Matching, Sorting and Wages

Jeremy Lise Costas Meghir Jean-Marc Robin

January 2012

Preliminary and Incomplete

Abstract

We develop an empirical search-matching model with productivity shocks so as to analyze policy interventions in a labour market with heterogeneous agents. To achieve this we develop an equilibrium model of wage determination and employment, which is consistent with key empirical facts. As such our model extends the current literature on equilibrium wage determination with matching and provides a bridge between some of the most prominent macro models and microeconometric research. The model incorporates long-term contracts, on-the-job search and counter-offers, and a vacancy creation and destruction process linked to productivity shocks. Importantly, the model allows for the possibility of assortative matching between workers and jobs, a feature that had been ruled out by assumption in the empirical equilibrium search literature to date. We use the model to estimate the potential gain from optimal labour market policies, as well as the efficiency and redistributive effects of typical unemployment insurance and minimum wage policies.

1 Introduction

Understanding the impact of labour market regulation is both important and complex. It is important because such regulations are pervasive and take the form of anything from firing restrictions to minimum wages all around the world. It is complex because regulations affect the equilibrium in the labour market, changing the wage structure, employment and the type of jobs that are available. The typical justification of such regulations are labour market imperfections and frictions: within a competitive framework any such regulation would be welfare reducing, would typically reduce employment and would increase wages. For a complete empirical understanding of the relative merits of such regulation we need a framework which at the same time allows for the possibility that some regulation is optimal, but does not necessarily lead to that conclusion, which should be an empirical outcome.

Our aim is to develop and estimate a search-matching model with productivity shocks so as to analyze policy interventions in a labour market with heterogeneous agents. To achieve this we specify an equilibrium model of wage determination and employment, which is consistent with key empirical facts. As such our model extends the current literature on equilibrium wage determination with matching and provides a bridge between some of the most prominent macro models (Diamond (1982); Mortensen (1982); Pissarides (2000)) and microeconometric research.

In our economy individual workers are matched to individual jobs. Workers differ from each other according to a productivity relevant characteristic, which is taken as fixed over time. Firms are also heterogeneous and their productivity is subject to possibly persistent shocks; these can be thought of as random shifts in the demand curve for their product or as price shocks in other inputs. The production function allows for complementarity between worker and job characteristics, leading to the possibility of sorting in the labour market. Jobs can remain completely idle, or post a vacancy, depending on a trade-off between the cost of posting and the expected discounted stream of benefits from filling the job. Any worker-job pair that meet will form a match when the surplus is non-negative - this accounts for the option value of keeping the vacancy and thus meeting a better partner. Finally, workers receive job offers while unemployed or employed at exogenously determined rates (that differ in the two states) and endogenously determined market tightness.

Wage contract are long-term contracts which can be renegotiated only by mutual agreement, i.e. when one of the partners can make a credible threat to walk away from the match, and employers counter outside offers. We use the sequential auction framework of Postel-Vinay and Robin (2002) extended to allow for an extra layer of Nash bargaining as in Dey and Flinn (2005) and Cahuc, Postel-Vinay, and Robin (2006). In addition, firm-level productivity shocks may move the current contract outside the bargaining set. We follow Postel-Vinay and Turon's (2010) extension of Postel-Vinay and Robin's sequential auction model to model contract renegotiation upon productivity shocks.

Our model offers an empirical framework for understanding employment and wage determination in the presence of firm-worker complimentarities, search frictions and productivity shocks; as a result it offers a way for evaluating the extent to which regulation can be welfare improving and can evaluate the impact of specific policies such as minimum wages or restrictions on firing. In a search framework protecting workers from being fired can have ambiguous effects on employment. Our framework will allow this effect to be quantified. But, in addition, it allows us to analyze the effect of regulation on the distribution of wages and profits thus showing who pays and who benefits from such a policy in this non-competitive environment.

The model generates wage dynamics. These arise from multiple sources: productivity shocks may lead to wage renegotiation or to match dissolution and a period of unemployment before reentering at a different wage; on-the-job search may lead to outside offers and hence to wage changes or mobility. The wage dynamics are thus explained by the structure of the model. The question is whether the dynamics match those documented in papers such as MaCurdy (1982); Abowd and Card (1989); Meghir and Pistaferri (2004). If they do this offers a structural interpretation to such processes based on productivity shocks and behavioral responses in wage setting.

It is now well understood that search models can give rise to wage dispersion even if workers are homogeneous (see Burdett and Judd, 1983; Burdett and Mortensen, 1998; Postel-Vinay and Robin, 2002). However, matching models with heterogeneous workers and jobs is a relatively new topic of interest fueled by the need to understand dispersion of wages of similar individuals. In general, workers differ by the numbers of years of education and experience, and jobs differ by the type of industry. There is thus a large body of empirical evidence showing that wages differ across industries, thus indicating that a matching process is at work in the economy (see for example Krueger and Summers, 1988). Static, competitive equilibrium models of sorting (Roy models) have been estimated by Heckman and Sedlacek (1985); Heckman and Honore (1990), while Moscarini (2001); Sattinger (2003) explore theoretical extensions of Roy models with search frictions. However, there must be an enormous amount of differences between workers and between jobs that are not accounted for by observables in the data. Marriage models with a continuum of heterogeneous agents in a frictional environment are studied in Sattinger (1995); Lu and McAffee (1996); Shimer and Smith (2000); Atakan (2006).[1] To the best of our knowledge, there have not yet been any empirical applications of assignment models with transferable utility in a frictional environment with realistic distributions of heterogeneity.

How much sorting is there with respect to these unobserved characteristics? Abowd, Kramarz, and Margolis (1999, AKM) and Abowd, Kramarz, Pérez-Duarte, and Schmutte (2009) use French and U.S. matched employer-employee data to estimate a static, linear log wage equation with employer and worker fixed effects (by OLS). They find a small, and if anything negative, cross-sectional correlation between job and worker fixed effects. Abowd, Kramarz, Lengermann, and Roux (2005) document the distribution of these correlations calculated within industries. In the U.S. 90% of these correlations range between -15% and 5%, and in France between -27% and -5%. These negative numbers, although hard to interpret, offer *prima faciae* evidence of no positive sorting. However, the evidence based on the log-linear decomposition used by AKM should not be interpreted as evidence that there is no sorting: the person and firm effects which are estimated from the linear log wage equation are complicated transformations of the underlying individual-specific, unobserved productivity-relevant characteristics; A structural model is thus required to recover the true underlying joint distribution of characteristics. **?** present some evidence that a matching model inspired from Shimer (2005) could both generate sorting on unobservables and the sort of empirical regularity that they find. More recently, de Melo (2009) has proposed a matching model, extending Shimer and Smith (2000) to allow for on-the-job search, that also produces the same prediction. Our framework allows us to investigate empirically whether sorting actually is important in practice.

[1]Sattinger develops a framework but does not prove the existence of an equilibrium. Lu and McAffee prove the existence for a particular production function ($f(x,y) = xy$). Shimer and Smith prove the existence of an equilibrium in a more general setup and derive sufficient conditions for assortative matching. Atakan shows that Becker (1973) complementarity condition for positive sorting is sufficient if there exist explicit search costs.

2 The Model

In the economy there are a fixed number of individuals. The number of jobs are determined so that the marginal entrant makes zero expected profits. Individuals may be matched with a job and thus working, or they may be unemployed job seekers. Jobs on the other hand may be in two different states. First they may be matched with a worker, in which case output is produced. Second, they may be vacant and waiting for a suitable worker to turn up. Individuals all have different levels of human capital, indexed by and jobs differ from each other according to some productivity relevant characteristic y. Output depends on the characteristics of both sides with possible complementarities. Crucially though, productivity follows a first order Markov process, which leads to the value of the match changing, with consequences for wage dynamics, worker mobility, job creation, and job destruction that are at the centre of our model.

Individuals are infinitely lived and maximize their discounted income; jobs maximize profits. When a job and a worker meet and the total match surplus is positive, the worker is hired and thus the match is formed. At this point the worker is paid a wage consistent with the reservation value of the best alternative option, *plus* a share of the excess match surplus. This process is discussed in detail in the next section.[2] A further important feature is that a shock to job productivity may trigger a wage renegotiation. This will happen if under the new productivity the match surplus remains positive but the wage is too high under the new conditions. On the other hand there is no incentive for the firm to accept to renegotiate when there is a shock increasing the surplus, but as we shall show this will increase the value of being employed in this job because of the prospect of future wage increases. Finally we close the model by a free entry condition: all production lines, whether active or not face a cost of posting a vacancy. The marginal job, which has the lowest productivity y, has zero expected profits from posting a vacancy.

[2] In parallel work Lentz (2010) and Bagger and Lentz (2008) consider a model of on-the-job search with endogenous search intensity, where all workers match with any job when transiting from unemployment and sorting is the result of differing returns to search effort by worker type. An important distinction from our work is that in the limit if frictions disappear there is no sorting (all workers are at the best firm) while the limit as frictions disappear in our model implies perfect sorting.

2.1 Heterogeneity and production

Each individual worker is characterized by a level of permanent productivity, which we denote by x, observable by all agents, but not the researcher. We can normalize x to be uniformly distributed on $[0, 1]$, without any loss of generality. There are L individuals of which U are unemployed. We also denote by $u(x)$ the (endogenous) measure of x among the unemployed.

Jobs are characterized by a productivity parameter y, which is also observable by all agents. Given the unobservable nature of y in the data we use, we normalize y to be uniformly $[0, 1]$ distributed, and allow the scale to be absorbed by the production function parameters, as will become clear later.

There are N jobs in the economy, a number that is endogenously determined. The endogenous measure of vacant posts is $v(y)$ and the resulting number of vacancies is denoted by V.

In a given job there are idiosyncratic productivity shocks. These may reflect changes in the product market (shifts in the demand curve) or events such as changes in work practices or investments (which here are not modeled). We assume that y fluctuates according to a jump process: At a rate δ such shocks arrive and a new productivity level y' is drawn from the uniform distribution on $[0, 1]$. Thus persistence of the shocks is controlled for by the size of δ.

A match between a worker of type x and a job of productivity y produces a flow of output $f(x, y)$; we will specify this function to allow for the possibility that x and y are complementary in production, implying that sorting will increase total output. However, we wish to determine this empirically, as it is important both for understanding the labor market and for evaluating the potential effects of regulation.

Equilibrium will result in a joint distribution of x and y, the distribution of matches. We denote this by $h(x, y)$. A simple adding up constraint implies that all matched workers and the unemployed sum up to L:

$$\int h(x, y)\, dy + u(x) = L. \tag{1}$$

Similarly all matched jobs, those posted vacancies and the idle jobs should add up we can write an equivalent relationship between the distribution of job productivities, active matches and vacancies,

$$\int h(x,y)\,dx + v(y) = N. \tag{2}$$

In both cases the relationship is essentially an accounting identity. Finally, matches can end both endogenously, as we characterize later, and exogenously. We denote by ξ the rate of exogenous job destructions, defined as an event that is unrelated to either worker or job characteristics.

2.2 Match formation and rent sharing

Individuals and jobs are risk neutral and we assume efficiency, in the sense that any match where the surplus is positive will be formed. Under these conditions we can characterize the set of equilibrium matches and their surplus separately from the sharing of the surplus between workers and jobs.

Let $W_0(x)$ denote the present value of unemployment for a worker with characteristic x; this will reflect the flow of income (or money metric utility of leisure) when out of work and the expected present value of income that will arise following a successful job match. Similarly $\Pi_0(y)$ denotes the present value to a job of posting a vacancy arising from the expected revenues of employing a suitable worker.

There are a number of reasons for matches to have a surplus in our economy. First, the distribution y is exogenous in the sense that an entrant cannot decide to invest to acquire a particular level of productivity. This causes a source of surplus, which can be further reinforced by the possible complimentarities between x and y in production. The surplus of an (x,y) match is defined by

$$S(x,y) = \Pi_1(w,x,y) - \Pi_0(y) + W_1(w,x,y) - W_0(x). \tag{3}$$

Feasible matches (x,y) are such that $S(x,y) \geq 0$. When an unemployed worker x finds a vacant job y a match is formed if and only if $S(x,y) \geq 0$. Let $M(x) = \{y : S(x,y) \geq 0\}$ define the set of acceptable job types for any given x. The wage for a worker transiting from unemployment is $w = \phi_0(x,y)$ and we assume that it is set to split the surplus according to Nash bargaining with worker's bargaining parameter β:

$$W_1(\phi_0(x,y), x, y) - W_0(x) = \beta S(x,y). \tag{4}$$

Through the process of on-the-job search and offers/counteroffers her wage will subsequently grow. We allow β to be determined empirically to fit observed wages, since a $\beta = 0$ often implies negative wages: although in practice some workers may pay to get a job, because of the option value of such a move, this is never observed in the data.

2.3 Renegotiation

Wages can only be renegotiated by mutual agreement. This will happen when: either side has an interest to separate if they do not obtain an improved offer; and such an offer is possible without making the match non-viable for the party that has to give up some of their payoff. The events that can trigger renegotiation occur when a suitable outside offer is made, or when a productivity shock reduced the value of the surplus sufficiently. Consider first the impact of an outside offer.

2.3.1 Poaching

Workers get contacted by other jobs when employed, at a rate that is determined by search intensity and (endogenous) market tightness. We assume that incumbent employers respond to outside offers: a negotiation game is then played between the worker and both jobs as in Dey and Flinn (2005) and Cahuc, Postel-Vinay, and Robin (2006). If a worker x, currently paired to a job y such that $S(x, y) \geq 0$, finds an alternative job y' such that $S(x, y') \geq S(x, y)$, the worker moves to the alternative job; this is because the poaching firm can always pay more than the current one can match. Alternatively, if the alternative job y' produces less surplus than the current job, but more than the worker's share of the surplus at the current job, $W_1 - W_0(x) < S(x, y') < S(x, y)$, the worker uses the outside offer to negotiate up her wage.

In either case, the worker ends up in the higher surplus match, and uses the lower surplus match as the outside option when bargaining. The bargained wage in this case is $w = \phi_1(x, y, y')$ such that the worker obtains the entire surplus of the incumbent job and a share of the incremental surplus between the two jobs, i.e.

$$W_1\left(\phi_1\left(x, y, y'\right), x, y\right) - W_0(x) = S\left(x, y'\right) + \beta\left[S(x, y) - S\left(x, y'\right)\right], \tag{5}$$

where $S(x, y) \geq S(x, y')$. The share of the increased surplus β accruing to the worker will be determined empiricall. Finally, if $S(x,y') \leq W_1 - W_0(x)$, the worker has nothing to gain from the competition between y and y' because she cannot make a credible threat to leave, and the wage does not change.

The present value of the new wage contract $W_1(\phi_1(x,y,y'),x,y)$ does not depend on the last wage contracted with the incumbent employer, but only on the total surplus of the previous outside option. The continuation value for workers when the match is destroyed, by the worker moving to unemployment or an alternative job, is not a function of the last negotiated contract. For this reason the total surplus $S(x,y)$ is only a function of x and y. Therefore, a simple rent splitting mechanism applies. This follows from the Bertrand competition between the incumbent and the poaching firm that is induced by on-the-job search and disconnects the poached employees' outside option from both the value of unemployment and their current wage contract. Shimer (2006) provides an analysis where incumbent employers do not match outside offers and where the current wage determines the new contract.

In our approach there is an asymmetry between workers and firms because the latter do not search when the job is filled. As a result they do not fire workers when they find an alternative who would lead to a larger total surplus, nor do they force wages down when an alternative worker is found whose pay would imply an increased share for the firm. We decided to impose this asymmetry because in many institutional contexts it is hard for the firm to replace workers in this way. Moreover, we suspect that even when allowed firms would be reluctant to do so in practice.

2.3.2 Productivity shock

Another potential source of renegotiation is when a productivity shock changes y to y' thus altering the value of the surplus. If y' is such that $S(x,y') < 0$, the match is endogenously destroyed: the worker becomes unemployed and the job will post a vacancy.

Suppose now that $S(x,y') \geq 0$. The value of the current wage contract becomes $W_1(w,x,y')$ because future pay negotiations whether due to productivity shocks or competition with outside offers will be affected by the the new value of the match. However the current wage may or may not change. If the wage w is such that she is still obtaining at least as much as her outside option, without taking more than the new surplus ($0 \leq W_1(w,x,y') - W_0(x) \leq S(x,y')$), neither the worker nor the job has a credible threat to force renegotiation: both are better off with the current wage w being paid to the worker than walking away from the match to unemployment and to a vacancy respectively. In this case there will be no renegotiation. If, however, $W_1(w, x, y') - W_0(x) < 0$ or $W_1(w, x, y') - W_0(x) > S(x, y')$ (with $S(x, y') \geq 0$) then renegotiation will take place because a wage can be found that keeps them both bettr off within the match.

To define how the renegotiation takes place and what is the possible outcome we use a setup similar to that considered by MacLeod and Malcomson (1993) and Postel-Vinay and Turon (2010). The new wage contract is such that it moves the current wage the smallest amount necessary to put it back in the bargaining set. Thus, if at the old contract $W_1(w, x, y') - W_0(x) < 0$, a new wage $w' = \psi_0(x, y')$ is negotiated such that

$$W_1(\psi_0(x, y'), x, y') - W_0(x) = 0,$$

which just satisfies the worker's participation constraint. If at the new y', $W_1(w, x, y') - W_0(x) > S(x, y')$, a new wage $w' = \psi_1(x, y')$ such that the firm's participation constraint is just binding

$$W_1(\psi_1(x, y'), x, y') - W_0(x) = S(x, y').$$

Separations and pay changes may happen following both good shocks that increase the value of productivity y and bad shocks that decrease it. It is all about missmatch: what matters is what happens to the overall surplus. A good shock, for example, can imply that the quality of the match becomes worse and the surplus declines, since the outside option of the firm has changed and it may be worthwhile remaining vacant to find a better match. Conversely a bad shock can improve the surplus as the quality of the match improves. A shock that reduces the surplus can still lead to a wage increase to compensate the worker who is now matched with a job with fewer future prospects of wage increases. Thus what really matters as far as the viability of the match and the possible options for renegotiation is whether a shock improves or worsens a particular match leading to an increase or a decrease, respectively of the surplus. Wages respond to job specific productivity shocks, but not always in an obvious direction.

2.4 Value Functions

The next step in solving the model is to characterize the value functions of workers and jobs, which have been kept implicit up to now. These define the decision rules for each agent. We proceed by assuming that time is continuous.

2.4.1 Unemployed workers

Unemployed workers are always assumed to be available for work at a suitable wage rate. While unemployed they receive income or money-metric utility depending on their ability x and denoted by $b(x)$. Thus the present value of unemployment to a worker of type x is $W_0(x)$, which satisfies the option value equation:

$$rW_0(x) \quad = \quad b(x) + s_0\kappa\beta \int S(x,y)^+ v(y)dy, \tag{6}$$

where The discount rate is denoted by r and where we define in general $a^+ = \max(a, 0)$. Thus the integral represents the expected value of the surplus of feasible matches given the worker draws from the distribution of vacant jobs $v(y)$. She gets share β of this surplus and contacts a job at a rate $s_0\kappa$, the latter depending on aggregate market conditions.

2.4.2 Vacant jobs

Using (3), (4), and (5), the present value of profits for an unmatched job meeting a worker with human capital x from unemployment is

$$\Pi_1(\phi_0(x,y), x, y) - \Pi_0(y) = (1-\beta) S(x,y).$$

where $1-\beta$ represents the proportion of the surplus retained by the firm and $\Pi_0(y)$ is the value of a vacancy. A job meeting a worker who is already employed will have to pay more to attract her. Thus the value of a job with productivity y meeting an employed worker in a lower surplus match with productivity y' is

$$\Pi_1(\phi_1(x,y,y'), x, y) - \Pi_0(y) = (1-\beta)\left[S(x,y) - S(x,y')\right]$$

Based on this notation, the present value of an vacancy for a job with productivity y is[3]

$$\begin{aligned}(r+\delta)\Pi_0(y) = -c + \delta \int \Pi_0(y')dy' + s_0\kappa(1-\beta)\int S(x,y)^+ u(x)dx \\ + s_1\kappa(1-\beta)\iint \left[S(x,y) - S\left(x,y'\right)\right]^+ h\left(x,y'\right) dx\, dy' \end{aligned} \quad (7)$$

where the $[]^+$ notation is as defined above. The c is a per-period cost of keeping a vacancy open. In (7) the second term reflects the impact of a change in productivity from y to y'. The third term is the flow of benefits from matching with a previously unemployed worker. The fourth term is the flow of benefits from poaching a worker who is already matched with another job; the integration is over all possible y' that are less attractive to worker type x and would thus move to the job with type y. We shall later show that $\Pi_0'(y) > 0$ so that we can assume that the marginal job makes zero profit:

$$\Pi_0(0) = 0. \quad (8)$$

or

$$\begin{aligned}c = \delta \int \Pi_0(y')dy' + s_0\kappa(1-\beta)\int S(x,0)^+ u(x)dx \\ + s_1\kappa(1-\beta)\int \left[S(x,0) - S\left(x,y'\right)\right]^+ h\left(x,y'\right) dx\, dy'. \end{aligned} \quad (9)$$

This condition will determine the number of potential jobs N in the economy.[4]

[3] In the following integrals we have used the assumption that the distribution of productivities is uniform [0,1]

[4] Note that we could (and we had in a previous version) assumed N exogenously given and large enough so there would exist a positive threshold $\underline{y}$ such that $\Pi_0(\underline{y}) = 0$. All vacant job with $\Pi_0(y) < 0$ would thus decide to remain "idle" and not paying the cost of posting a vacancy until the next shock moving $\Pi_0(y)$ above 0.

2.4.3 The match output and joint surplus

In the web appendix we show that $S(x,y)$ is defined by the fixed point in the following equation

$$\begin{aligned}(r+\xi+\delta)\,S(x,y) = f(x,y) - b(x) - s_0\kappa\beta\int S\left(x,y'\right)^{+} v(y')dy' \\ + c - s_0\kappa(1-\beta)\int S(x,y)^{+} u(x)dx \\ - s_1\kappa(1-\beta)\iint \left[S(x,y) - S\left(x,y'\right)\right]^{+} h\left(x,y'\right)dx\,dy' \\ + s_1\kappa\beta\int \left[S\left(x,y'\right) - S(x,y)\right]^{+} v\left(y'\right)dy' + \delta\int S\left(x,y'\right)^{+} dy'. \quad (10)\end{aligned}$$

The important point to note from this expression is that the surplus of an (x,y) match never depends on the wage; the Bertrand competition between the two jobs for the worker ensures this, as discussed above. Thus the Pareto possibility set for the value of the worker and the job is convex in all cases, implying that the conditions for a Nash bargain are satisfied. This contrasts with Shimer's (2006) model, where jobs do not respond to outside offers and where the actual value of the wage determines employment duration in a particular job. This feature also has a computational advantage since the equilibrium distribution of matches can be determined without computing the pay rates.

In the web appendix we also show that the expected profits of a vacant job increase with productivity ($\Pi_0^{'}(y) > 0$) so long as the marginal priduct of y is positive ($\partial f(x,y)/\partial y > 0$).

2.4.4 Employed workers

The present value to the worker of a wage contract w for a feasible match (x,y) is denoted by $W_1(w,x,y)$. The wage $w = \phi_0(x,y)$ for someone being hired out of unemployment solves the equation $W_1(\phi_0(x,y),x,y) - W_0(x) = \beta S(x,y)$ for a feasible match (x,y). Following a job to job transition the new wage will depend on the productivity of both the previous job and the new one because the worker's outside option is the previous job. In this case the wage $w = \phi_1(x,y,y')$ solves $W_1(\phi_1(x,y,y'),x,y) - W_0(x) = S(x,y') + (1-\beta)[S(x,y) - S(x,y')]$. In either case we need to determine the function $W_1(w,x,y)$ for any w and at all feasible pairs (x,y). There is however, no need to define $W_1(w,x,y)$ for the case where the surplus is negative $(S(x,y) < 0)$: if a productivity shocks moves y to y' such that $S(x,y') < 0$ then the worker and the firm separate irrespective of the wage.

The value of working at wage rate w in a feasible match (x, y) can be written as:

$$
\begin{aligned}
(r+\delta+\xi+s_1\kappa v\left(A\left(w,x,y\right)\right))\left[W_1\left(w,x,y\right)-W_0\left(x\right)\right] &= w-rW_0\left(x\right)\\
&+\delta\int\left[\min\left\{S\left(x,y'\right),W_1\left(w,x,y'\right)-W_0(x)\right\}\right]^+ dy'\\
+s_1\kappa\int_{A(w,x,y)}\left[\beta\max\left\{S\left(x,y\right),S\left(x,y'\right)\right\}+(1-\beta)\min\left\{S\left(x,y\right),S\left(x,y'\right)\right\}\right] & v\left(y'\right)dy', \quad (11)
\end{aligned}
$$

where

$$
A\left(w,x,y\right)=\left\{y':W_1\left(w,x,y\right)-W_0\left(x\right)<S\left(x,y'\right)\right\},
$$

is the set of jobs that can lead to a wage change (either by moving or renegotiation).

In (11) the effective discount rate is the interest rate, increased by the rate of arrival of shocks δ, the job destruction rate ξ and the rate of arrival of offers that can lead to a wage change and/or job change. On the right hand side the flows are the wage net of the flow value of unemployment; the expected value to the worker of a productivity shock (times the probability that it occurs): in this case the worker either ends up with the entire new surplus or the new value or indeed nothing if the match is no longer feasible; the third term is the expected value following an outside offer. The integral is over all offers that can improve the value (whether the worker moves or not).

2.5 Steady-state flow equations

In a stationary equilibrium the number of jobs N, the distribution of matches as well as the distribution of x among the unemployed and of y among the posted vacancies are all constant over time. Thus to solve for the equilibrium distribution we need to define the steady state flow equations.

The total number of matches in the economy will be

$$
L-U=N-V=\int h\left(x,y\right)dx\,dy. \quad (12)
$$

Existing matches, characterized by the pair (x, y), can be destroyed for a number of reasons. First, there is exogenous job destruction at a rate ξ; second, with probability δ, the job component of match productivity changes to some value y' different from y, and the worker may move to unemployment or may keep the job; third, the worker may change job, with probability $s_1 \kappa v \left(\overline{B}(x,y)\right)$ - i.e., a job offer has to be made and has to be acceptable for x currently matched with y. On the inflow side, new (x,y) matches are formed when some unemployed or employed workers of type x match with vacant jobs y, or when (x,y') matches are hit with a productivity shock and exogenously change from (x,y') to (x,y). In a steady state all these must balance leaving the match distribution unchanged. Thus formally we have for all (x,y) such that the match is feasible , i.e. $S(x,y) \geq 0$:

$$
\begin{aligned}
\left[\delta + \xi + s_1 \kappa v \left(\overline{B}(x,y)\right)\right] h(x,y) = \delta \int h\left(x,y'\right) dy' \\
+ \left[s_0 u(x) + s_1 h(x, B(x,y))\right] \kappa v(y) . \quad (13)
\end{aligned}
$$

The left hand side of (13) describes the total destruction rate of (x,y) matches. The right hand side has two components: the first accounts for existing matches (x,y') that receive a shock and become y. The second accounts for those among the unemployed and those among the employed at matches inferior to (x,y) who receive acceptable offers for an (x,y) match, i.e. the inflow into (x,y) matches due to job mobility. This equation defines the steady-state equilibrium, together with the following accounting equations for the workers:

$$
\begin{aligned}
u(x) &= L - \int h(x,y) \, dy \\
&= L \frac{\xi + \delta \# \left(\overline{M}(x)\right)}{\xi + \delta \# \left(\overline{M}(x)\right) + s_0 \kappa v (M(x))}
\end{aligned}
\quad (14)
$$

where $\delta \# \left(\overline{M}(x)\right) = \delta \int \mathbf{1}\{S(x,y) < 0\} \, dy$ is the endogenous job destruction rate and $s_0 \kappa v (M(x))$ is the job finding rate, and for the jobs

$$
v(y) = N - \int h(x,y) \, dx. \quad (15)
$$

The total number of vacancies V is then obtained as

$$
V = \int v(y) dy. \quad (16)
$$

2.6 Equilibrium

In equilibrium all agents follow their optimal strategy and the steady state flow equations defined above hold. The exogenous parameters of the model are the number of workers L, the form of the matching function $M(\cdot,\cdot)$ as well as the arrival rate of shocks δ, the job destruction rate ξ, the search intensities for the unemployed s_0 and employed workers s_1, the discount rate r, the value of leisure $b\,(x)$, the cost of posting a vacancy c, bargaining power β, and the production function $f\,(x,y)$. The distribution of worker types, of productivities and of the shocks have all been normalized to uniform $[0,1]$, without loss of generality. The web appendix provides a simple iterative algorithm that uses these equations and that of the surplus to compute the equilibrium objects.

2.7 Minimum wages

Incorporating a minimum wage puts a constraint on the ability of workers and jobs to make transfers. As a result, the condition for match feasibility will depend on the match surplus being high enough to cover a minimum wage contract $\underline{w}$, and still provide positive surplus to the job. Given this constraint, an (x,y) match is feasible if and only if $S\,(x,y)>0$ and $W_1\,(\underline{w},x,y)-W_0\,(y)\leq S\,(x,y)$. This requires defining the value to a filled job of an (x,y) match paying a wage $\underline{w}$. The key practical difficulty is that the matching set for the unemployed depends both on whether the surplus is positive and on the value of filling the vacancy at the minimum wage. Before, $S(x,y)>0$ was sufficient to determine the feasible matches. For all matches that are feasible subject to the minimum wage constraint, the wage is determined as

$$w=\max\left\{\underline{w},w^{*}\right\},$$

where w^{*} solves (11) with either (4) or (5) on the left hand side depending on whether the worker is hired away from another job or hired from unemployment.

3 The Data

We use the 1979 to 2002 waves of the National Longitudinal Survey of Youth 1979 (NLSY) to construct all moments used in estimation. The NLSY consists of 12,686 individuals who were

14 to 21 years of age as of January, 1979. The NLSY contains a core nationally representative random sample, as well as an over-sample of black, Hispanic, the military, and poor white individuals. For our analysis, we keep only white males from the core sample. Since the schooling decision is exogenous to the model, we only include data for individuals once they have completed their education. We also drop individuals who have served in the military, or have identified their labor force status as out of the labor force. The majority of individuals who are not in the labor force report being disabled and are clearly not searching for employment. We subdivide the data into three education groups: less than high school education, high school degree or some college, and college graduate.

Individuals are interviewed once a year and provide retrospective information on their labour market transitions and their earnings. From this it is possible to construct a detailed weekly labour market history. However, we choose to work at at monthly frequency aggregating the data as follows: we define a worker as employed in a given week if he worked more than 35 hours in the week. We define a worker's employment status in a month as the the activity he was engaged in for the majority of the month. In other words, we treat unemployment spells of two weeks or less as job-to-job transitions. After sample selection, we are left with an unbalanced panel of 2,022 individuals (325,462 person months).

We remove aggregate growth from wages in the NLSY , using estimates from the CPS data 1978-2008.[5] We select white men in the CPS and regress the log wage on a full set of year dummies and age dummies, separately for our four education groups. To remove aggregate growth from the NLSY79 data, we subtract the coefficient on the corresponding year dummy from log wage. Our definition of wage in the NLSY is weekly earnings. We trim the data to remove a very small number of outliers when calculating wage changes. When calculating year to year wage changes, we exclude observations where the wage changes by more than a factor of 10. For year to year wage growth at the same job, we exclude observations where the wage changes by more than a factor of two. Similarly for wage changes via unemployment, we exclude wages that either increase by more than a factor two or decrease by more than a factor than ten. Finally, for job-to-job wage growth, we exclude observations where the wage increases by a factor greater than ten or decreases by a factor greater than two. This results in variances of wage changes that are broadly consistent with Meghir and Pistaferri (2004). Finally when

[5] The trend cannot be estimated from the NLSY, which is a cohort and has an aging structure.

calculating transition rates we average over the entire time period, thus averaging out as much as possible the cyclical fluctuations.

4 Specification

Estimating the model involves first parameterizing the production and the matching functions. For the production function we have chosen a CES specification that allows us to estimate the degree of sorting in the data. Thus we have

$$f(x,y) = \exp(f_1)\left(\exp\left(f_2\Phi^{-1}(x)\right)^{f_4} + \exp\left(f_3\Phi^{-1}(y)\right)^{f_4}\right)^{1/f_4},$$

where Φ is the standard normal cdf, f_2 is the standard deviation of x and f_3 is the standard deviation of productivity y. Thus we take the uniformly distributed characteristics and transform them to log normally distributed variables with variances to be estimated. The parameter f_1 determines the scale of wages. Finally the parameter f_4 determines the degree of substitutability between x and y with $\rho_{xy} = 1/(1-f_4)$ being the elasticity of substitution. When $f_4 = 1$ the production function becomes additive and there are no complimentarities between the inputs.

For the matching function we specify

$$M = \alpha\left[U + s_1(L-U)\right]^{0.5} V^{0.5}$$

where the search intensity for the unemployed has been normalized to one. The matching function allows the number of matches to adapt to the number of vacancies and the unemployed and as such is important for evaluating counterfactual policies.

5 Estimation Method

Constructing the likelihood function for this model is intractable. So to estimate the model we use Simulated Method of Moments (SMM). Because the simulated moments are not necessarily a smooth function of the parameters we use an MCMC based method for classical estimation developed by Chernozhukov and Hong (2003), which does not require derivatives of the criterion function.

The SMM approach works as follows. The data sample allows to calculate a vector of moments $\widehat{m}_N = \frac{1}{N}\sum_{i=1}^{N} m_i$, for example, mean wage for a given experience or the probability of being employed, or wage change after t periods for job movers, etc. Given a value θ for the vector of structural parameters one can simulate S alternative samples of wage trajectories yielding average simulated moments $\widehat{m}_S^M(\theta) = \frac{1}{S}\sum_{s=1}^{S} m_s^M(\theta)$, where the superscript M denotes a model generated quantity. The SMM aims at finding a value $\widehat{\theta}$ maximising

$$L_N(\theta) = -\frac{1}{2}\left(\widehat{m}_N - \widehat{m}_S^M(\theta)\right)^{\mathsf{T}} \Omega_{NS}(\theta)^{-1}\left(\widehat{m}_N - \widehat{m}_S^M(\theta)\right),$$

where $\Omega_N(\theta)$ is an estimator of the variance of $\widehat{m}_N - \widehat{m}_S^M(\theta)$. Since the estimates and the simulations are (assymptotically) independent we decompose the weight matrix as

$$\Omega_{NS}(\theta) = \Sigma_N + W_S(\theta).$$

We can estimate the variance of $\widehat{m}_S^M(\theta)$ ($W_S(\theta)$) as $\frac{1}{S}$ times the empirical variance of $m_i^s(\theta)$, that is,

$$W_S(\theta) = \frac{1}{S}\left(\frac{1}{S}\sum_{s=1}^{S}[m_s^M(\theta) - \widehat{m}_S^M(\theta)][m_s^M(\theta) - \widehat{m}_S^M(\theta)]^{\mathsf{T}}\right).$$

For the variance of m_i (Σ_N) in the actual sample, we use the bootstrap. However as shown by Altonji and Segal (CITE) estimating the variance of second moments is subject to substantial small sample bias causing minimum distance type estimators to be seriously biased. Thus, instead of using the variance of the bootstrap draws, which are sensitive to outliers, we base our calculation on the interquartile range, by using the fact that the averages are asymptotially normal.[6] Thus denote by $(m_i^b, i = 1, ..., N), b = 1, ..., B$, the B bootstrap samples obstained by resampling with replacement from $(m_i, i = 1, ..., N)$. Denote by $\widehat{m}_{kN}^b$ the average of the kth element of m_i^b. Using the fact that $\widehat{m}_{kN}^b$ is approximately normal we can estimate the assymptotic standard deviation by rescaling appropriately the interquartile range, i.e.

$$\widehat{\sigma}_{kN} = IQR(\widehat{m}_{kN}^b)/\left(\Phi^{-1}(3/4) - \Phi^{-1}(1/4)\right),$$

where Φ is the CDF of the standard normal distribution.[7] The same idea can be applied to

[6]We thank Han Hong for suggesting this procedure.

[7]So $\Phi^{-1}(3/4) = 0.6745$ and $\Phi^{-1}(1/4) = -0.6745$.

the all the relevant statistics, allowing estimation of the entire covariance matrix of the data moments (Σ_N), while avoiding the problems raised by Altonji and Segal.

The estimation procedure of Chernozhukov and Hong consists of simulating a chain of parameters that (once converged) has the quasi-posterior density

$$p(\theta) = \frac{\mathrm{e}^{-L_N(\theta)}\pi(\theta)}{\int \mathrm{e}^{-L_N(\theta)}\pi(\theta)\,d\theta}.$$

A point estimate for the parameters is obtained as the average of the N_S elements of the converged MCMC chain:

$$\hat{\theta}_{MCMC} = \frac{1}{N_S}\sum_{j=1}^{N_S}\theta^j,$$

and standard errors are computed as the standard deviation of the sequence of θ^j. To simulate a chain that converges to the quasi posterior, we use the Metropolis-Hastings algorithm. The algorithm generates a chain $\left(\theta^0, \theta^1, ..., \theta^{N_S}\right)$ as follows. First, choose a starting value θ^0. Next, generate ψ from a proposal density $q\left(\psi|\theta^j\right)$ and update θ^{j+1} from θ^j for $j = 1, 2, ...$ using

$$\theta^{j+1} = \begin{cases} \psi & \text{with probability} \quad d\left(\theta^j, \psi\right) \\ \theta^j & \text{with probability} \quad 1 - d\left(\theta^j, \psi\right), \end{cases}$$

where

$$d(\gamma, \zeta) = \min\left(\frac{\mathrm{e}^{L_N(\zeta)}\pi(\zeta)\,q(\gamma|\zeta)}{\mathrm{e}^{L_N(\gamma)}\pi(\gamma)\,q(\zeta|\gamma)}, 1\right).$$

This procedure is repeated many times to obtain a chain of length N_S that represents the ergodic distribution of θ. Choosing the prior $\pi(\theta)$ to be uniform and the proposal density to be a random walk ($q(\gamma|\zeta) = q(\zeta|\gamma)$), results in the simple rule

$$d(\gamma, \zeta) = \min\left(\mathrm{e}^{L_N(\zeta) - L_N(\gamma)}, 1\right).$$

The main advantage of this estimation strategy is that it only requires function evaluations, and thus discontinuous jumps do not cause the same problems that would occur with a gradient based extremum estimator. Additionally, the converged chain provides a direct way to construct valid confidence intervals or standard errors for the parameter estimates. The drawback of the procedure is that it requires a very long chain, and consequently a very large number of function evaluations, each requiring the model to be solved and simulated. In practice, we simulate 100 chains in parallel, each of length 10,000, and use the last 2,000 elements (pooled over the 100 chains) to obtain parameter estimates and standard errors.[8]

In practice, we start the chain with a weighting matrix $\Omega_N(\theta)$ that consists in only the diagonal of Σ_N. When the chain seems to stabilise, we calculate a first estimate $\widetilde{\theta}_1$ by averaging and we then set the weighting matrix equal to $\Omega_N(\widetilde{\theta}_1)$. Then, we let the chain unfold until it stabilises again. We thus obtain a second estimate $\widetilde{\theta}_2$, which is like a two-stage feasible GMM estimator. Finally, we let $\Omega_N(\theta)$ update at each new iteration to obtain the MCMC equivalent of the continuously updated GMM estimator.

5.1 Measurement error

Wages are likely to be measured with error. Indeed if we do not take this into account we will distort both the variance of the shocks to productivity and the transitions, both of which drive the variance of wage growth. We use the monthly records of wages in the NLSY. Thus, while it may be reasonable to assume that measurement error is independent from one year to the next it may not be so within the year, as all records are reported at the same interview with recall. Having experimented with a number of alternatives, including a common equicorrelated component across all months, we settled on a measurement error structure that is AR(1) within year and independent across years. The serial correlation coefficient together with the variance of measurement error is estimated alongside the other parameters of the model.

6 Identification

The moments we choose are the transitions between employment states, the level of wages and their cross sectional variance one year, eleven years and twenty-one years following labor market entry (discussed below), wage growth and its variance within and between jobs ,with and without an intervening unemployment spell, as well as the variance and first two autocovariances of wage residual wage growth, once we remove age and time effects; the latter mimic the autocovariance used in papers that consider the dynamics of earnings.

[8] Details pertaining to tuning the MCMC algorithm, a parallel implementation, and related methods in statistics can be found in Robert and Casella (2004); Vrugt, ter Braak, Diks, Higdon, Robinson, and Hyman (2009); Sisson and Fan (2011).

Although it is hard to argue that some moments are informative only for some parameters, we can provide a bit of information on how the parameters are identified. The transitions in and out of work and between jobs play a key role in identifying the job destruction rate (ξ), the intensity of on-the job search (s_1) and the overall level of matches (α). In particular α is identified by h_{UE}, s_1 is identified by $h_{\Delta J}$ (relative to h_{UE}) and that ξ is identified by h_{EU}. The wage level is directly related to the scale of output and hence the average wage identifies $f_{1.}$ The variance of wages is directly linked to the variance of unobserved heterogeneity (or equivalently its weight in the production function) and hence both f_2 and f_3 are directly linked to the variance of wages. Through the effect of productivity shocks or changes in matches from job to job, y is also linked to the variance of wage growth and hence the identification of f_3 is also driven by the variance of wage growth, whether unconditionally or conditional on all types of transition. Similarly the within and between job variance of wage growth identifies the rate of arrival of productivity shocks δ. The bargaining parameter β is primarily identified by wages following a job change. So the key moment for this is wage growth with or without an intervening unemployment spell. Finally the measurement error variance (σ^2) is primarily identified by the cross sectional variance of wages and variance of wage growth, while the autocovariance of wages identifies the within year serial correlation of wages ρ. Finally, in the current version of the model the parameter b in $b(x) = f(x, b)$, reflecting the income flow while out of work, is identified by starting wages in the new jobs after unemployment.

This schematic description only provides some basic intuition. In practice the model is heavily over-identified and most parameters affect the fit of the model in many dimensions. Figure 1 shows how the pseudo-likelihood $L(\theta)$ and a few moments change when one varies the main parameters of interest, f_4, δ, β and b. Of course, this is not a proof, yet a strong indication that the moments are informative. We now discuss the resulting fit.

7 The Fit of the Model

Table 1 compares simulated and observed transition rates and wage growth. We also report the standard error of the data moment to give a sense of how significant the difference between the two is and to show the weight that the moment has in estimation - the lower the standard error the larger the weight. The frequency of the transitions is monthly and within firm wage growth

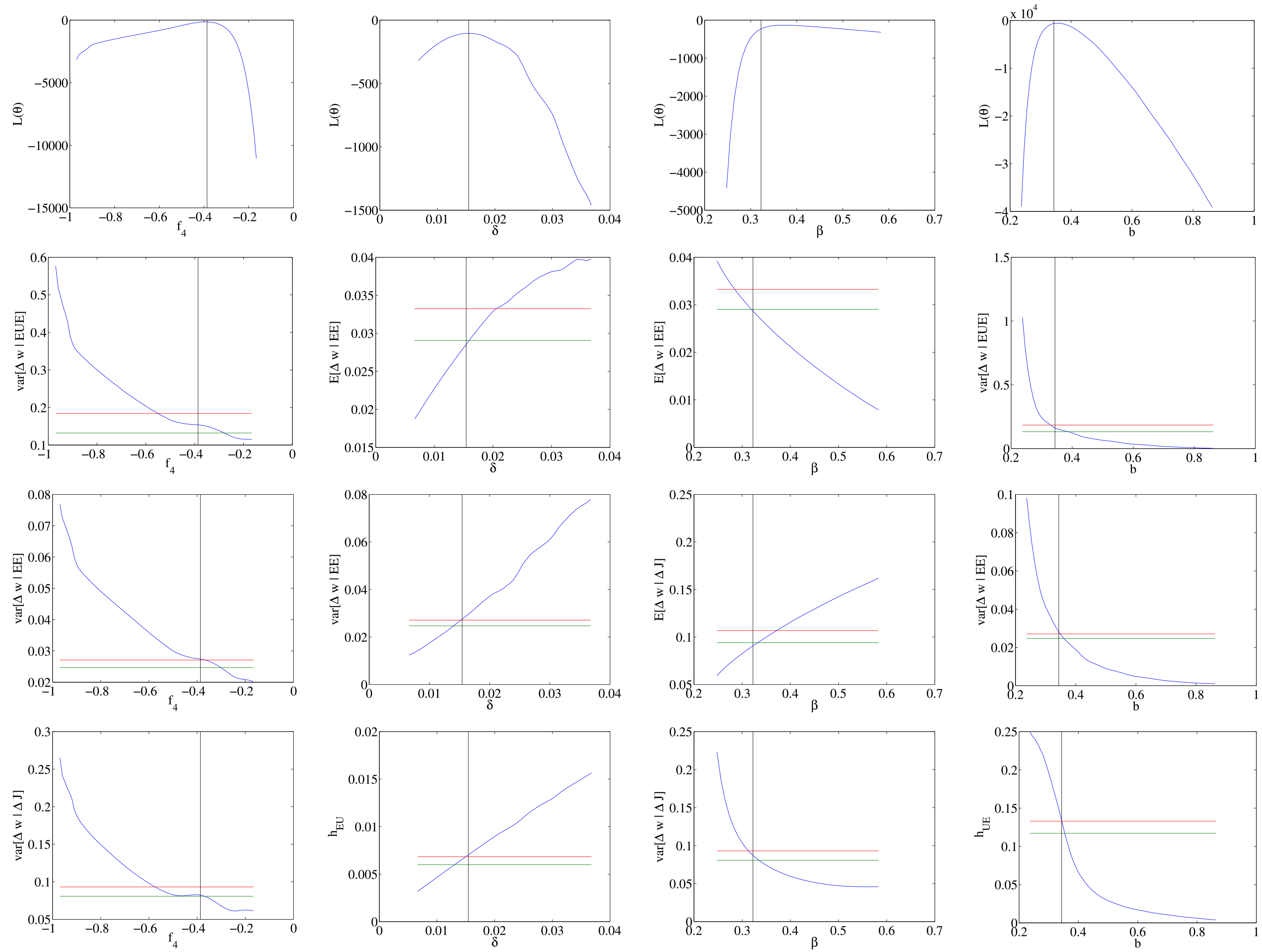


Figure 1: Identification

Table 1: Model Fit

	Less than high school			High school graduate			College graduate		
	Model	Data	St err	Model	Data	St err	Model	Data	St err
h_{EU}	0.030	0.028	0.002	0.017	0.017	0.000	0.007	0.006	0.000
h_{UE}	0.117	0.114	0.007	0.130	0.124	0.008	0.136	0.125	0.008
$h_{\Delta J}$	0.062	0.066	0.003	0.035	0.043	0.001	0.022	0.023	0.001
$\mathbb{E}\left(\Delta w \| EUE\right)$	-0.184	-0.104	0.020	-0.276	-0.050	0.028	-0.378	-0.074	0.029
$\mathbb{E}\left(\Delta w \| EE\right)$	0.022	0.032	0.005	0.036	0.031	0.002	0.029	0.031	0.002
$\mathbb{E}\left(\Delta w \| \Delta J\right)$	0.058	0.060	0.006	0.070	0.064	0.007	0.091	0.100	0.006
$\mathbb{E}\left(w_1\right)$	5.915	5.646	0.040	6.241	6.063	0.021	6.625	6.565	0.020
$\mathbb{E}\left(w_{11}\right)$	5.981	6.053	0.038	6.386	6.504	0.030	6.949	7.063	0.028
$\mathbb{E}\left(w_{21}\right)$	5.976	6.319	0.050	6.387	6.664	0.095	6.963	7.173	0.098

Note: The first three rows are the monthly hazard rates for employment to unemployment (h_{EU}), unemployment to employment (h_{UE}), and job-to-job ($h_{\Delta J}$). The mean wage changes in the next three rows are conditional on employment-unemployment-employment (EUE), one year continuous employment at the same job (EE), and job-to-job transitions (ΔJ). The last three rows report average log wages one, 11 and 21 years after school from the data. These are compared to average wages 1, 11 and 21 years after a spell of unemployment from the model.

has been annualized. All results are presented for the three educational groups.

The predicted transition rates fit remarkably well, with the actual numbers being very close to each other and with no significant differences when it comes to exits from employment to unemployment h_{EU} and re-entry from unemployment h_{UE}. There seem to be some significant divergence between predicted and actual transitions from job to job with an intervening spell of unemployment, particularly for the lowest skill group of the high school dropouts. However, even there the divergence is only small.

The model does a good job at replicating the qualitative pattern of wage growth with a cost of job loss, a wage gain for those moving jobs and a relatively low but positive wage growth for those remaining with the same employer. However, the model does overestimate the cost of job loss. The fit of this moment is in part driven by the share of the surplus that goes to the worker (β). However, changing this value will worsen the fit of the between job wage growth, which has a lower standard error. This fits remarkably well implying that the model replicates well the wage gains to job mobility. Finally, the wage growth within job ($\overline{\Delta w}_{|EE}$) and between job with no unemployment spell in between ($\overline{\Delta w}_{|\Delta J}$). The former is somewhat underestimated by the model but the latter fits exceptionally well. Thus the gains to wages from job mobility are replicated by the model very closely. The within job wage growth is however underestimated, which perhaps points to the need to introduce human capital considerations in the model, a subject of future research (also see Bagger, Fontaine, Postel-Vinay, and Robin, 2011).

The overall implications of the transition rates and wage growth can be summarized at least in part by the comparing the wage profiles we obtain following an unemployment spell. Because human capital does not grow in our model the implied wage profile can be compared to the life-cycle profile of wages as observed following the end of formal education. In the last three rows of Table 1 we show average observed wages for those one eleven and 21 years into a labor market career. These are compared to average wages one, eleven and twenty one years following an unemployment spell. While wages predicted by the model do exhibit some growth, particularly for the highest education group it is not enough to match that of the data. In principle our model could match steeper growth by changing the distribution of firm productivities (reflected in the production function parameters), increasing the arrival rate of outside offers and reducing the exits to unemployment; however this would not allow us to fit the transitions as well as we do. It seems that search and matching in itself captures only part of the wage growth we observe, despite the strong matching effects we will document below for all but the lowest education group.

In Table 2 we present the fit of various second moments. The model moments include the effects of measurement error so that they are consistent with the actual data. The first three rows relate to the variance of wage growth, followed by the cross sectional variance of wages. These are followed by the variance/autocovariance structure of residual earnings, net of wage growth.

The model fits remarkably well the within firm and between firm variance of wage growth for all education groups. However it underestimates the variance of wage growth around an unemployment spell; this feature is much improved among those with the highest education level, where the difference of the model moment from that of the data is not significant. It seems that in our model a large loss and a large disparity of wages following a spell in unemployment cannot be made consistent with the needs of fitting transitions and indeed wage movements that do not involve unemployment changes. One interpretation is that returns to tenure are important. However, most studies seem to suggest that these are either zero or very small[9] Another possibility may relate to depreciation of skills and duration of unemployment, both of which are not accounted for here.

[9]see Altonji and Williams (1997) for the US.

Table 2: Model Fit

	Less than high school			High school graduate			College graduate		
	Model	Data	St err	Model	Data	St err	Model	Data	St err
var $(\Delta w \vert EUE)$	0.160	0.166	0.022	0.092	0.121	0.028	0.154	0.158	0.026
var $(\Delta w \vert EE)$	0.049	0.040	0.003	0.036	0.033	0.001	0.027	0.026	0.001
var $(\Delta w \vert \Delta J)$	0.097	0.083	0.008	0.050	0.073	0.006	0.083	0.087	0.006
var (w_1)	0.235	0.349	0.089	0.237	0.278	0.053	0.193	0.321	0.053
var (w_{11})	0.293	0.283	0.065	0.341	0.321	0.029	0.390	0.357	0.028
var (w_{21})	0.292	0.314	0.057	0.347	0.324	0.081	0.414	0.400	0.086
var (Δu_t)	0.092	0.101	0.008	0.048	0.081	0.005	0.043	0.071	0.005
cov $(\Delta u_t, \Delta u_{t-1})$	-0.008	-0.013	0.003	-0.007	-0.010	0.002	0.002	-0.003	0.002
cov $(\Delta u_t, \Delta u_{t-2})$	-0.018	-0.018	0.003	-0.008	-0.013	0.002	-0.007	-0.012	0.002

Note: The first three rows represent the variance of wage growth between jobs (with intervening unemployment), within jobs (annualized) and between jobs with no intervening unemployment respectively. The next three rows is the cross section variance of wages one, eleven and twenty one years following the start of the career (following an unemployment spell in the model). The last three rows represent variances and autocovariances of residual wage growth, having removed "deterministic" growth and refers to annual frequencies

The following three rows consider the fit of cross sectional variances of wages at the three different points in the "life-cycle" discussed above. In all but the lowest skill groups the initial cross sectional variance of wages is explained well by the model. The model also captures the general pattern of decline in the variance for the lowest skill group and an increase everywhere else. However, other than for the college group the magnitude of the variance growth is not captured well pointing to the possibility of heterogeneous effects of experience or permanent shocks to individual human capital x.

One of the issues we set out to investigate is whether our model can justify and explain the dynamics of earnings we see in reduced form analyses such as those of MaCurdy (1982); Abowd and Card (1989); Meghir and Pistaferri (2004). This question has also been looked at in a simpler model by Postel-Vinay and Turon (2010) who find that their structural search model is able to replicate the observed dynamics. In our case the variance of wage growth is underestimated. Beyond that, the first and second order autocovariances are negative and small as in the data. Based on these estimates one would conclude that wages are a random walk plus a transitory shock with quite low variance.

8 Parameter Estimates

The parameters implied by our estimation are presented in Table 3. Turning first to the production function we see that for the least skilled workers $f_4 = 1$, implying an additive production function and no gains from sorting in that labor market. Interestingly, because of search frictions there are still gains from on-the-job search: if an employed unskilled worker (who would be accepted by any firm) meets a higher productivity firm she will be offered a higher wage to move because the new firm will prefer (to an extent) an expensive unskilled worker to remaining vacant longer; of course this is subject to the premium they need to pay and to the expected waiting time to finding either an unemployed unskilled worker or one working at a lower value job.

For the higher skill groups the substitution elasticities are (in ascending order of skill) ## and 0.56 respectively: in all three cases there are strong complementarities and gains from sorting; we look at the implications below.

The exogenous monthly job destruction rate (ξ) is quite high for the unskilled, but low for the other three groups: since sorting plays no role for the unskilled the only way they would enter unemployment is either if the firm decides to close down and remain idle (following a big negative productivity shock) or if the job is destroyed exogenously. Effectively the absence of sorting has all but closed down the endogenous job destruction implied by productivity shocks that turn previously viable matches to ones with a negative surplus; while this is an important mechanism for job termination for the three top groups is plays a much reduced role for the unskilled because any viable job hires them.

Productivity shocks are drawn from a uniform distribution and arrive in any period with probability δ. The probability of such shocks is lowest for the lowest and highest skill groups but highest for the middle group. Thus for example productivity in college graduate jobs are the most persistent with a change occurring approximately every 32 months. The other key parameter for our model is the bargaining parameter β, with the results implying that workers get between 2.7% and 22% of the surplus. While nonzero, it is interesting to see how small this share is even for college graduates. Finally, search intensity (s_1) is generally lower for the employed than for the unemployed with the exception of unskilled workers.

The last two parameters relate to measurement error. The standard deviation of monthly

Table 3: Parameter Estimates

		Less than HS	High school	College
Production function parameters	f_1	4.27	9.98	9.99
		(0.09)	(0.04)	(0.14)
	f_2	1.31	1.10	1.14
		(0.08)	(0.02)	(0.03)
	f_3	5.05	5.17	6.42
		(0.07)	(0.06)	(0.17)
	f_4	0.82	-0.23	-0.39
		(0.07)	(0.01)	(0.04)
Matching function parameters	α	1.49	0.86	0.75
		(0.06)	(0.07)	(0.06)
	s_1	2.90	1.11	0.92
		(0.22)	(0.07)	(0.090)
Probability of exogenous job destruction	ξ	0.024	0.0003	0.0011
		(0.001)	(0.0002)	(0.0004)
Probability of shock to y	δ	0.028	0.033	0.015
		(0.003)	(0.001)	(0.001)
Home production parameter	b	0.60	0.39	0.34
		(0.01)	(0.01)	(0.003)
Worker bargaining power	β	0.12	0.27	0.32
		(0.011)	(0.01)	(0.02)
Variance of measurement error	σ	0.042	0.070	0.020
		(0.014)	(0.006)	(0.014)
	ρ	0.96	0.94	0.45
		(0.56)	(0.09)	(0.27)
Probability of shock to y		0.028	0.033	0.015
Mean separation rate conditional on shock		0.220	0.489	0.454
Variance of separation rate conditional on shock		0.034	0.024	0.026
Probability of job contact when unemployed		0.118	0.174	0.155
Mean matching rate given a contact		0.990	0.739	0.876
Variance of matching rate given a contact		0.003	0.089	0.015
Probability of job contact when employed		0.342	0.194	0.142
Mean matching rate given a contact		0.177	0.169	0.130
Variance of matching rate given a contact		0.048	0.034	0.024
Elasticity of substitution: $(1-f_4)^{-1}$		5.56	0.81	0.72

Table 4: Implications for wage dynamics

	Less than high school		High school graduate		College graduate	
Share of wage changes						
shocks	0.195		0.244		0.250	
counter offers	0.805		0.756		0.750	
Autocovariance (autocorrelation) of order s						
s	*cov*	*corr*	*cov*	*corr*	*cov*	*corr*
0	0.102	1.000	0.053	1.000	0.045	1.000
1	-0.008	-0.075	-0.007	-0.141	0.002	0.038
2	-0.019	-0.190	-0.009	-0.165	-0.007	-0.146
3	-0.011	-0.105	-0.005	-0.088	-0.005	-0.112
4	-0.006	-0.059	-0.002	-0.047	-0.004	-0.082
5	-0.003	-0.031	-0.001	-0.024	-0.003	-0.057
6	-0.002	-0.014	-0.001	-0.017	-0.002	-0.040
7	-0.001	-0.008	0.000	-0.008	-0.001	-0.030
8	-0.001	-0.007	0.000	-0.004	-0.001	-0.021
9	0.000	-0.004	0.000	-0.002	-0.001	-0.012
10	0.000	-0.003	0.000	-0.002	0.000	-0.010

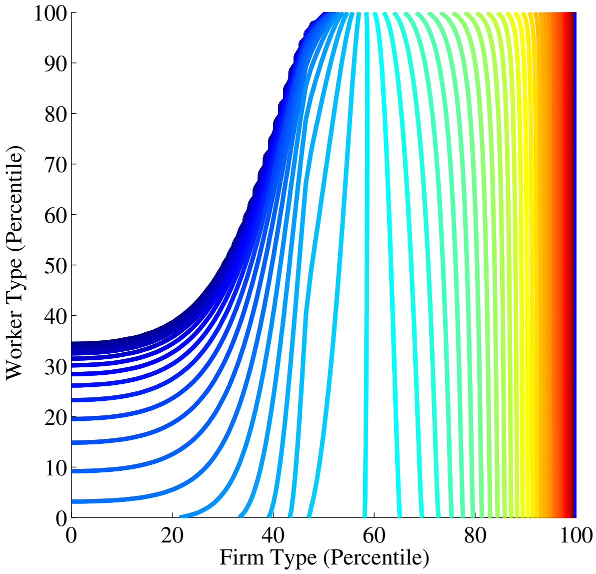
Note: The autocovariances are based on simulating from the stationary distribution,and, as a result, differ slightly from those presented in Table 2.

measurement error is below 0.1 and as expected the correlation between any two months within a reporting year is reasonably high.

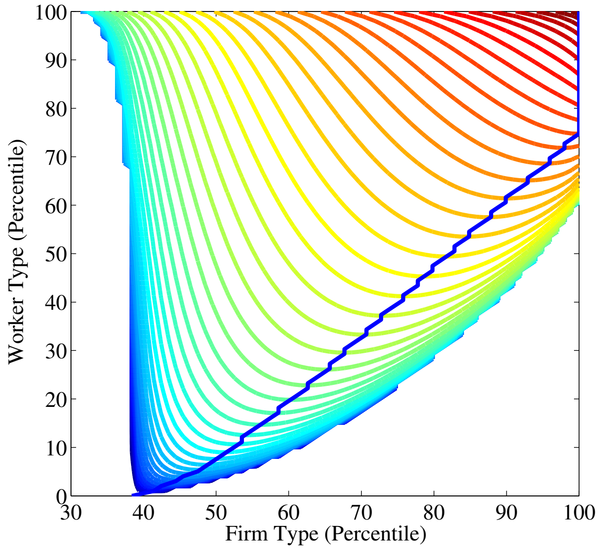
9 Sorting in the labour market

The key conclusion of the empirical results is that there are productive complementarities between job and worker characteristics for but the lowest skill workers. We thus expect sorting for the higher skill groups. In Figure 5a we illustrate these implications by presenting information on the joint distribution of matches. For each skill group, the graphs on the left represent the cumulative distribution for firm type conditional on worker type; on the right we show the cumulative distribution of worker type conditional on various quantiles of the distribution of firm productivity.

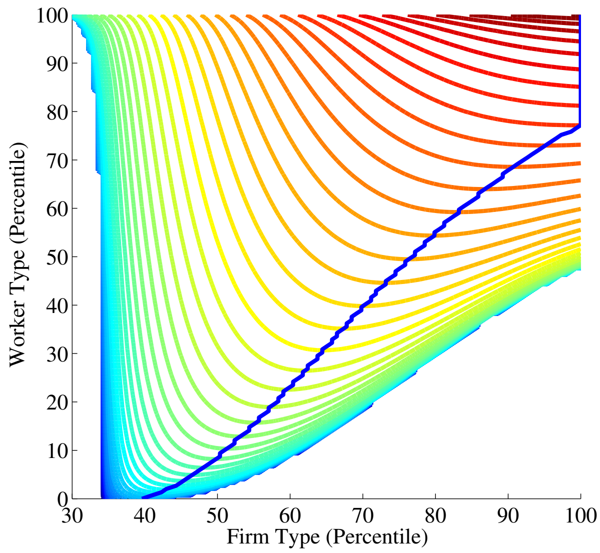
There is very strong evidence of sorting: as we move from left to right on these graphs we move to higher quantiles of the conditioning characteristic (worker or firm respectively): the distributions of firm productivities relating to higher levels of worker productivity stochastically dominate the ones conditional on lower worker productivity. The same is true if we look at



(a) Less than high school



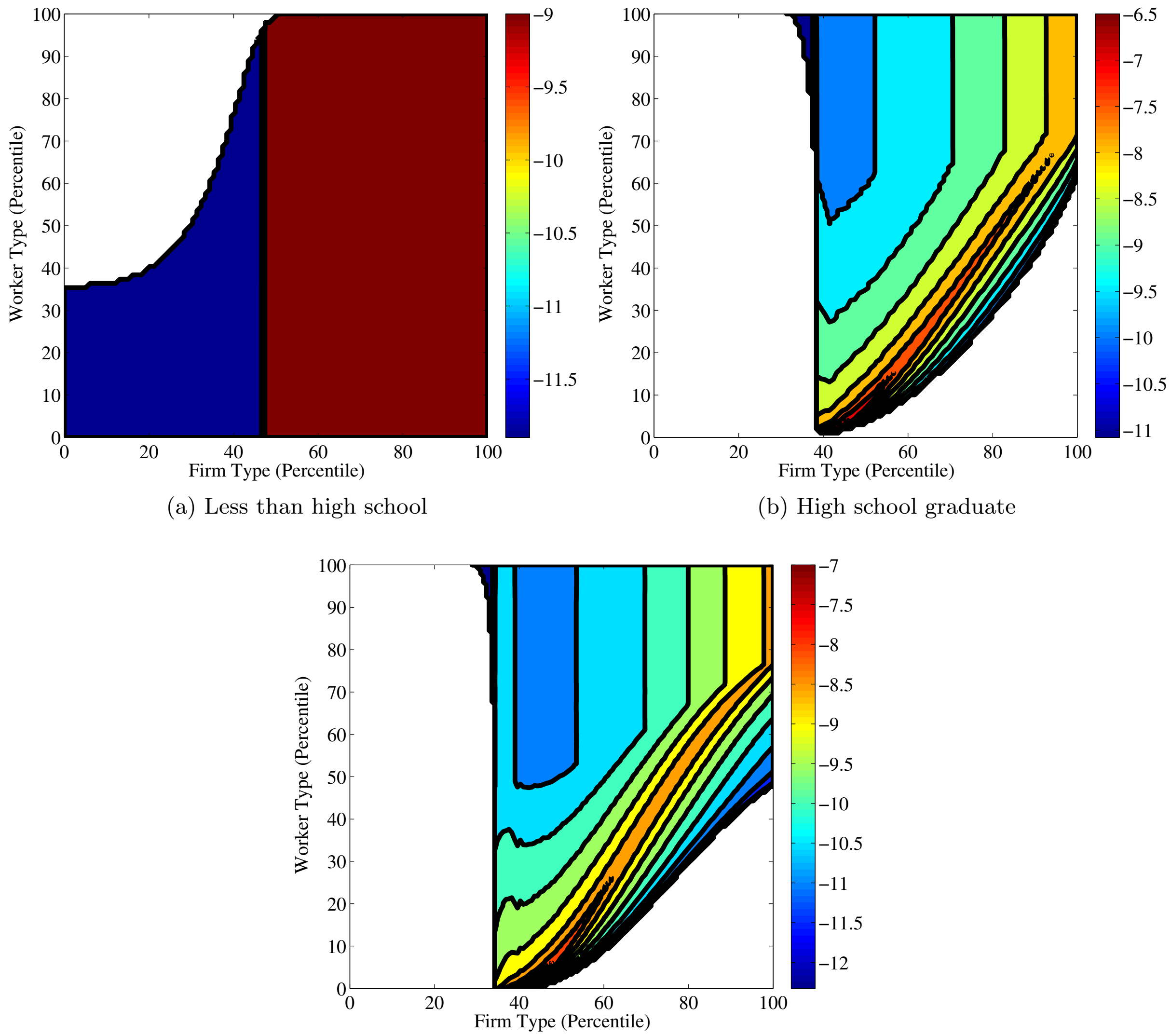
(b) High school graduate



(c) College graduate

Figure 2: Match surplus

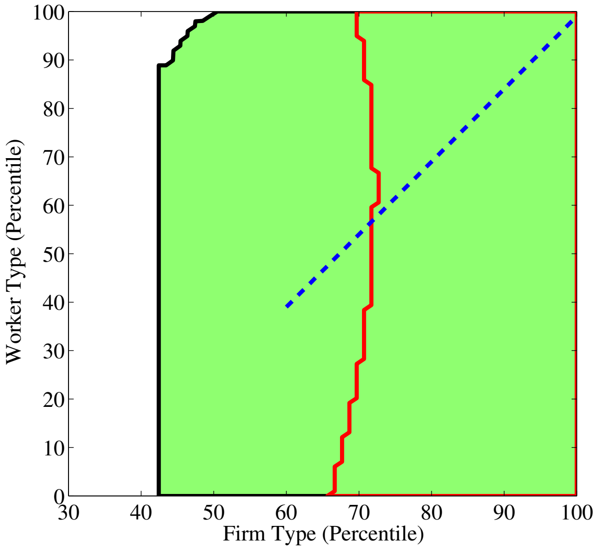
Notes: Here we plot the contours of the match surplus, $S(x, y)$. The blue line passing through the contours represents, for each worker type x, $y(x) = \arg\max S(x, y)$.



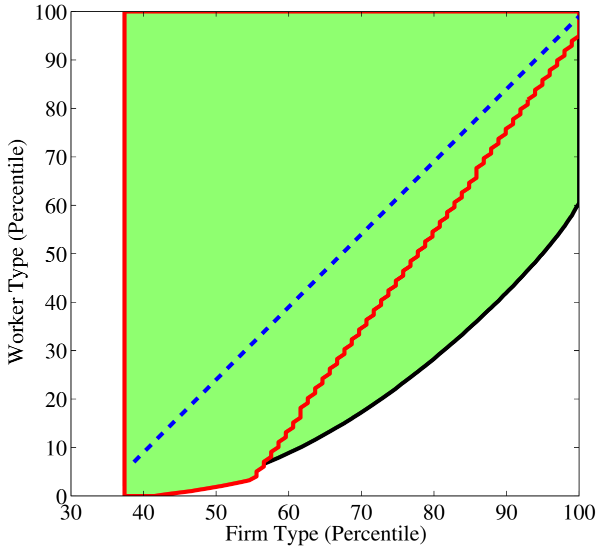
(c) College graduate

Figure 3: Stationary distribution of worker firm matches

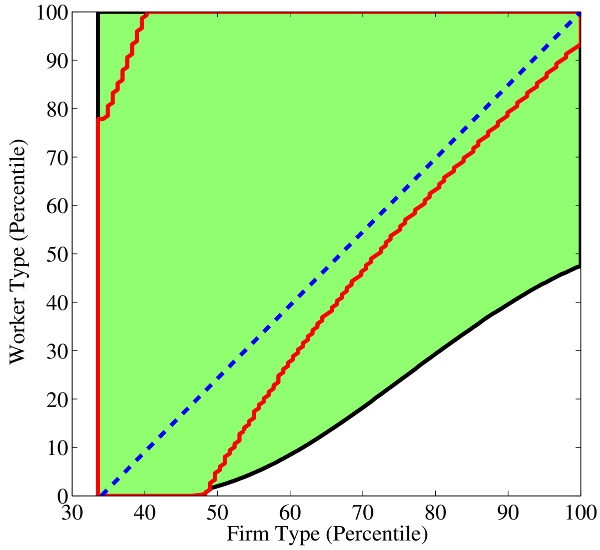
Notes: Here we plot the log-density of (x, y) matches. The positive density of matches to the left of the vacancy posting cut-off are the result of shocks to y. For these matches, $S(x, y) \geq 0$, and therefore remain intact after the shock, however, due to the cost of posting vacancies they would never form as new matches.



(a) Less than high school



(b) High school graduate



(c) College graduate

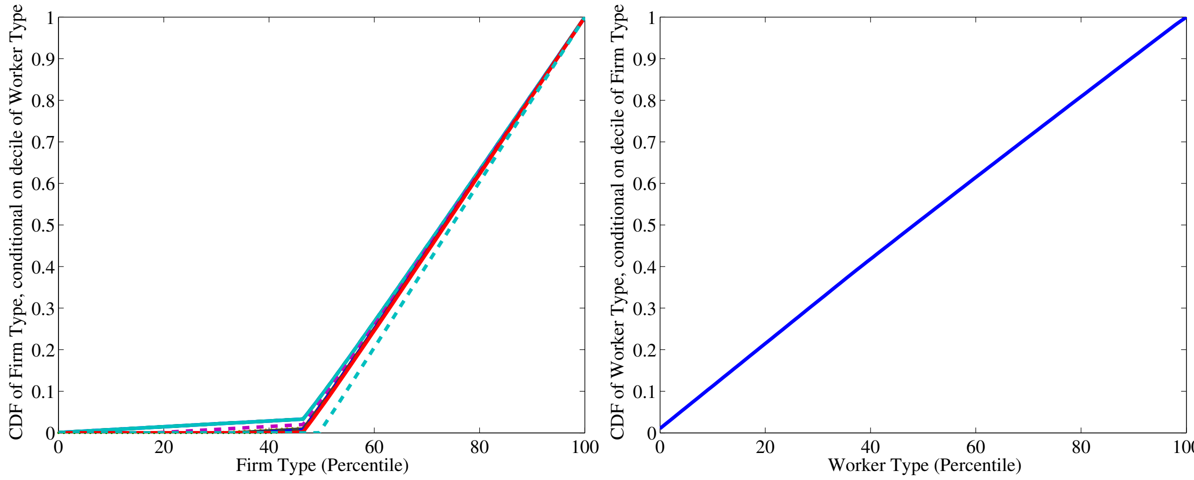
Figure 4: Matching Sets: Decentralized, Constrained Planner, Frictionless

Notes: Here we plot the set of matches that are formed between unemployed workers and vacant jobs. The green shaded area represents the decentralized matching set, the red lines are the boundaries of the constrained social planner's matching set, and the dotted blue line is the frictionless allocation.

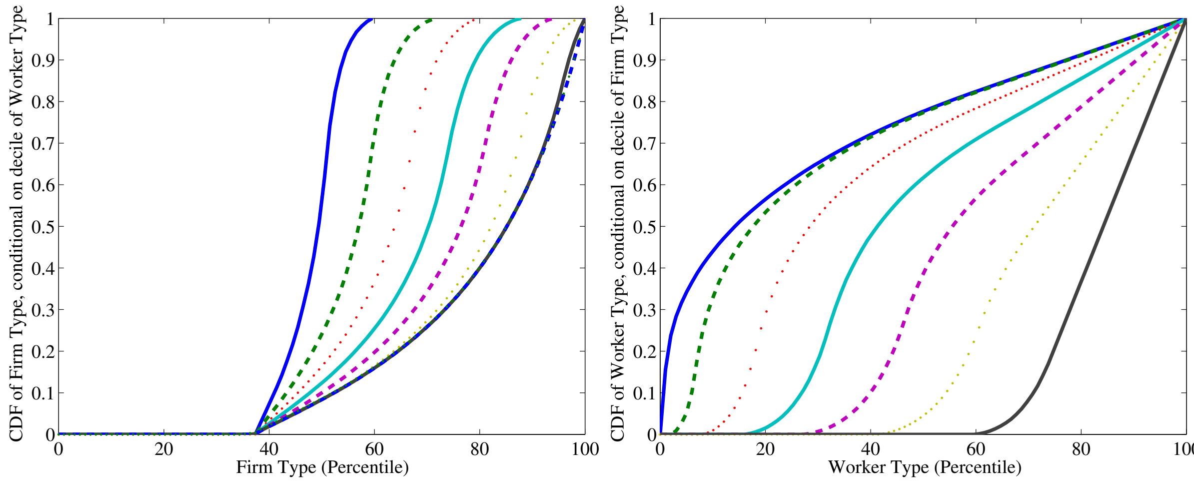
the distributions of worker productivity conditional on the firms' y value. The sole exception is the case of the unskilled where no such effect is observed and basically the two resulting equilibrium distributions are independent, resulting in no sorting. The kink in the distribution of productivities for jobs is due to the fact that no vacant job with productivity below that kink will find it profitable to post a vacancy. The ones that actually exist and operate are ones that acquired a worker when more productive but subsequently received a negative shock.

Turning now to the higher education groups: the graphs illustrate the amount of sorting and the fact that there is no perfect sorting. For example relatively low skill workers are associated with firms on a broad range of productivities and many different types of workers work in the same type of job in terms of y. It is however true that certain matches never happen. For example, the lowest skill worker among those with high school degrees do not work in firms above the 60th percentile. The other feature of these graphs is that as we move up the skill distribution they tend to spread out more, particularly the distribution of worker types conditional on firm types. This reflects increasing amounts of sorting at the higher education level. Moreover, the shift of the the shape of the cumulative distribution functions of workers (rhs) from concave to convex as we move from low productivity firms to higher ones is again the result of sorting, as lower productivity jobs have a higher concentration of lower skill (x) workers. At the same time these relationships are asymmetric between worker and firm types: the impact on the distribution of workers of moving to a higher quantile of firm type is stronger than the reverse (on the left). This is because while workers can search on the job, firms cannot search when the position is filled (by assumption). Generally, looking through these graphs it is clear from the support of these cumulative distributions that a number of types of matches never occur. Nevertheless, there is plenty of mismatch in the economy in terms of the optimal allocation of workers to jobs and in the next section we will quantify the welfare implications of this.

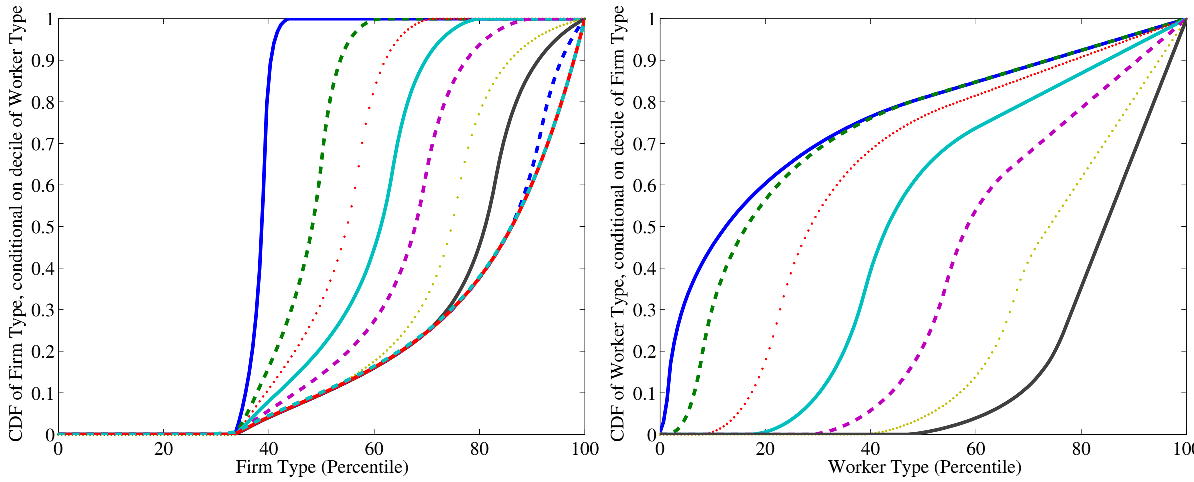
In Figure 6 we present the employment rate for workers and the activity of the firms. For all education groups (but the high school drop outs) the higher ability (x) workers have the highest employment rates - these increase with education, with very little unemployment for the highest education level, at least above the 10th percentile of ability. The right hand side panel shows the share of producing firms and those posting vacancies: to the left of the discontinuity no vacancies are posted because productivity is too low. Operating jobs in that range were filled



(a) Less Than High school



(b) High school graduate



(c) College

Figure 5: CDF of Match Distribution conditional on type .

Note: The left hand panels plot the CDF of firm type conditional on decile of worker type $H_y\left(y|x=X_d\right)$. In each of the plots on the left hand side, the initial linear segment corresponds to firm types for which vacancy posting is zero. The only flow into these matches is due to iid shocks to y. The outflow from any of these matches is the same (the probability of contacting a vacancy producing positive surplus). The shape to the right of the kink is dictated by the differential flow in and out of matches. The right hand panels plot the CDF of worker type, conditional on decile of firm type (and on positive vacancy posting) $H_x\left(x|y=Y_d,v\left(y\right)>0\right)$. For any firm type with zero vacancy posting, the CDF of worker types is (approximately) uniform due to the iid shocks to y.

when the firm was above the productivity threshold. The dotted line represents idle firms. All firms would be idle to the left of the threshold for posting vacancies if they had no worker, so this graph is zero after the threshold and one minus the producing firms to the left of it. The share of jobs actively looking for a worker is represented by the lowest line and to the right of the threshold increases with productivity because higher productivity firms are more choosy and take longer to fill a vacancy.

10 Welfare and Policy

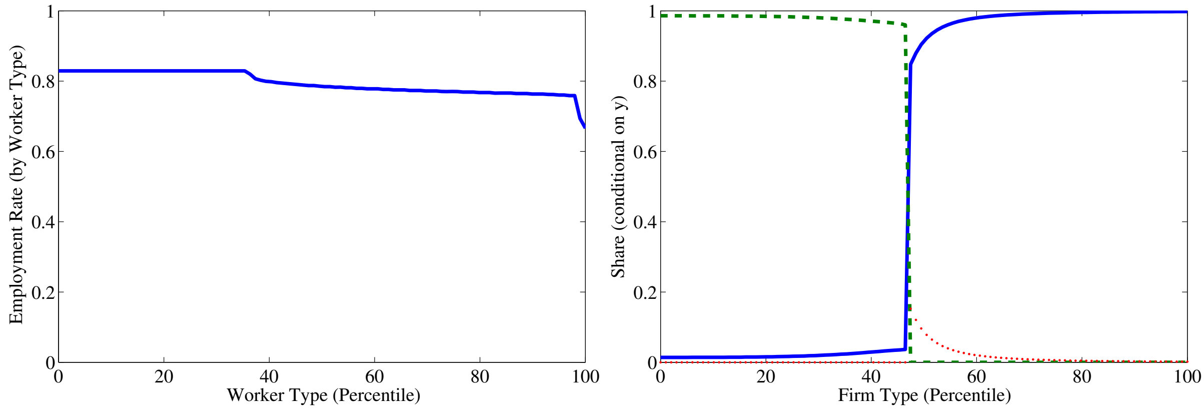
10.1 The impact of search frictions

The potential for labour market regulation arises from the job search frictions and the externalities they cause during the job allocation process. These externalities arise both from the classic issue of "overcrowding" among job seekers, i.e. when an extra person seeks a job it reduces the arrival rate for others, as implied by the matching function. An extra dimension arises in our model because of heterogeneity and sorting: by having low quality jobs compete for workers they lengthen the time it takes to fill higher productivity ones, without adding much when they are filled (because they have zero or near zero surplus). This implies that because of complementarity, cutting some low productivity jobs may increase welfare, even if this means that some very low productivity workers never work. An additional inefficiency comes from the fact that workers seek outside offers to increase their share of the match surplus. As a result, workers are willing to form matches even when the current output of the match is below the what they could produce at home, even after accounting for the cost of the vacancy: $f(x, y) < b(x) - c$. Even though the worker and firm are producing less together than they would separately, the fact that the firm has monopsony power up front means it is willing to hire the worker and extract most of the match surplus during the early periods of the match. The worker is willing to be in the match as it provides the needed outside option to extract surplus from poaching firms.

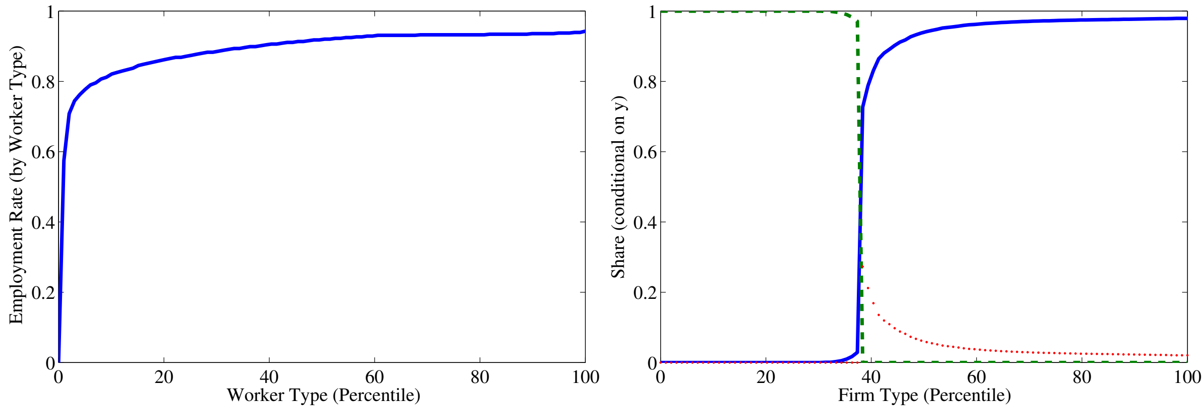
It is instructive to define the total potential output for the frictionless limit of the economy. This is calculated as

$$\int \max\{f(x, y(x)), b(x)\}\, dx,$$

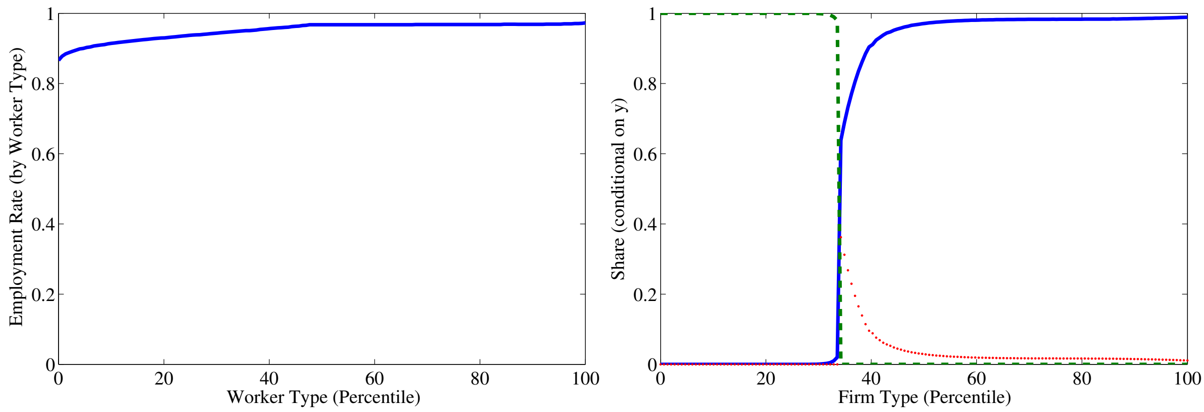
where $y(x)$ denotes the optimal job associated with worker type x.



(a) Less than high school

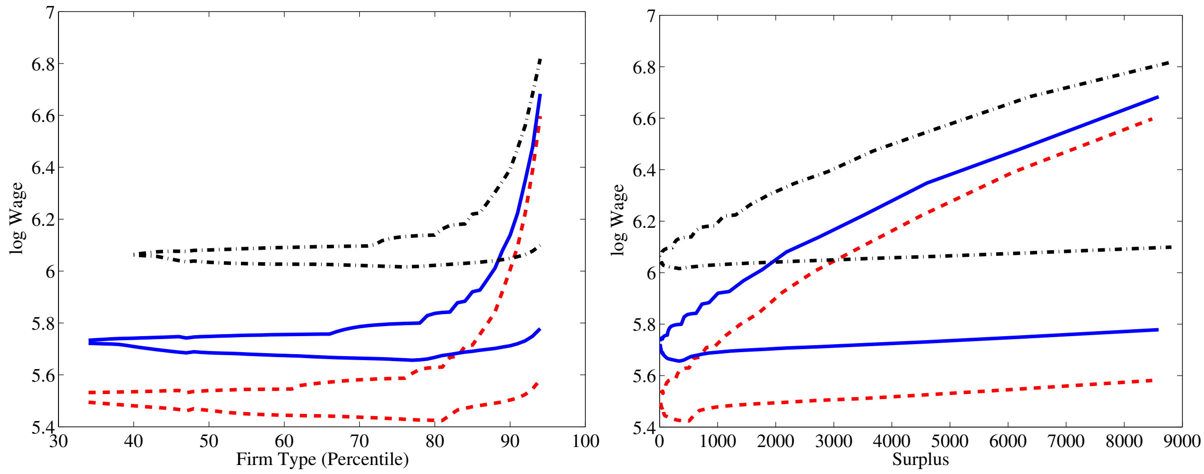


(b) High school graduate

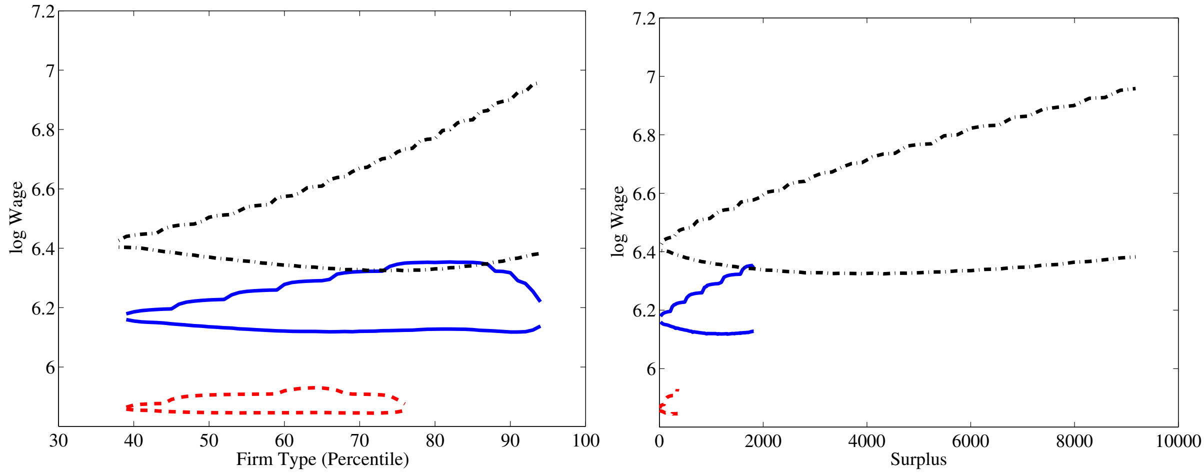


(c) College

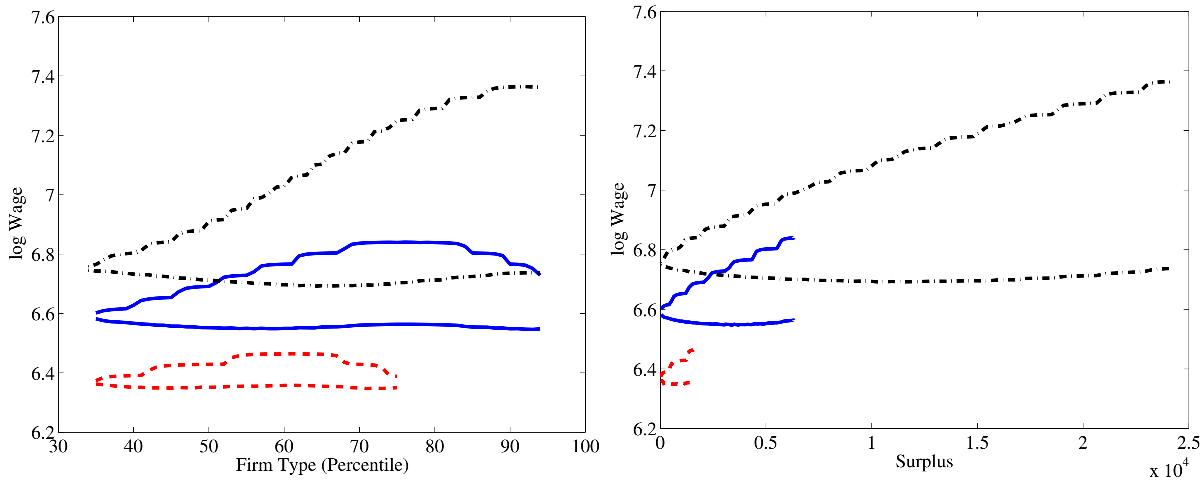
Figure 6: Employment Rate by Worker Type and Share of Producing, Vacant and Idle Jobs by Firm Type



(a) Less than high school



(b) High school graduate



(c) College

Figure 7: Wages, Firm Type and Match Surplus

Note: The lowest and highest possible wage in each match is plotted against firm type (left) and match surplus (right). The three wage bands are calculated at the 25th, 50th, and 75th percentiles of the worker type distribution. The lowest and highest wage in an (x, y) match are defined by $W_1\left(w^l, x, y\right) - W_0\left(x\right) = \beta S\left(x, y\right)$ and $W_1\left(w^h, x, y\right) - W_0\left(x\right) = S\left(x, y\right)$ respectively.

Other than specific attempts to change such frictions by improving the job search technology any intervention in the labor market will improve welfare only to the extent that it can address such externalities, if of course they are significant. Thus the planners problem offers an upper bound to welfare improvements through regulation, such as minimum wages, severance pay etc. The planner maximizes total output and home production subject to the flow constraints implied implied by the frictions, i.e.

$$\max_{h^{SP},u^{SP},V^{SP}} \int f(x,y)\, h^{SP}(x,y)\, dxdy + \int b(x)\, u^{SP}(x)\, dx - cV^{SP} \tag{17}$$

subject to the equations (12), (13), (14), (15) and (16). In the above the superscript SP is used to distinguish the endogenous objects chosen by the planner as opposed to those arising in a decentralized economy. An equivalent formulation to directly choosing the distribution of matches is to have the planner choose the set of admissible matches $\mathcal{M}(x,y)$. We approximate the solution to the planner's problem by allowing the planner to choose the boundry around the frictionless allocation: $\mathcal{M}^{SP}(x) = \left\{ y | \left(y - y(x)\right)^2 \leq \tau_0 + \tau_1 x + \tau_2 y + \tau_3 x^2 + \tau_4 y^2 + \tau_5 xy \right\}$.

Table 5 shows the breakdown of contributions to total welfare (steady state output) under different scenario. The first column relates to the fully decentralized economy we observe from the data. The second column shows the results of the planner maximizing welfare as in (17). For all but the lowest education groups this leads to an increase in welfare of about 3-5%. For the lowest education group there is a negligible increase in welfare. This has a key implication: we cannot justify any type of labor market regulation for the lowest skill group on the basis that it is welfare improving. Some corrective policy could be designed for the other groups but at best this could achieve an effect of 5% on welfare, if it was really well tuned.

For the lowest skill group other policies that improved the search technology would be ineffectual as well since there are no gains from sorting. Perhaps this result underlies the consistent failure to improve the labor market outcomes of the least skilled (for example see the analysis of the Canadian self-sufficiency program by Card and Hyslop (2004). However, when we look at the third column of the table for the other three, higher education groups we see that eliminating skill mismatch due to search frictions could lead to an 50% welfare increase relative to the benchmark. This is an astounding figure and seems to suggest that improving of matching technologies could have large effects on welfare if only we know how to achieve this. On the other hand the model dos not allow for mobility costs explicitly and these should be taken into account in a more general welfare calculation.

In the fourth column we repeat the exercise of optimally allocating workers, but in this case we take as given the distribution of unemployed from the decentralized economy. This column reallocates the employed workers, and leaves the unemployed as they were. Comparing columns 3 and 4 we have a measure of the unemployment cost of search frictions; between 0.03 and 2.8 percent as we move from the lowest to highest educated.

The final column presents the welfare (output) gains from an optimal unemployment insurance policy. We approximate unemployment insurance as a payment proportional to home production, which is positively related to permanent income: $(1 + b_0)\, b(x)$. The unemployment benefits are funded by a proportional tax on match output: $b_0 \int b(x)\, u(x)\, dx = \tau \int f(x, y)\, h(x, y)\, dxdy$. Due to the negligible gains for the lowest education group we find it is optimal to have no unemployment insurance for this group. In interpreting this result note that utility here is linear and UI has no insurance benefit per-se but can improve welfare by altering the job acceptance probabilities and indirectly eliminating very low value matches. For the other education groups, we find that optimal unemployment insurance can deliver between 10 and 20 percent of the potential welfare gains (the gains attainable by the constrained planner). This involves increasing the baseline flow utility of being out of work by between 68 and 147%.

10.2 The effects of labor market policy

Labor market regulation can only have a very limited role to play in improving overall welfare. However governments may use such policies for redistributive purposes, even if these are inefficient. So an interesting question is who actually benefits and who pays for such policies. In Table 6 we show the efficiency implications of two minimum wage policies, one where the minimum wage is set to half the pre-reform median wage and one which is set to a fifth of the median wage. The output declines are at most 1.7% (for the high school group). So from an efficiency point of view the policy is not particularly costly - indeed at 20% of the median the cost is negligible. However, the higher minimum wage implies large increases of unemployment. What happens is that a large amount of low value jobs close down. This has little impact on welfare because the surplus is minimal. It does however imply that a large proportion of unskilled workers remain

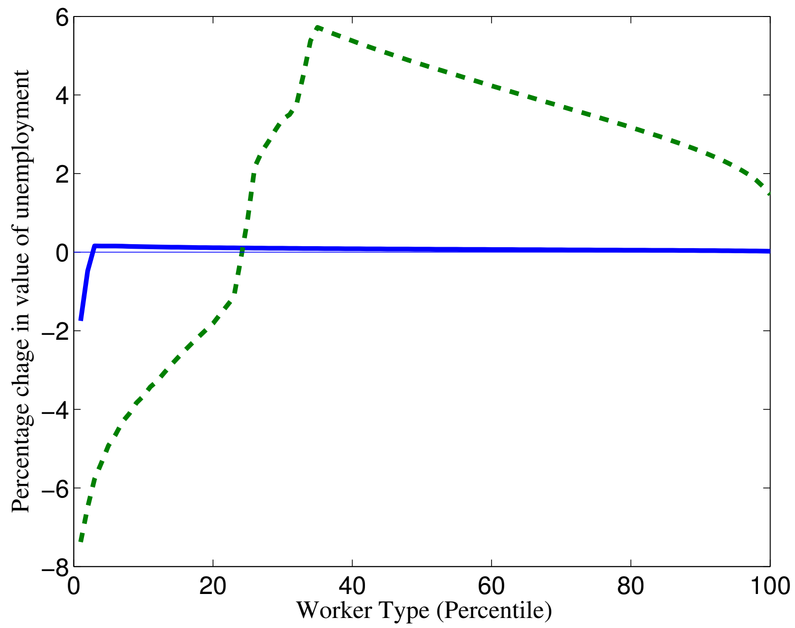


Figure 8: Redistributive Effect of Minimum Wage, Less than high school education
INCOMPLETE: BASED ON PREVIOUS ESTIMATES
Note: The percentage change in the value of unemployment is calculated as $100\frac{W_0(x|\underline{w})-W_0(x)}{W_0(x)}$. The solid line represents a minimum wage equal to one fifth the median wage (across all education groups) and the dotted line represents a minimum wage of one half the median.

out of the labor market. Since unemployment may have other impacts, not accounted for here, such as crime or implications for families and role models for children, one can only conclude that such a minimum wage policy at best has no benefits and may have longer term implications associated with joblessness among the lowest skills. Another way to document the redistributive impact of the minimum wage is provided in Figure 8 where we show the impact of the policy on the value of unemployment (which includes the benefits of future work opportunities) across the percentiles of worker types among the lowest skill group that is usually targeted for such policies. The solid line shows the impact of the low minimum wage, while the dashed line of the higher minimum wage. The lower minimum wage has only small negative effects for the lowest skill individuals and negligible positive effects for some. However, the higher minimum wage redistributes welfare from those below the 25 percentile of ability to those above the 25th percentile, with those at the 40th benefiting most. It is easy to comprehend why the minimum wage is both popular and harmful for those it is supposed to be benefiting.

Table 5: Output and Employment Effects of Policy

	Decentralized	Constrained Planner	Frictionless Benchmark	Frictionless conditional on unemployment	Optimal Unemployment Insurance
			Less Than High school		
Steady-state output	100.00	100.14	100.33	100.30	100.00
Market Production	99.95	100.03	100.28	100.25	99.95
Home Production	0.05	0.11	0.06	0.05	0.05
Recruiting Costs	-0.001	-0.0003	-	-	-0.001
Employment Rate	79.35	50.17	60.00	79.35	79.35
corr (x, y)	0.04	0.13	1.00	1.00	
			High school		
Steady-state output	100.00	117.86	147.94	143.69	100.60
Market Production	99.47	117.11	147.78	143.16	99.99
Home Production	0.54	0.76	0.15	0.54	0.56
Recruiting Costs	-0.007	-0.011	-	-	-0.007
Employment	88.58	83.85	93.00	88.58	85.94
corr (x, y)	0.61	0.78	1.00	1.00	
			College		
Steady-state output	100.00	115.06	131.51	129.09	100.63
Market Production	99.86	114.78	131.51	128.94	100.39
Home Production	0.14	0.28	0.00	0.00	0.23
Recruiting Costs	-0.001	-0.002	-	-	-0.001
Employment	95.10	89.75	100.00	95.10	90.02
corr (x, y)	0.66	0.87	1.00	1.00	

INCOMPLETE: PRELIMINARY ESTIMATES

Notes: The frictionless benchmark is the positive assortative allocation $\{x, y(x)\}$, and full employment. The frictionless conditional on given employment takes the distribution of unemployed from the decentralized economy as given and allocates the remaining workers according to $\{x, y(x)\}$. The constrained planner chooses chooses admissible matches to maximize steady state output. Here we restrict the planner to choosing a range of admissible matches around the frictionless allocation: $\mathcal{M}^{SP}(x) = \left\{y | (y - y(x))^2 \leq \tau_0 + \tau_1 x + \tau_2 y + \tau_3 x^2 + \tau_4 y^2 + \tau_5 xy\right\}$. . Optimal unemployment insurance is modeled as the tax and benefit pair that maximizes steady-state output. The unemployment insurances is modeled as proportional to home production:$(1 + b_0) b(x)$, and we impose the balanced budget $b_0 \int b(x) u(x) dx = \tau \int f(x, y) h(x, y) dxdy$. ε is the elasticity of participation with respect to a change in benefits. less than high school: $\tau = 0$, $b_0 = 0$, $\varepsilon = .$,high school: $\tau = 0.001$, $b_0 = 0.12$, $\varepsilon = -0.25$, and college: $\tau = 0.01$, $b_0 = 0.54$, $\varepsilon = -0.10$

Table 6: Counterfactual Minimum Wage Experiments

	Less Than High school	High school	College
	Minimum wage equal to half median		
Steady-state output	99.14		99.57
Market Production	98.32		99.55
Home Production	0.91		0.22
Recruiting Costs	-0.09		-0.20
Employment rate	64.80		92.64
	Minimum wage equal to one fifth median		
Steady-state output	99.99		100.00
Market Production	99.40		100.01
Home Production	0.65		0.19
Recruiting Costs	-0.07		-0.19
Employment rate	81.64		94.39

INCOMPLETE: BASED ON PREVIOUS ESTIMATES
Notes: Steady-state output is expressed relative to the decentralized economy.

11 Conclusion and Further Work

We develop an equilibrium model of employment and wage determination, which builds on the work of Postel-Vinay and Robin (2002) and Shimer and Smith (2000). In our model both workers and firms are heterogeneous, their productivity characteristics are potentially complimentary in production creating the possibility of sorting. However, firms are subject to productivity shocks. Workers can search both on and off the job. This creates an environment where there may be potential for welfare improving labor market regulation. Moreover our framework is well suited to consider the redistributive (as well as efficiency) implications of policy. The scope for and impact of policy is thus an empirical issue in our model.

We estimate the model based on NLSY data and find strong evidence of sorting among all but the labor market for the lowest education workers; for the latter there is no sense in which they would be mismatched and a different allocation of firms and workers could improve output. However for all other educational groups these complementarities imply large efficiency losses due to mismatch between job and worker productivities caused by search frictions.

Mismatch is a source of inefficiency that labor market regulation cannot correct; this would require changing the job search technology improving job finding rates and enabling more mobility following shocks. Policies such as minimum wages can improve efficiency to the extent that they address the externalities induced by search frictions. We established that these are very small. We then show that the optimal minimum wage is zero for the lowest skill group; moreover we show that a minimum wage policy setting the minimum at half the median wage would have low efficiency losses but would increase unemployment substantially and would redistribute wealth from the lowest skill to workers above the lower quartile of ability.

Our model opens up an empirical research agenda on which to build and address important issues. We have demonstrated the importance of heterogeneity, sorting and search frictions. Among these are the welfare and labour market effects of risk and the role of assets in determining the wage offer distribution and the role of investment in human capital. Similarly, an important extension of such a model is considering investment decisions by forms and how this can affect productivity y which we took as given. Finally, this kind of model is well suited to interpreting matched employer employee data. Indeed such data could aid identification by providing direct information on firm level productivity. However, when we move to such data new, important and difficult questions arise when defining wage setting in an environment with sorting and multiple workers per firm. This is of course an important future research area.

References

Abowd, J., F. Kramarz, P. Lengermann, and S. Roux (2005): "Persistent inter-industry wage differences: Rent sharing and opportunity costs," .

Abowd, J., F. Kramarz, and D. Margolis (1999): "High wage workers and high wage firms," *Econometrica*, 67(2), 251–333.

Abowd, J. M., and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–445.

Abowd, J. M., F. Kramarz, S. Pérez-Duarte, and I. Schmutte (2009): "A Formal Test of Assortative Matching in the Labor Market," NBER Working Papers 15546, National Bureau of Economic Research, Inc.

Altonji, J., and N. Williams (1997): "Do wages rise with job seniority? A reassessment," p. 34.

Atakan, A. E. (2006): "Assortative Matching with Explicit Search Costs," *Econometrica*, 74, 667–680.

Bagger, J., F. Fontaine, F. Postel-Vinay, and J.-M. Robin (2011): "Tenure, Experience, Human Capital and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," .

Bagger, J., and R. Lentz (2008): "An Equilibrium Model of Wage Dispersion and Sorting," (March), 1–25.

Becker, G. S. (1973): "A Theory of Marriage: Part I," *The Journal of Political Economy*, 81, 813–846.

Burdett, K., and K. L. Judd (1983): "Equilibrium Price Dispersion," *Econometrica*, 51(4), 955–969.

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

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006): "Wage Bargaining with On-the-Job Search: Theory and Evidence," *Econometrica*, 74(2), 323–364.

Card, D., and D. Hyslop (2004): "Estimating the Dynamic Treatment Effects of an Earnings Subsidy for Welfare-Leavers," *Econometrica*, 73(6).

Chernozhukov, V., and H. Hong (2003): "An MCMC approach to classical estimation," *Journal of Econometrics*, 115(2), 293–346.

de Melo, R. L. (2009): "Sorting in the Labor Market: Theory and Measurement," Discussion paper, University of Chicago.

Dey, M., and C. Flinn (2005): "An Equilibrium Model of Health Insur- ance Provision and Wage Determination," *Econometrica*, 73(2), 571–627.

Diamond, P. A. (1982): "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy*, 90(5), 881–894.

Heckman, J. J., and B. E. Honore (1990): "The Empirical Content of the Roy Model," *Econometrica: Journal of the Econometric Society*, 58, 1121–1149.

Heckman, J. J., and G. Sedlacek (1985): "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market," *The Journal of Political Economy*, 93, 1077–1125.

Krueger, A., and L. Summers (1988): "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, 56(2), 259–293.

Lentz, R. (2010): "Sorting by search intensity," *Journal of Economic Theory.*

Lu, X., and R. McAffee (1996): *Matching and expectations in a market with heterogeneous Agents.* Advances in Applied Microeconomics.

MacLeod, W. B., and J. M. Malcomson (1993): "Investments, Holdup, and the Form of Market Contracts," *The American Economic Review*, 83(4), 811–837.

MaCurdy, T. E. (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in Longitudinal Data Analysis," *Journal of Econometrics*, 18, 82–114.

Meghir, C., and L. Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1), 1–32.

Mortensen, D. T. (1982): "Property Rights and Efficiency in Mating, Racing, and Related Games," *The American Economic Review*, 72, 968–979.

Moscarini, G. (2001): "Excess Worker Reallocation," *The Review of Economic Studies*, 68, 593–612.

Pissarides, C. (2000): *Equilibrium Unemployment Theory.* MIT Press.

Postel-Vinay, F., and J.-M. Robin (2002): "Equilibrium W age Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70(6), 2295–2350.

Postel-Vinay, F., and H. Turon (2010): "On-the-job search, productivity shocks, and the individual earnings process," *International Economic Review.*

Robert, C. P., and G. Casella (2004): *Monte Carlo Statistical Methods.* Springer, New York, 2 edn.

Sattinger, M. (1995): "Search and the Efficient Assignment of Workers to Jobs," *International Economic Review*, 36(2), 283–302.

Sattinger, M. (2003): "A Search Version of the Roy Model," .

Shimer, R. (2005): "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," *Journal of Political Economy*, 113, 996–1025.

——— (2006): "On-the-Job Search and Strategic Bargaining," *European Economics Review*, 50, 811–830.

Shimer, R., and L. Smith (2000): "Assortative Matching and Search," *Econometrica*, 68, 343–369.

Sisson, S. A., and Y. Fan (2011): "Likelihood-free Markov chain Monte Carlo," in *Handbook of Markov Chain Monte Carlo.* Chapman and Hall.

Vrugt, J., C. ter Braak, C. Diks, D. Higdon, B. Robinson, and J. Hyman (2009): "Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling," *International Journal of Nonlinear Sciences and Numerical Simulation*, 10(3), 273–290.

A Web Appendix

A.1 Derivation of expressions

A.1.1 The Continuation value and the equation for the surplus

Let $P(x,y)$ be the value of joint production of an (x,y) match. Then the surplus is defined by $P(x,y) - W_0(x) - \Pi_0(y) = S(x,y)$. Assume jos can also be destroyed at an exogenous rate ξ. Tehn we have that

$$
\begin{aligned}
rP(x,y) &= f(x,y) + \xi\left[W_0(x) + \Pi_0(y) - P(x,y)\right] \\
&\quad + s_1\kappa \int \left[\max\left\{P(x,y), \Pi_0(y) + W_0(x) + S(x,y) + \beta\left[S\left(x,y'\right) - S(x,y)\right]\right\} - P(x,y)\right] v\left(y'\right) dy' \\
&\quad + \delta \int \left[\max\left\{P\left(x,y'\right), W_0(x) + \Pi_0\left(y'\right)\right\} - P(x,y)\right] dy' \\
&= f(x,y) - \xi S(x,y) \\
&\quad + s_1\kappa \int \left[\max\left\{0, \beta\left[S\left(x,y'\right) - S(x,y)\right]\right\}\right] v\left(y'\right) dy' \\
&\quad + \delta \int \left[\max\left\{P\left(x,y'\right) - W_0(x) - \Pi_0\left(y'\right), 0\right\} - P(x,y) + W_0(x) + \Pi_0\left(y'\right)\right] dy' \\
&= f(x,y) - \xi S(x,y) \\
&\quad + s_1\kappa\beta \int \left[S\left(x,y'\right) - S(x,y)\right]^+ v\left(y'\right) dy' \\
&\quad + \delta \int S\left(x,y'\right)^+ dy' - \delta S(x,y) + \delta \int \left[\Pi_0\left(y'\right) - \Pi_0(y)\right] dy'
\end{aligned}
$$

Then, substituting out $rW_0(x)$ and $r\Pi_0(y)$, we have $S(x,y)$ defined by the fixed point

$$\begin{aligned}(r+\xi+\delta)\,S(x,y) = f(x,y) - b(x) - s_0\kappa\beta \int S(x,y)^+\, v(y)dy \\ + c - s_0\kappa(1-\beta)\int S(x,y)^+\, u(x)dx \\ - s_1\kappa(1-\beta)\iint \left[S(x,y) - S(x,y')\right]^+ h(x,y')\,dx\,dy' \\ + s_1\kappa\beta \int \left[S(x,y') - S(x,y)\right]^+ v(y')\,dy' + \delta \int S(x,y')^+\,dy'. \quad (18)\end{aligned}$$

Note that

$$\left[r+\delta+\xi+s_1\kappa v\left(\overline{B}(x,y)\right)\right]\frac{\partial S(x,y)}{\partial y} = \frac{\partial f(x,y)}{\partial y} - (r+\delta)\Pi_0'(y), \quad (19)$$

where

$$\overline{B}(x,y) = \{y' : S(x,y') \geq S(x,y)\},$$

and $\mu(A) = \int_A \mu(y)\,dy$, for any set A and any measure density μ. Therefore, for any x, the set of y's maximising the surplus $S(x,y)$ is the set of y's maximising the surplus flow $f(x,y) - rW_0(x) - r\Pi_0(y)$.

A.1.2 The expected profits of a vacant job increase with productivity

From the previous expression note also that

$$\begin{aligned}(r+\delta)\Pi_0'(y) = s_0\kappa(1-\beta)\int_{S(x,y)\geq 0} \frac{\partial S(x,y)}{\partial y} u(x)dx \\ + s_1\kappa(1-\beta)\int h(x,B(x,y))\frac{\partial S(x,y)}{\partial y}dx, \quad (20)\end{aligned}$$

where $B(x,y)$ is the set of jobs with a productivity y' leading to a lower surplus than the pair (x,y),

$$B(x,y) = \left\{y' : S(x,y) \geq S(x,y')\right\}.$$

Hence, plugging the above expression for $\partial S(x,y)/\partial y$ in equation (20) shows that $\Pi_0'(y)$ is positive if $\partial f(x,y)/\partial y$ is positive.

A.1.3 Computing the equilibrium

The equilibrium is characterized by knowledge of the number of jobs N, the labour market tightness $\kappa(U, V)$, the joint distribution of active matches $h(x, y)$, and the surplus function $S(x, y)$. A fixed point iterative algorithm operating on (κ, h, S) can be constructed as follows.

First, with inputs κ, $h(x, y)$ and $S(x, y)$,

1. Calculate $u(x)$ using (14) and calculate $U = \int u(x)\, dx$.

2. Solve for V in equation
$$\kappa = \frac{M\left(s_0 U + s_1(L-U), V\right)}{\left[s_0 U + s_1(L-U)\right] V}.$$

3. Calculate $N = V + L - U$ and calculate $v(y)$ with equation (15).

Second,

1. Update h using equation (13) as

$$h(x,y) \leftarrow \frac{\delta \int h(x,y')\, dy' + \left[s_0 u(x) + s_1 h\left(x, B(x,y)\right)\right] \kappa v(y)}{\delta + \xi + s_1 \kappa v\left(\overline{B}(x,y)\right)} \mathbf{1}\{S(x,y) \geq 0\}. \tag{21}$$

2. Update S using equation (10).

3. Update κ using the free entry equation (9)

$$\kappa \leftarrow \frac{(c - \delta \overline{\Pi}_0)/(1-\beta)}{s_0 \int S(x,0)^+ u(x)\, dx + s_1 \int \left[S(x,0) - S(x,y')\right]^+ h(x,y')\, dx\, dy'}. \tag{22}$$

Alternatively, one can solve for (h, S) for a given κ and search for the κ that satisfies the free entry condition. The full iterative fixed point algorithm does not indeed guaranty positive updates for κ.

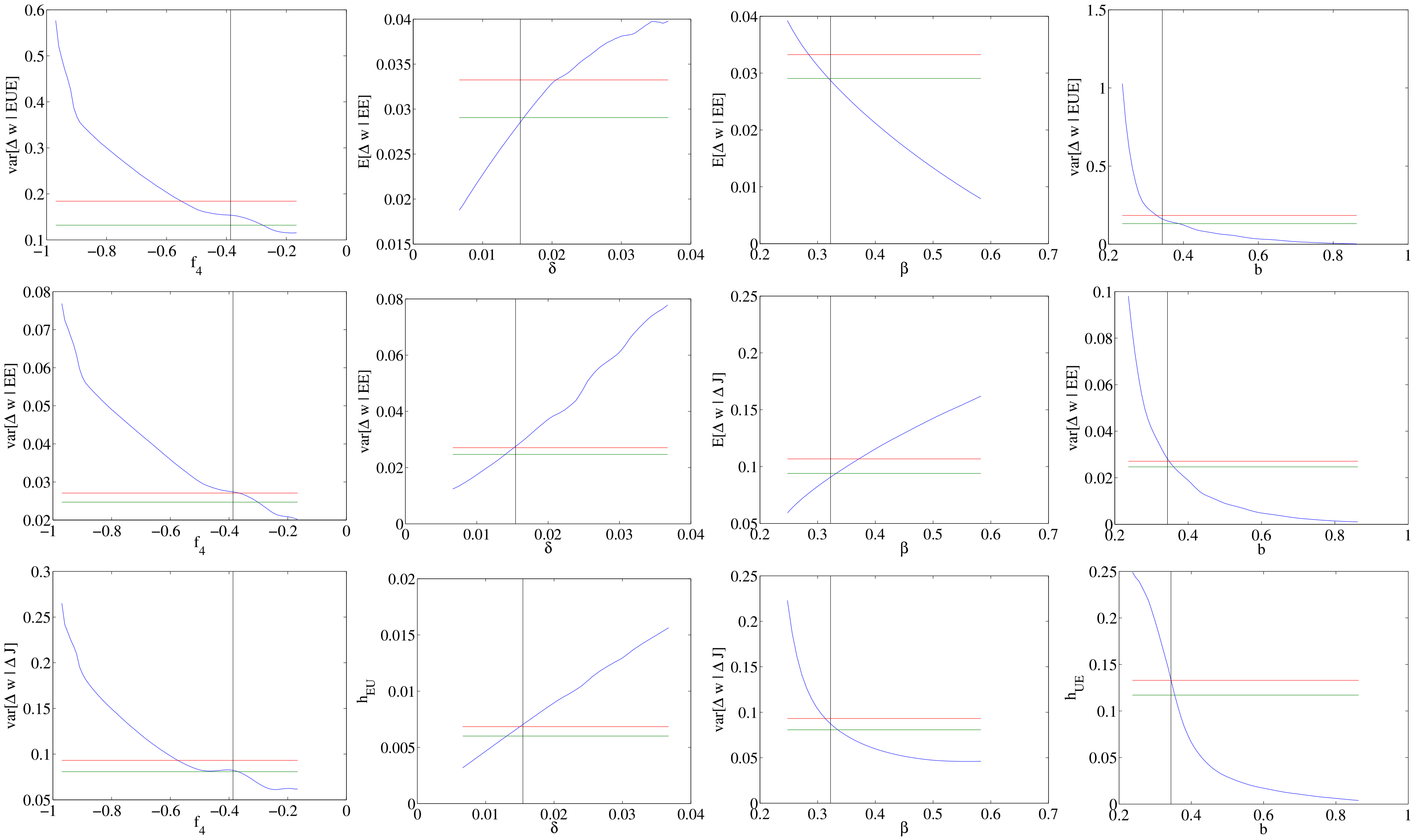


Figure 9: Identification of complementarity, productivity shocks, bargaining power and home production

Note: Based on the parameter estimates for college educated workers. The vertical black line represents the estimated value of the parameter on the horizontal axis. The horizontal red and green lines represent the value of the data moment on the vertical axis, plus and minus two standard errors.