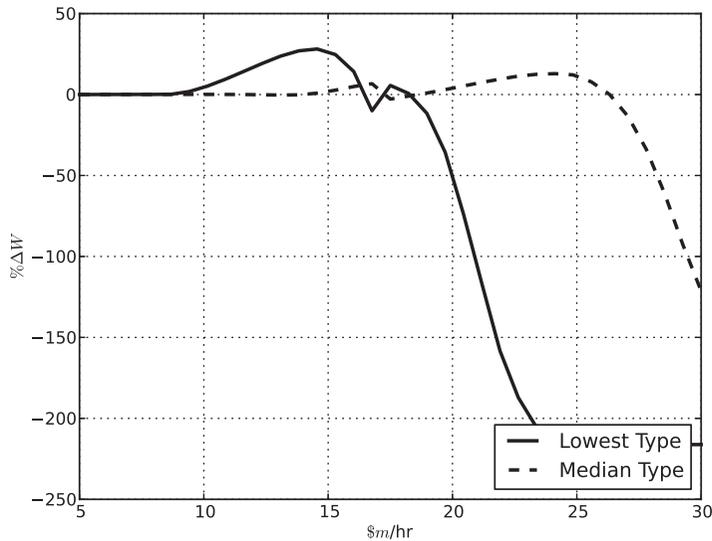


FIGURE 5

PERCENTAGE CHANGE IN WORKER WELFARE IN GENERAL EQUILIBRIUM FOR THE LOWEST AND MEDIAN WORKER TYPES

to the partial equilibrium case; low types suffer dramatically at choices of m close to what we would consider the “optimal” minimum wage for aggregate welfare.

We have seen that the minimum wage can bring improvements to worker welfare. However, are these changes efficient from the perspective of a social planner who attempts to maximize aggregate resources? We answer this question by computing the aggregate effects of minimum wages between 5 and 23 dollars.

Following our approach for worker welfare, we first conduct this experiment by fixing the schooling cutoff decision at the a^* implied by our model estimates. Figure 6 shows the high- and low-schooling components of the planner’s welfare criterion, W_1 , along with the contact rates $\lambda_{U,s}$ for $s = 0, 1$. We see that increasing the minimum wage brings efficiency gains in both markets. The optimal minimum wage in the low-schooling market is approximately 18, whereas the optimal minimum wage in the high-schooling market is approximately 28. Even though contact rates decrease in both markets, aggregate resources improve. This suggests that congestion externalities are being resolved in both markets. That is, firms are posting too many vacancies from the perspective of the planner. Imposing a minimum wage transfers surplus from the firm to the worker and corrects this, reducing the cost of vacancies in both markets. In Appendix A.2, we see that similar gains can be achieved by increasing worker bargaining power, similar to the well-known condition of Hosios (1990). Clearly, the hump shape in the planner’s welfare is produced by a trade-off, since at some point the loss in resources generated by lower contact rates will outweigh the reduction in vacancy posting costs. Additionally, the rising minimum wage rules out progressively more matches that generate positive social surplus but are unprofitable to the firm.

We next consider the general equilibrium case in which schooling choices are endogenized. Figure 7 shows the unemployment rate, total output (excluding vacancy costs), the schooling rate, and the planner’s measure of welfare, W_1 , as defined in Equation (12). The aggregate effects of this policy intervention are once again quite striking. We highlight several points of interest. First, the schooling rate is highly nonmonotonic, displaying a hump shape before jumping to 100% at around 15. This extreme response is produced in large part due to our assumption that all individuals face the same cost of schooling, c . Second, the function exhibits two similar “humps.” The first appears approximately at this jump point as workers shift discontinuously into the high-schooling market. The second, and optimal, point occurs at approximately 22, at

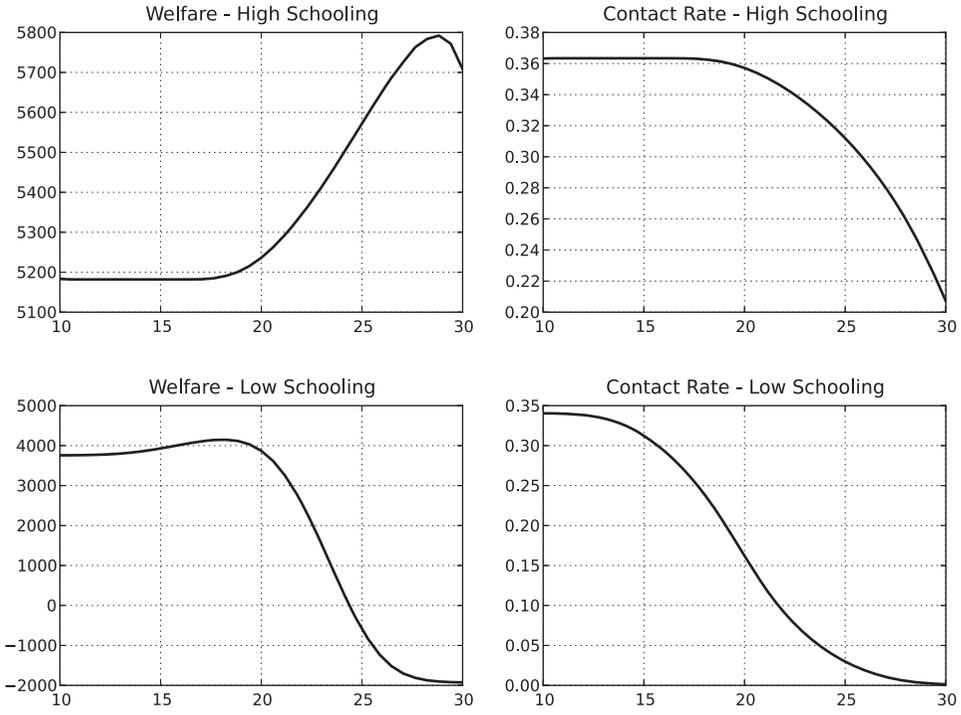


FIGURE 6

PLANNER WELFARE AND CONTACT RATES FOR THE MINIMUM WAGE POLICY EXPERIMENT WITH FIXED SCHOOLING

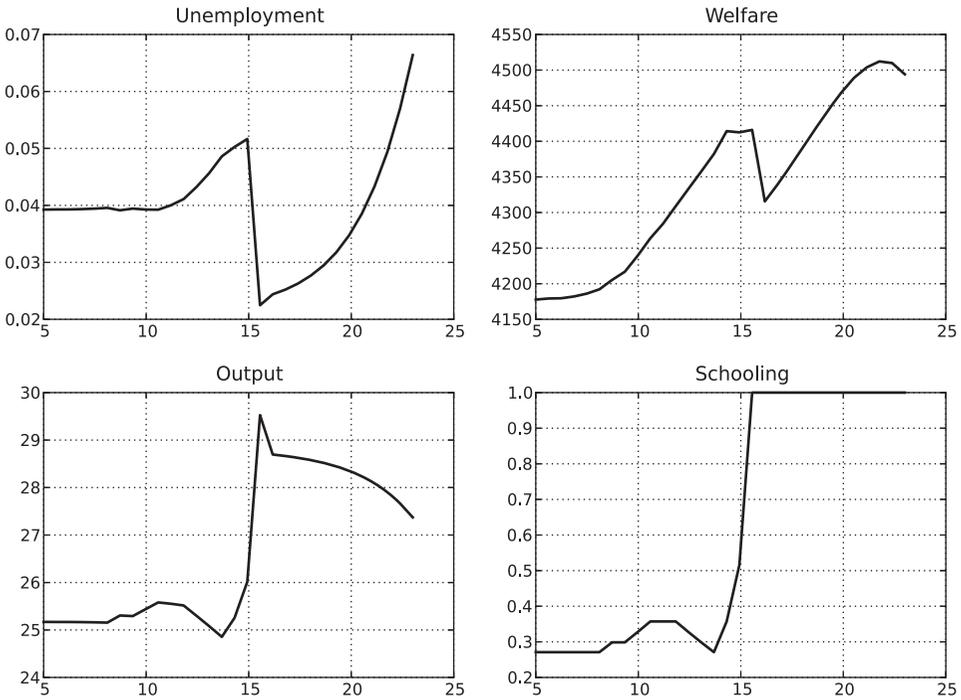


FIGURE 7

PLANNER WELFARE, OUTPUT, UNEMPLOYMENT, AND THE SCHOOLING RATE FOR THE MINIMUM WAGE POLICY EXPERIMENT WITH ENDOGENOUS SCHOOLING

which point everyone in the market has chosen to invest in schooling. Aside from the dramatic shift in schooling, the general pattern of the results is the same as for the case of fixed schooling.

To understand the nonmonotonicity in schooling rates, we look at the countervailing forces at play. By equalizing all bargained wages less than m in both markets, the relative value of the low-schooling market should increase. However, for any worker of type ν , the lower bound on viable matches will increase as m increases, as per Equation (13). This effect will reduce the relative value of the low-schooling choice, since more matches will bind against this constraint than in the high-schooling market. Finally, although m reduces the value of vacancy posting to firms in both markets, the extent of this is more severe in the low-schooling market, where average worker skills are lower. This implies that the effect on contact rates through vacancy posting is stronger in the low-schooling market, a result that is in line with the results of our example with fixed schooling shown in Figure 6.

Why are the effects on the schooling rate so extreme in this model? This relates to an instability produced by selection effects. As the schooling rate increases, due to increases in m , the relative mass of abilities in the low-schooling market shifts downward. This decreases the relative profitability of successful matches as well as the probability of finding an acceptable match from the pool of unemployed workers. To satisfy the free-entry condition, the contact rate must decrease, which forces the schooling rate to increase further. Eventually, we reach a point at which it is impossible for an active low-schooling market to exist.

The extreme predictions from the general equilibrium experiment make it difficult to interpret the welfare results of the minimum wage experiment normatively. It is obviously not true that a minimum wage of \$15 or more would induce all workers to attend college. These results rely quite stringently on the steady state and timing assumptions of worker and firm choices as well as the parametric assumptions of the model. In particular, workers are pushed into the schooling decision because firms abandon the low-schooling market altogether and vice versa. The disappearance of the low-schooling market is in large part an artifact of our assumption of constant costs of schooling. As in models of labor market signalling, there is good reason to expect that the costs of higher education are not constant in the population, and that the cost of schooling is negatively related to one's ability endowment. Unfortunately, the identification of a bivariate distribution of ability and schooling cost is beyond the realm of possibility with the data to which we have access.

9.2. Schooling Subsidies. We now examine the welfare consequences of offering a subsidy that reduces the cost of schooling c . We assume that a subsidy (or tax) τ is offered, so that workers who choose to invest will pay $c - \tau$, whereas the planner pays the difference τ for those who take up schooling. We assume that the subsidy is financed by lump sum transfers, the nature of which we do not specify, except that we require a balanced budget. This ensures that welfare gains are not created artificially by violating the resource constraint. In this case, the natural criterion for welfare is W_1 , the planner's value, since this criterion is not affected by the choice of how the lump sum tax burden is assigned. In fact, it is easy to see that in this case, since financing of the subsidy does not appear in the output equation, this policy is equivalent in aggregate welfare terms to forcing a change in a^* .¹⁶

Looking to Figure 8 for the results of this experiment, we see that a positive schooling subsidy improves planner welfare, reduces aggregate unemployment, and increases production in the economy. There is a sharp contrast between this aggregate result and the effect of the subsidy on each individual schooling market. In Figure 9, we plot the unemployment rate, U_s , and the flow of total resources, Y_s , as defined in Equation (11), for each market $s = 0, 1$. We see that in each submarket the subsidy increases unemployment and decreases total resources. To see why, recall that as τ increases, more workers are shifted from the low-schooling to the high-schooling market. As noted already, this shifts the conditional expectation of worker ability in

¹⁶ Of course, if we assumed that the subsidy was financed by a tax on labor earnings or vacancy posting, then this would no longer be the case.

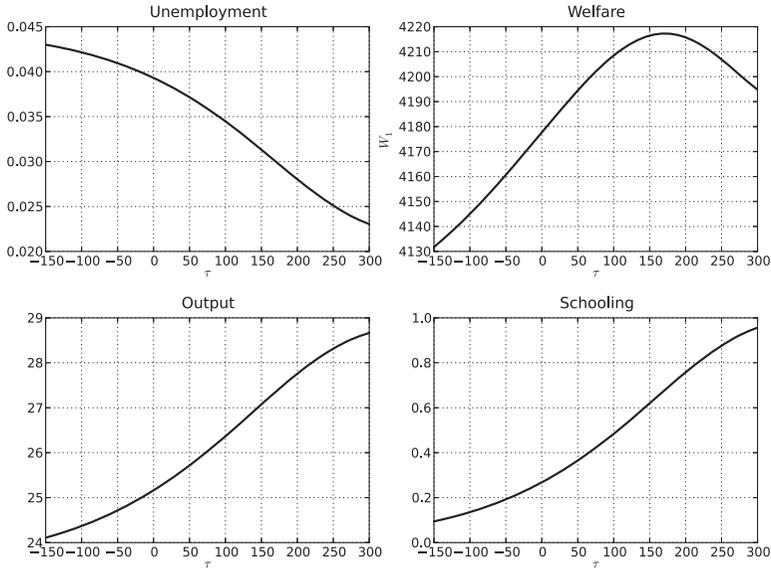


FIGURE 8

PLANNER WELFARE, OUTPUT, UNEMPLOYMENT, AND THE SCHOOLING RATE FOR THE SCHOOLING SUBSIDY POLICY EXPERIMENT (ENDOGENOUS SCHOOLING)

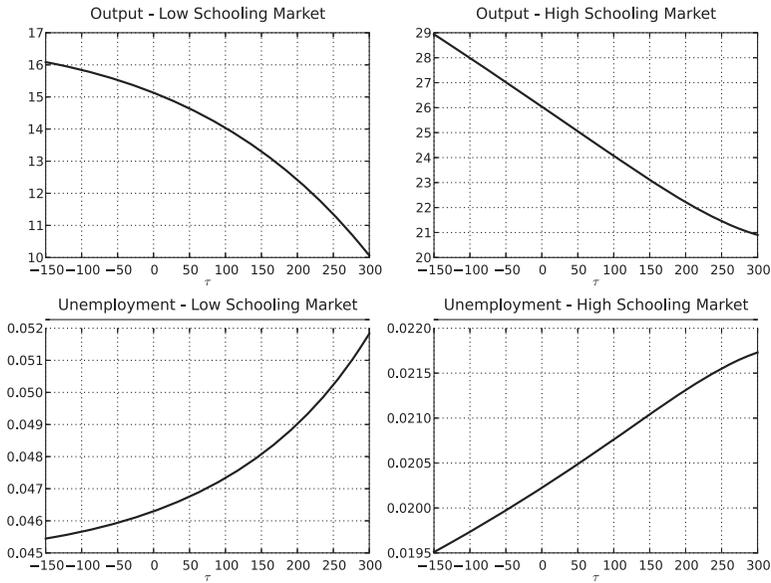


FIGURE 9

TOTAL RESOURCES AND UNEMPLOYMENT IN EACH SCHOOLING MARKET

both markets downwards. Accordingly, the free entry condition requires that firms post fewer vacancies and the contact rate decreases. The end result is that the mean flow of resources produced by each submarket decreases and the unemployment rate increases.

If this is true of each submarket, how can the aggregate effect across both markets look so different? The answer here combines the fact that (1) the flow of resources is higher and the unemployment rate is lower in the high-schooling market and (2) the schooling subsidy shifts

workers from low to high schooling. Therefore, there is a *composition effect* that improves output and unemployment in the average across both submarkets. To make this clear, consider the expression for the flow of total resources

$$Y = (1 - P_1) \cdot Y_0 + P_1 \cdot Y_1.$$

Taking a derivative with respect to the policy variable τ we get

$$\frac{\partial Y}{\partial \tau} = (1 - P_1) \cdot \frac{\partial Y_0}{\partial \tau} + P_1 \frac{\partial Y_1}{\partial \tau} + \underbrace{\frac{\partial P_1}{\partial \tau} (Y_1 - Y_0)}_{\text{Composition Effect}}.$$

Hence, in our application, even though Y_0 and Y_1 decrease in response to increases in τ , the composition effect from moving workers into the high-schooling market dominates the first two terms in the expression above.

10. CONCLUSION

In this article, we have developed a labor market model of search, matching, and bargaining that allows for pre-entry productivity-enhancing investments by workers. Although the model is highly stylized, it incorporates elements of the schooling investment decision that have not previously been investigated empirically. Individuals on the supply side of the model made investment decisions, which were assumed to be binary, and their choices depended on all parameters characterizing the labor market environment as well as institutional characteristics such as the presence of a minimum wage and schooling subsidies. Schooling investment enhanced the native ability of the individual, and flow productivity in the model was the product of individual ability and match productivity. We estimated two specifications of the model that includes on-the-job search, which differed in their bargaining protocols. Our estimates indicated that the population variances of ability and match heterogeneity were roughly the same. We opted to utilize estimates from the model in which firms did not renegotiate employment contracts with workers, since the overall goodness of fit was better under this specification and it produced more reasonable estimates of key model parameters.

We used the estimates to conduct two types of policy experiments. In the first, we examined the impact of minimum wages on two welfare measures, one of which was a general measure of worker welfare and the other was a measure of aggregate productivity. Since wages are relatively high for both schooling classes, only very high minimum wages were binding, and we found that there were welfare gains in terms of both welfare measures from imposing binding minimum wages. This result persists whether one considers contact rate parameters as fixed or not and allowing adjustment in schooling levels or not. Many of these results are qualitatively similar to what was found by Flinn (2006), although he assumed no on-the-job search, no individual heterogeneity in ability, and did not allow for schooling decisions. The key difference, however, is that we allow for individual heterogeneity in flow productivity and find that minimum wages have very unequal welfare consequences across the ability distribution. Low-ability individuals, even after increasing their schooling attainment, are the ones who are increasingly priced out of the market as the minimum wage increases. The use of a Rawlsian criterion for choosing the minimum wage that maximizes the ex ante welfare of the least able labor market participant would yield very different implications regarding the optimal minimum wage.

The results of our schooling cost subsidy experiment yielded similar conclusions, since the schooling subsidy induces more individuals to enter the high-schooling market. This increases average productivity, which is beneficial, and it also decreases vacancy creation in both markets. Given our estimates, this is actually beneficial, since we found that firms were initially posting too many vacancies in each market. We did not explicitly consider how the schooling subsidies were to be financed, other than suggesting the use of a lump-sum tax. If, more realistically, the subsidy was paid for through a tax on labor earnings, the impact of the subsidy may be less

beneficial due to the distortion. We did not consider the possibility of using both a minimum wage and a schooling subsidy as policy instruments, but such an analysis is feasible and may produce results of some policy interest.

We have estimated one of the first general equilibrium models of schooling investment in a search environment and believe that our results are of importance for policy debates regarding minimum wages and tuition subsidies. Although both types of policies may have beneficial effects at the aggregate level, measured either in terms of individual welfare or efficiency, their payoffs are very different across subpopulations. High levels of the minimum wage may disproportionately harm those they are designed to help while proving more beneficial for those with intermediate levels of ability. Similarly, untargeted tuition subsidies are pure rents to those who would have gone to school without the subsidy and disproportionately help those with intermediate levels of ability. These results suggest that more careful analysis of the distributional impacts of these types of policies is required prior to their implementation.

APPENDIX: EFFICIENCY PROPERTIES OF THE MODEL

How might the two policy interventions we consider in this article produce efficiency gains? To understand this question, we examine three sources of inefficiency in the model.

A.1. The holdup Problem. When worker bargaining power is less than one ($\alpha < 1$), one concern is that the worker's return to education is less than that of the planner and hence the schooling take-up rate in equilibrium is too low. This argument does not consider the fact that the subsequent labor markets suffer their own inefficiencies. In particular, firms in either market may post too many or too few vacancies.¹⁷ The relative size of these inefficiencies may offset or even outweigh the benefit of moving workers into the high education market. In partial equilibrium, it is simply the relative difference in vacancy costs per worker that factors into welfare calculations.

We examine the change in welfare when the cutoff decision in schooling is adjusted. To make the above point clear, we calculate W_1 both with and without the vacancy posting cost in (11), which we call VC_s . Further, to isolate both effects, we do not recompute the contact rates in each market in general equilibrium.

Figure A1 produces a dramatic contrast. Ignoring the vacancy posting cost (and hence the relative inefficiencies of each market) suggests that the optimal schooling rate is about 70%. However, once we include the vacancy posting cost, we obtain the opposite conclusion: that the equilibrium schooling rate is too high. In this experiment, the cost to the planner of moving a worker from the low- to the high-schooling market is $(VC_1 - VC_0)/\rho + c$. We calculate the difference in present value vacancy cost per worker, $(VC_1 - VC_0)/\rho$, to be \$388.46 in hourly wage value, which is nearly one-third of the schooling cost. This magnitude explains why including these costs dramatically changes our analysis.

This contrast highlights an important result. In this model, vacancy posting decisions are just as relevant for welfare as the schooling decision. The question of whether or not this firm decision is efficient is the subject of the next section.

A.2. Congestion Externalities. Although addressing the holdup problem is a key policy goal, our assumptions concerning the labor market introduces another important dynamic. A well-known inefficiency in search and matching models arises from two wedges between the firm's vacancy posting decision and that of the planner. The first is the wedge between the joint surplus to the worker and the firm and the surplus to the planner of a match. Since the firm claims only a fraction $(1 - \alpha)$ of the surplus generated by a match, the surplus to the firm tends to be less than the joint surplus in the planner's problem. This wedge would induce firms to post too few vacancies. To countervail this effect, firms do not consider the marginal effect

¹⁷ See the next section on congestion externalities for further discussion.

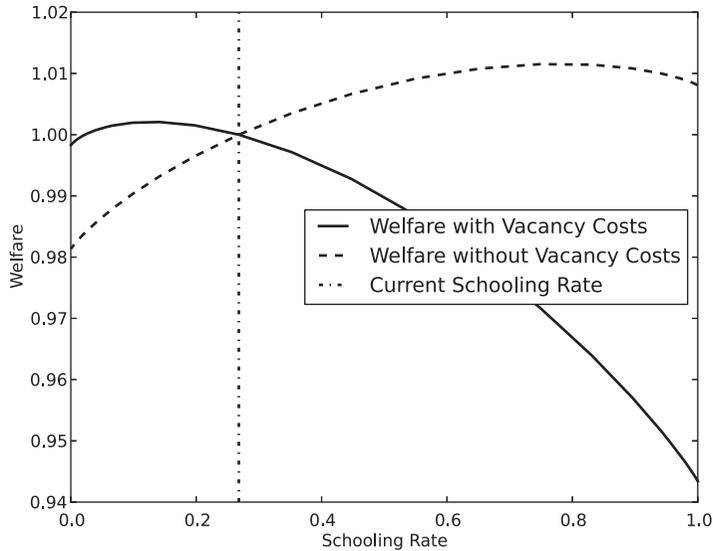


FIGURE A1

EFFECT OF CHANGES IN THE SCHOOLING RATE ON PLANNER WELFARE IN PARTIAL EQUILIBRIUM (FIXED k) INCLUDING AND EXCLUDING THE COST OF VACANCIES

that their posting decisions have on the contact rate in equilibrium. This congestion externality leads firms to post too many vacancies. In rare cases, these two effects will perfectly counteract each other and the market equilibrium is efficient. In the market without on-the-job search and homogeneous matches, this occurs when $\alpha = \delta$, the well-known Hosios (1990) condition.

In our application, such a clean condition is not obtainable. However, the same principle leads us to conjecture that adjustments in the bargaining power will result in welfare improvements. To see this, we solve for equilibrium market tightness k_s for each market $s = 0, 1$, keeping the schooling rate fixed. Figure A2 shows the effect of congestion externalities at work. As surplus is transferred from the firm to the worker, welfare improves. This suggests that, ceterus paribus, firms post too many vacancies in equilibrium. In addition, we see that these welfare improvements are more significant when the distance between α and δ , the elasticity of the firm's contact rate with respect to k , is greater. This is suggestive of a "Hosios-like" condition. These observations are all important when considering the welfare effect of minimum wages in Section 9.1.

A.3. Selection Effects. The final effect we must consider takes place as follows. Since the schooling decision is characterized by a cutoff type a^* that is indifferent between each level of schooling, the average stock of human capital in each market is characterized by a pair of truncated expectations, $E[a \mid a < a^*]$, $E[a \mid a > a^*]$. As the cutoff a^* decreases, the schooling rate increases and both expectations decrease.¹⁸

Since increasing the schooling rate lowers mean ability in each market, this lowers the mean payoff to a match for the firm and hence decreases the contact rate for workers in both markets. This effect will temper any welfare gains from improving the schooling rate. We see this effect at play in Section 9.2, when evaluating schooling subsidies.

A.4. Robustness: Different Elasticities. In this section, we briefly examine whether our welfare experiments are robust to different choices in the match elasticity, δ_s . In particular, we would like to know how sensitive our results are to the assumption that $\delta_0 = \delta_1$ across both

¹⁸ Of course, the mean stock of human capital in the economy will increase.

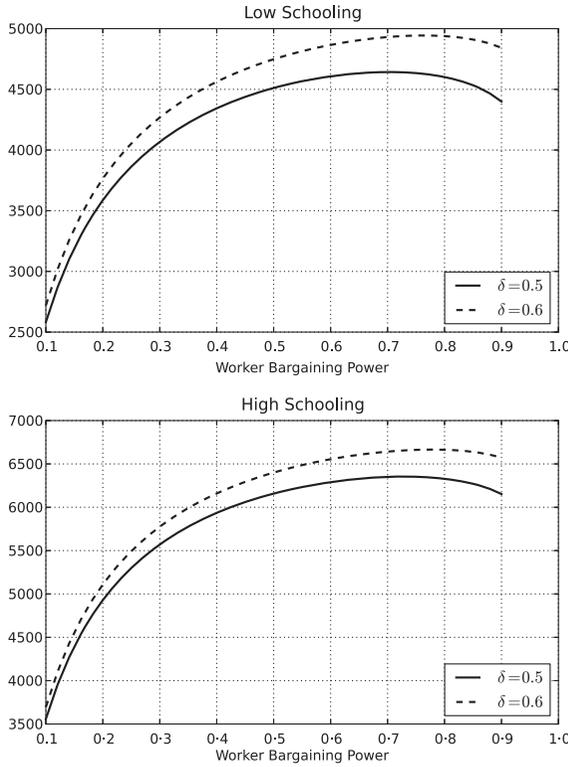


FIGURE A2

EFFECT OF CHANGES IN WORKER BARGAINING POWER ON PLANNER WELFARE IN GENERAL EQUILIBRIUM WITH FIXED SCHOOLING FOR THE HIGH- AND LOW-SCHOOLING MARKETS

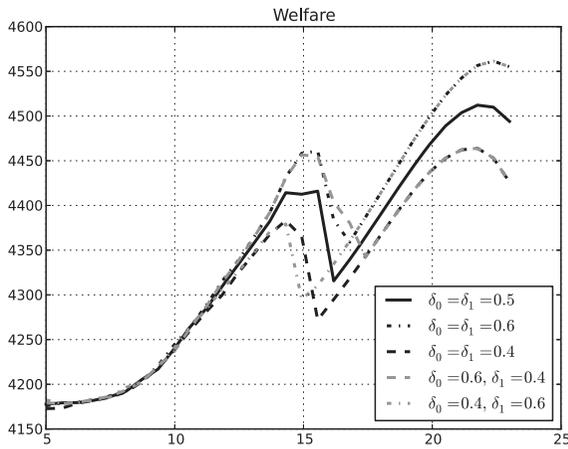


FIGURE A3

MINIMUM WAGE EXPERIMENT IN GENERAL EQUILIBRIUM FOR DIFFERENT CHOICES OF (δ_0, δ_1)

markets. To evaluate this, we repeat the minimum wage experiment from Section 9.1 with several choices of elasticities (δ_0, δ_1) . Figure A3 shows the results.

We find that the results are qualitatively the same. It does appear that the choice of the pair (δ_0, δ_1) affects the magnitude of welfare gains but does not seem to greatly affect the choice of an optimal minimum wage. The one pattern that emerges was predicted by our analysis in

Section A.2; the greater the distance between α and δ , the more significant the welfare gains from increasing m . Hence, welfare gains to the right of the discontinuity (where the schooling rate jumps to 100%) are greatest when $\delta_1 = 0.6$ and smallest when $\delta_1 = 0.4$. Similarly, to the left of this discontinuity, we see the greatest gains in welfare when $\delta_0 = 0.6$. The upper and lower bounds on the “optimal” minimum wage remain fairly tight between approximately 21.50 and 22.50.

REFERENCES

- ACEMOGLU, D., “A Microfoundation for Social Increasing Returns in Human Capital,” *Quarterly Journal of Economics* 111 (1996), 779–804.
- , “Training and Innovation in an Imperfect Labor Market,” *Review of Economic Studies* 64 (1997), 445–64.
- , AND R. SHIMER, “Holdups and Inefficiencies with Search Frictions,” *International Economic Review* 40 (1999), 827–49.
- BARTOLUCCI, C., “Gender Wage Gaps Reconsidered: A Structural Approach Using Match Employer–Employee Data,” *Journal of Human Resources* 48 (2013), 998–1034.
- CAHUC, P., F. POSTEL-VINAY, AND J.-M. ROBIN, “Wage Bargaining with On-the-Job Search: Theory and Evidence,” *Econometrica* 74 (2006), 323–64.
- CHARLOT, O., AND B. DECREUSE, “Self-Selection in Education with Matching Frictions,” *Labour Economics* 12 (2005), 251–67.
- , AND ———, “Over-Education for the Rich, Under-Education for the Poor: A Search-Theoretic Microfoundation,” *Labour Economics* 17 (2010), 886–96.
- CUNNINGHAM, J., “The Impact of Minimum Wages on Youth Employment, Hours of Work, and School Attendance: Cross-Sectional Evidence from the 1960 and 1970 Censuses,” in S. Rottenberg, ed., *The Economics of Legal Minimum Wages* (Washington, DC: American Enterprise Institute, 1981), pp. 88–123.
- DEY, M., AND C. FLINN, “An Equilibrium Model of Health Insurance Provision and Wage Determination,” *Econometrica* 73 (2005), 571–27.
- ECKSTEIN, Z., AND K. WOLPIN, “Estimating a Market Equilibrium Search Model from Panel Data on Individuals,” *Econometrica* 58 (1990), 783–808.
- , AND ———, “Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model,” *Review of Economic Studies* 62 (1995), 263–86.
- EHRENBERG, R., AND A. MARCUS, “Minimum Wages and Teenagers’ Enrollment–Employment Outcomes: A Multinomial Logit Model,” *Journal of Human Resources* 17 (1982), 39–58.
- FLINN, C., “Labour Market Structure and Inequality: A Comparison of Italy and the U.S.,” *Review of Economic Studies* 69 (2002), 611–45.
- , “Minimum Wage Effects on Labor Market Outcomes under Search, Bargaining, and Endogenous Contact Rates,” *Econometrica* 74 (2006), 1013–62.
- , AND J. HECKMAN, “New Methods for Analyzing Structural Models of Labor Force Dynamics,” *Journal of Econometrics* 18 (1982), 115–68.
- , AND J. MABLI, “On-the-Job Search, Minimum Wages, and Labor Market Outcomes in an Equilibrium Bargaining Framework,” working paper, NYU, 2009.
- GEMICI, A., AND M. WISWALL, “Evolution of Gender Differences in Post-Secondary Human Capital Investments: College Majors,” *International Economic Review* 55 (2014), 23–56.
- HAGEDORN, M., AND I. MANOVSKII, “The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited,” *American Economic Review* 98 (2008), 1692–706.
- HECKMAN, J., L. LOCHNER, AND C. TABER, “Explaining Rising Wage Inequality: Explanations with a Dynamic General Equilibrium Model of Earnings with Heterogeneous Agents,” *Review of Economic Dynamics* 1 (1998), 1–58.
- HORNSTEIN, A., P. KRUSELL, AND G. VIOLANTE, “Frictional Wage Dispersion in Search Models: A Quantitative Assessment,” *American Economic Review* 101 (2011), 2873–98.
- HOSIOS, A., “On the Efficiency of Matching and Related Models of Search and Unemployment,” *Review of Economic Studies* 57 (1990), 279–98.
- KEANE, M., AND K. WOLPIN, “The Career Decisions of Young Men,” *Journal of Political Economy* 105 (1997), 473–522.
- KRUEGER, A., “Measuring Labor’s Share,” *American Economic Review* 89 (1999), 45–51.
- LEE, D., “An Estimable Dynamic General Equilibrium Model of Work, Schooling, and Occupational Choice,” *International Economic Review* 46 (2005), 1–34.
- , AND K. WOLPIN, “Intersectoral Labor Mobility and the Growth of the Service Sector,” *Econometrica* 47 (2006), 1–46.

- MALCOMSON, J., "Contracts, holdup, and Labor Markets," *Journal of Economic Literature* 35 (1997), 1916–57.
- MORTENSEN, D., AND C. PISSARIDES, "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies* 61 (1994), 397–415.
- NEUMARK, D., AND O. NIZALOVA, "Minimum Wage Effects in the Longer Run," *Journal of Human Resources* 42 (2007), 435–52.
- , AND W. WASCHER, "Minimum Wage Effects on Employment and School Enrollment," *Journal of Business and Economics Statistics* 13 (1995), 199–206.
- , AND ———, "Minimum Wages and Skill Acquisition: Another Look at Schooling Effects," *Economics of Education Review* 22 (2003), 1–10.
- PETRONGOLO, B., AND C. PISSARIDES, "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature* 39 (2001), 390–431.
- PISSARIDES, C., *Equilibrium Unemployment Theory*, 2nd ed. (Cambridge, MA: MIT Press, 2000).
- POSTEL-VINAY, F., AND J.-M. ROBIN, "Equilibrium Wage Dispersion with Heterogeneous Workers and Firms," *Econometrica* 70 (2002), 2295–350.
- SHIMER, R., R. ROGERSON, AND R. WRIGHT, "Search Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature* 43 (2005), 959–88.
- SULLIVAN, P., "A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search," *International Economic Review* 51 (2010), 289–317.
- VAN DEN BERG, G., AND G. RIDDER, "An Empirical Equilibrium Search Model of the Labor Market," *Econometrica* 66 (1998), 1183–222.