# MISMATCH, SORTING AND WAGE DYNAMICS

By

Jeremy Lise, Costas Meghir and Jean-Marc Robin

January 2013

COWLES FOUNDATION DISCUSSION PAPER NO. 1886

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

# Mismatch, Sorting and Wage Dynamics*

Jeremy Lise[†] Costas Meghir[‡] Jean-Marc Robin[§]

January 2013

**Abstract**

We develop an empirical search-matching model which is suitable for analyzing the wage, employment and welfare impact of regulation in a labor market with heterogeneous workers and jobs. To achieve this we develop an equilibrium model of wage determination and employment which 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 productivity shocks, long-term contracts, on-the-job search and counter-offers. Importantly, the model allows for the possibility of assortative matching between workers and jobs due to complementarities between worker and job characteristics. We use the model to estimate the potential gain from optimal regulation and we consider the potential gains and redistributive impacts from optimal unemployment insurance policy. The model is estimated on the NLSY using the method of moments.

KEYWORDS: Sorting, mismatch, search-matching, wage dynamics

---

*We have benefitted greatly from comments and questions from seminar participants at Berkeley, Chicago, Columbia, Georgetown, Harvard, Indiana, Minneapolis Fed, NYU, Northwestern, Princeton, Stanford, Wisconsin, UBC, Simon Fraser University, IIES, Stockholm School of Economics, Cergy-Pontoise, Edinburgh, York, Southampton, Bristol, CEMFI, Oslo, Oxford, Tilburg, EUI, Uppsala, Queen Mary, Carlos III, SMU, Central Bank of Turkey, Koc, SITE, NBER summer meeting, SED meetings, Cowles Summer Conference, Crest-Thema SaM, and SOLE meetings in NY and from numerous individuals including Joe Altonji, Thibaut Lamadon, Dale Mortensen, Guisepe Moscarini, John Moore, Bentley McLeod, Rob Shimer and many others. Costas Meghir thanks the ESRC for funding under the Professorial Fellowship RES-051-27-0204 and the Cowles foundation at Yale. Jean-Marc Robin gratefully acknowledges financial support from the Economic and Social Research Council through the ESRC Centre for Microdata Methods and Practice grant RES-589-28-0001, and from the European Research Council (ERC) grant ERC-2010-AdG-269693-WASP. The authors are responsible for all errors or interpretations.

[†]University College London and IFS. j.lise@ucl.ac.uk

[‡]Yale University, NBER and IFS. c.meghir@yale.edu

[§]Sciences Po, Paris and University College of London. jeanmarc.robin@sciences-po.fr.

# 1 Introduction

Labor market imperfections justify labor market interventions. Within a competitive framework regulation would be welfare reducing, as it would typically reduce employment and increase insiders' wages. By contrast, any friction constraining the allocation of workers to jobs inevitably spawns many mismatched job-worker pairs and allows some agents to appropriate a greater share of the rent than a central planner would deem fit. Indeed, if there are important complementarities in production, mismatch may produce substantial welfare losses relative to the first best or a constrained planner. In the now large body of literature using search models to describe labor markets with heterogeneous agents, no paper, as far as we know, has estimated a search-matching model with two-sided heterogeneity and used it to infer the welfare cost of mismatched workers and firms. A fortiori, no empirical study has evaluated the role of standard labor market regulations, such as unemployment insurance, in augmenting or reducing mismatch. Our aim in this paper is to lay a first foundation stone to fill this gap.

We develop a search-matching model in which workers with different abilities are assigned to different tasks. Our model links with the current literature on equilibrium search and assortative matching. It extends the model of Shimer and Smith (2000) in various directions (wage determination, on-the-job search, shocks to individual characteristics) and provides a bridge between an essentially theoretical literature on allocation of heterogeneous workers to jobs[1], and the large microeconometric literature on job mobility and wage dynamics.[2]

We estimate our model using panel data on workers and use the estimates to quantify how much sorting there is with respect to unobserved characteristics. A first set of studies interested in this question follow Abowd, Kramarz, and Margolis (1999, AKM) and assess the degree of assortative matching in the labor market by calculating the correlation between worker and firm fixed effects estimated by a panel-data regression of individual wages on workers' and employers' indexes.[3] They typically find non-significant or negative correlations. However, Andrews, Gill, Schank, and Upward (2008) show that this estimated correlation is contaminated by a spurious statistical bias and propose a bias-corrected estimator. Using German IAB data, they find that the negative OLS estimate is turned into a positive number after applying the bias correction (0.23 instead of somewhere in the range $[-0.19, -0.15]$).

However, Eeckhout and Kircher (2011) strongly argue against the possibility of identifying assortative matching in labor market by this approach, bias-corrected or not. This is because the surplus of a match is not a monotonic function of worker and firm characteristics in general. Depending on the distribution of matches around the optimal, Beckerian allocation, a positive or a negative AKM correlation can be estimated irrespective of the sign of the correlation between

[1] See also Shi (2001), Teulings and Gautier (2004), Moscarini (2005)and Gautier, Teulings, and van Vuuren (2010) for alternative approaches to two-sided matching models without and with on-the-job search.

[2] See amongst others MaCurdy (1982), Altonji and Shakotko (1987), Abowd and Card (1989), Topel (1991), Topel and Ward (1992), Meghir and Pistaferri (2004), Altonji and Williams (2005), Guvenen (2007), Bonhomme and Robin (2009), Altonji, Smith, and Vidangos (2009), Guvenen (2009), Low, Meghir, and Pistaferri (2010), Lise (2012).

[3] See Goux and Maurin (1999) and Abowd, Kramarz, Pérez-Duarte, and Schmutte (2009) who present results for French and U.S. matched employer-employee data, and Gruetter and Lalive (2009) for Austrian data.

workers' and firms' true unobserved characteristics in the population of active matches. We also find that the AKM correlation is misleading and show evidence that positive sorting with respect to unobserved characteristics may induce a negative correlation in the worker and firm effects estimated on a panel of wages. Lopes de Melo (2009), Bagger and Lentz (2012), Hagedorn, Law, and Manovskii (2012) reach similar conclusions with different models.

Identification of complementarity and sorting, given that we use a panel of workers' wages and labor market transitions extracted from the NLSY,[4] is also hard to prove or disprove theoretically. However, we argue that sorting can be seen in the way wage and employment mobility vary as a function of the length of time spent working following an unemployment spell. Due to search frictions, the cohort of workers entering the labor market following a spell out of work will start off mismatched, but through on-the-job search they will become better and better matched with time spent in the labor market. If the tendency to sort in equilibrium is strong this will result in wages spreading out as workers sort themselves. Monte Carlo simulations for the Simulated Method of Moment estimator that we have implemented here seem to suggest that the set of moments we match do the degree of identify sorting. We find that the NLSY data are best fitted by our model assuming no complementarity and zero sorting for high-school dropouts and positive complementarity and sorting for high-school graduates and college-graduates.

Finally, our model offers an empirical framework for understanding employment and wage determination in the presence of firm-worker complementarities, search frictions and productivity shocks. As a result it offers a way for evaluating the extent to which regulation may be welfare improving and can evaluate the impact of specific policies such as unemployment insurance. In a search framework with match complementarities unemployment insurance can have ambiguous effects on employment and total output. On the one hand, it allows workers to be more picky and form better matches, which comes at the cost of longer unemployment spells and higher unemployment. On the other hand, the fact that higher quality matches will be formed may induce firms to create more jobs, increasing the contact rate and potentially reducing unemployment duration. Our framework allows this effect to be quantified. And, in addition, it allows us to analyze the effect of such policies on the distribution of welfare thus showing who pays and who benefits from such a policy in this non-competitive environment.

We find that the degree to which labor market interventions are justified depends on whether we are looking at the low, medium or high skilled markets. Our finding that the market for unskilled labor (high school drop outs) is characterized by an extremely low degree of complementarity implies that mismatch is not very costly, indeed a constrained social planner would not choose to alter the allocation. In contrast, the high degree of complementarity we find in the markets for higher skilled labor (high school graduates and college graduates) implies that the mismatch of workers to jobs arising in the decentralized equilibrium is quite costly. As a result, a constrained planner can improve total output by trading off the level of employment against

[4]We preferred to stick to standard worker panel data for two reasons. First, matched employer-employee data are less universal and not always easily accessible to researchers. Second, value-added per worker (used by Cahuc et al., Bagger et al. and Bagger and Lentz) does not measure well the labor productivity of a single job isolated from the other jobs in a firm.

the quality of matches—choosing a lower employment level (and fewer costly vacancies) and higher output per match by preventing low quality matches from forming (which also reduces the congestion externality). Finally, we find that if we limit the planner to an optimal unemployment insurance program, this can go a long way to realizing the potential welfare gains for the high school group. However, it is ineffectual for the college educated group as the resulting employment distortions outweigh the gains from improved match quality.

The paper proceeds as follows. Section 2 describes the model and Section 3 the estimation procedure. Section 4 presents the results of estimation, analyses the fit and discusses estimation of the degree of sorting. Section 5 presents the welfare analysis. Section 6 analyses in detail unemployment insurance policy. Section 7 concludes.

## 2 The Model

Our model is an equilibrium model of the labor market in which all agents behave optimally and wages are endogenous. It produces rich individual level dynamics, and endogenous distributions of worker-job matches, unemployed workers, and vacant jobs. There is no aggregate uncertainty, and hence we focus on stationary equilibrium.

In this economy there are a fixed number of infinitely lived individuals, while the number of jobs is endogenously determined so that the marginal entrant makes zero expected profits. Workers are matched pairwise to jobs or tasks. They differ from each other according to ability, and ability differences are permanent. As in Mortensen and Pissarides (1994), jobs are subject to persistent idiosyncratic productivity shocks, which can trigger separations and wage renegotiations. The production function allows for complementarity between worker and job characteristics, leading to the possibility of sorting in the labor market. Any worker-job pair that meet will form a match when the surplus is non-negative — this accounts for the option value of each partner of continuing to search for a better partner. Worker and firm heterogeneity may be unobserved by the econometrician but they are observed by all parties when meetings take place. Workers receive job offers while unemployed or employed at rates that differ in the two states and depend on the endogenously determined market tightness. Search on the job may trigger quits and wage renegotiations. We assume random matching within a well defined labor market (defined by worker education), so while workers cannot direct their search to specific firm types within their labor market, they direct their search to jobs segmented by educational requirements. We do not explicitly model how this segmentation arises.

An important feature of the model is that it generates wage dynamics. We assume that wage contracts are long-term contracts which can be renegotiated only 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). On-the-job search may thus lead to outside offers and hence to wage changes and/or mobility. In addition, firm-level productivity shocks may move the current contract outside the bargaining set and lead to wage renegotiation or to match dissolution and a period of unemployment before reentering at a different wage (see Postel-Vinay and Turon, 2010). The wage dynamics are thus explained by the structure of the model. The question is whether the predicted dynamics match those documented in the empirical papers mentioned earlier. If they do this offers a structural interpretation to such processes based on productivity shocks and behavioral responses in wage setting. Note that in this paper we abstract from human capital accumulation, which, while important, would complicate matters beyond the scope of this first paper.[5]

## 2.1 Workers, jobs and matches

Each individual worker is characterized by a level of permanent ability, continuously distributed across workers, observable by all agents, but not by the researcher. Let $x$ denote the rank of ability across workers. There are $L$ individuals of which $U$ are unemployed. We also denote by $u(x)$ the (endogenous) measure density of $x$ among the unemployed.

Jobs are characterized by a technological factor, say labor productivity, which is also continuous and observable by all agents, but not the researcher. Let $y$ denote the rank of labor productivity across jobs, vacant or filled. There are $N$ jobs in the economy, a number that is endogenously determined. The endogenous measure density of vacant posts is $v(y)$ and the resulting number of vacancies is denoted by $V = \int v(y)\,dy$.

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: shocks arrive at a rate $\delta$, and a new productivity level $y'$ is drawn from the uniform distribution on $[0, 1]$. Thus the size of $\delta$ determines the persistence of the shocks.

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. Matches can end both endogenously, as we characterize later, and exogenously. We denote by $\xi$ the rate of exogenous job destructions, defined as an event that is unrelated to either worker or job characteristics.

Equilibrium will result in a joint distribution of $x$ and $y$, the distribution of matches. We denote this by $h(x, y)$. Because $x$ and $y$ are uniformly distributed in the whole populations of workers and jobs, 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 and posted vacancies should add up to the total number of jobs $N$:

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

[5] One possibility to make this extension tractable is to assume piece-rate contracts as in Barlevy (2008) and Bagger, Fontaine, Postel-Vinay, and Robin (2011).

## 2.2 Match formation

Individuals and jobs are risk neutral and we assume efficiency, in the sense that any match where the surplus is positive will be formed when the worker and the job meet. 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 when out of work (or home production) 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 net of expected posting costs. Let also $W_1(w, x, y)$ (respectively $\Pi_1(w, x, y)$) denote the present value of a wage contract $w$ for a worker $x$ employed at a job $y$ (respectively the firm's discounted profit).

The surplus of an $(x, y)$ match is defined by

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

For a pair $(x, y)$ a match is feasible and sustainable if the surplus is nonnegative, $S(x, y) \geq 0$.

## 2.3 Wage negotiation with unemployed workers

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 $\beta$:

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

Through the process of on-the-job search and offers/counteroffers her wage will subsequently grow. A simple assumption would be to set $\beta$ to zero. However this may lead to the counterfactual implication that for some matches initial wages are negative implying workers pay to obtain a high value job with the potential of future wage increases. Although this is not implausible (unpaid internships) this is never observed in our survey data. By allowing $\beta$ to be determined by the data we avoid this problem and allow greater flexibility in fitting the facts.

## 2.4 Renegotiation

Wages can only be renegotiated when either side has an interest to separate if they do not obtain an improved offer, assuming that the match remains viable for both parties. The events that can trigger renegotiation occur when a suitable outside offer is made, or when a productivity shock changes the value of the surplus sufficiently. We consider first the impact of an outside offer.

### 2.4.1 Poaching

A key source of wage renegotiation and growth in this model is on-the-job search. This generates alternative opportunities for workers and either job mobility or responses to the outside offers with resulting wage increases. 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, i.e. $W_1 - W_0(x) < S(x, y') < S(x, y)$ (where $W_1$ denotes the present value of the current contract), then the worker uses the outside offer to negotiate up her wage. Lastly, 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.

In either one of the first two cases, the worker ends up in the higher surplus match, and uses the lower surplus match as the outside option when bargaining. Assume $S(x, y) \geq S(x, y')$. 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 plus a share of the incremental surplus between the two jobs, i.e.

$$\begin{aligned} W_1\left(\phi_1(x, y, y'), x, y\right) - W_0(x) &= S(x, y') + \beta\left[S(x, y) - S(x, y')\right] \\ &= \beta S(x, y) + (1 - \beta) S(x, y'). \end{aligned} \tag{5}$$

The share of the increased surplus $\beta$ accruing to the worker will be determined empirically.

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. We do, however, allow firms to fire a worker when the current surplus becomes negative and then immediately search for a replacement.

### 2.4.2 Productivity shocks

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 $w$ becomes $W_1(w, x, y')$. 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 the worker 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 the match viable and each partner better off within the match relative to unemployment for the worker and a vacancy for the job.

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

$$W_1(\psi_0(x, y'), x, y') - W_0(x) = 0, \tag{6}$$

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

$$W_1(\psi_1(x, y'), x, y') - W_0(x) = S(x, y'). \tag{7}$$

Wages respond to job specific productivity shocks, but not always in an obvious direction. 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 mismatch: what matters is what happens to the overall surplus. A positive productivity 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 to separate from the current worker and post a vacancy to find a better worker. Conversely a negative productivity shock can improve the surplus if this means the job type is now closer to the optimal one sought by the worker. 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, measured by whether it leads to an increase or a decrease, respectively of the surplus. Later, in Subsection 4.4 we consider further the non-monotonic relationship between wages and firm productivity, and the implications for empirical estimates of the degree of sorting of workers across jobs.

### 2.5 The meeting function

We relate the aggregate number of meetings between vacancies and searching workers through a standard aggregate matching function $M(\cdot, \cdot)$. This takes as inputs the total number of vacancies $V$ and the total amount of effective job seekers $s_0 U + s_1 (L - U)$. The matching function is assumed to be increasing in both arguments and exhibit constant returns to scale.

For the purpose of exposition it is useful to define

$$\kappa = \frac{M(s_0 U + s_1(L-U), V)}{[s_0 U + s_1(L-U)]\, V},$$

which summarizes the effect of market tightness in a single variable. In a stationary equilibrium $\kappa$ is constant, but it is not invariant to policy, and it is important to allow it to change when evaluating interventions or counterfactual regulations. Note that while $M(\cdot,\cdot)$ governs the aggregate number of meetings, whether or not a meeting translates into a match will be determined by the decisions of workers and firms.

The matching parameter $\kappa$ allows us to calculate all relevant meeting rates. The instantaneous rate at which an unemployed worker meets a vacancy of type $y$ is $s_0\kappa V \cdot \frac{v(y)}{V} = s_0\kappa v(y)$. The instantaneous probability for any vacancy to make a contact with an unemployed worker is $s_0\kappa u(x)$. Employed workers are poached by jobs of type $y$ with instantaneous probability $s_1\kappa v(y)$. Finally, the rate at which vacancies are contacted by a worker $x$ employed at a job $y$ is $s_1\kappa h(x,y)$.

## 2.6 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.

### 2.6.1 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 (home production) 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

$$rW_0(x) = b(x) + s_0\kappa\beta \int S(x,y)^+ v(y)\, dy, \tag{8}$$

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 contacts a job of type $y$ at a rate $s_0\kappa v(y)$ and the match is consummated if the surplus $S(x,y)$ is non negative, in which case she gets a share $\beta$ of the surplus.

### 2.6.2 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. 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 a vacancy for a job with productivity $y$ is

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

The $c$ is a per-period cost of keeping a vacancy open. The second term reflects the impact of a change in productivity from $y$ to $y'$, assuming that productivity shocks continue to accrue at rate $\delta$ when the job is vacant and assuming that the distribution of productivity shocks is uniform [0,1]. The third term is the expected gain from matching with a previously unemployed worker. The fourth term is the expected gain from poaching a worker who is already matched with another job. The notation $[\cdot]^{+}$ ensures that integration is over all possible $x$ and $y'$ that increase the current surplus.

We show in Appendix A that $\Pi_0'(y) > 0$ so that we can assume that the last job that enters makes zero profit, $\Pi_0(0) = 0$, that is

$$\begin{aligned} c = \delta \int \Pi_0(y') dy' + s_0\kappa\,(1-\beta) \int S(x,0)^{+} u(x)\, dx \\ + s_1\kappa\,(1-\beta) \int \left[S(x,0) - S(x,y')\right]^{+} h(x,y')\, dx\, dy'. \end{aligned} \tag{10}$$

This condition will determine the number of potential jobs $N$ in the economy. Firms enter the market and draw a $y$ in the uniform distribution, with $N$ determined by the least productive firm making zero profit.[6]

[6] Alternatively, we could (and we had in a previous version) assume that $N$ is exogenously given and large enough so there would exist a positive threshold $\underline{y}$ such that $\Pi_0(\underline{y}) = 0$. All vacant jobs with $\Pi_0(y) < 0$ would thus decide to remain "idle", not paying the cost of posting a vacancy until a shock moves $\Pi_0(y)$ above 0. However, this approach creates a nasty singularity at $\underline{y}$, arising because a match with $y < \underline{y}$ may still be viable if it arises after a negative productivity shock to an existing match as the vacancy costs have already been paid for this match.

### 2.6.3 Employed workers

In order to derive the wage rates we need to define the value of a job to a worker $W_1(w,x,y)$. This is the present value to the worker of a wage contract $w$ for a feasible match $(x,y)$ (if $S(x,y)<0$ there is no match and no wage). Any wage contract $w$ delivers $w$ in the first unit of time. The continuation value depends on competing events: the job may be terminated with probability $\xi$; the job may be hit by a productivity shock with probability $\delta$; the worker may be poached with probability $s_1\kappa V$.

Given the above discussion, the flow value of working at wage rate $w$ in a feasible match $(x,y)$ is determined by the following Bellman equation,

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

where

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

is the set of jobs that can lead to a wage change (either by moving or renegotiation) and $\mu(A)=\int_A\mu(y)\,dy$, for any set $A$ and any density $\mu$.

On the right hand side the first term is the wage net of the flow value of unemployment. The second term is the expected excess 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 (see equations (6) and (7)). The third line is the expected excess value following an outside offer as in equation (5). The integral is over all offers that can improve the value (whether the worker moves or not).

### 2.6.4 The match output and joint surplus

Having defined the non-employment value for the worker, the vacancy value for the firm, the value of a contract $w$ to the worker (with a similar expression for the value to the firm), we can now calculate the surplus value $S(x,y)$ defined by equation (3). We show in Appendix A that the match surplus $S(x,y)$ is effectively independent of the current wage contract and is defined by the fixed point in the following equation,

$$
\begin{aligned}
(r+\xi+\delta)\,S(x,y) &= f(x,y)-b(x)+c \\
&\quad-s_0\kappa\beta\int S(x,y')^+v(y')\,dy'-s_0\kappa\,(1-\beta)\int S(x,y)^+u(x)\,dx \\
&\quad-s_1\kappa\,(1-\beta)\iint\left[S(x,y)-S(x,y')\right]^+h(x,y')\,dx\,dy' \\
&\quad+s_1\kappa\beta\int\left[S(x,y')-S(x,y)\right]^+v(y')\,dy'+\delta\int S(x,y')^+\,dy'. \quad (12)
\end{aligned}
$$

Note that the surplus of an $(x, y)$ match never depends on the wage. 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 employee's outside option from both the value of unemployment and her current wage contract. 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 simultaneously computing the wage rates for workers.

## 2.7 Steady-state flow equations

The final set of equations needed to pin down the surplus and the distribution of matches are those defining equilibrium in this economy. 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.

The total number of matches in the economy is

$$L - U = N - V = \int h(x, y)\, dx\, dy. \tag{13}$$

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 $\xi$; second, with probability $\delta$, 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 forming a new match $(x, y')$; third, the worker may receive an offer and quit to a higher surplus $(x, y')$ match. 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. $S(x, y) \geq 0$),

$$\left[\xi + \delta + s_1 \kappa v(\overline{\mathcal{B}}(x, y))\right] h(x, y) = \left[s_0 u(x) + s_1 h(x, \mathcal{B}(x, y))\right] \kappa v(y) + \delta \int h(x, y')\, dy', \tag{14}$$

where

$$\begin{aligned}
\mathcal{B}(x, y) &= \{y' : 0 \leq S(x, y') < S(x, y)\}, \\
\overline{\mathcal{B}}(x, y) &= \{y' : S(x, y') \geq S(x, y)\},
\end{aligned}$$

are the sets of jobs that would imply ($\overline{\mathcal{B}}$) or not imply ($\mathcal{B}$) an improvement of the match surplus for $x$. Thus, $s_1 \kappa v(\overline{\mathcal{B}}(x, y))$ is the probability of receiving an alternative offer, when employed, from a job that beats the current one.

This equation defines the steady-state equilibrium, together with the following accounting equations for the workers:

$$u(x) = L - \int h(x,y)\,dy = L\frac{\xi + \delta\int \mathbf{1}\{S(x,y) < 0\}\,dy}{\xi + \delta\int \mathbf{1}\{S(x,y) < 0\}\,dy + s_0\kappa v\left(\mathcal{M}(x)\right)}, \tag{15}$$

where $\delta\int \mathbf{1}\{S(x,y) < 0\}\,dy$ is the endogenous job destruction rate and $s_0\kappa v(\mathcal{M}(x))$ is the job finding rate ($\mathcal{M}(x) = \{y : S(x,y) \geq 0\}$).

For the jobs we have

$$v(y) = N - \int h(x,y)\,dx. \tag{16}$$

The aggregate number of vacancies is then determined as $V = \int v(y)\,dy$. Note that one can either say that the free entry condition fixes the number of jobs $N$ and $V$ follows, or that it fixes the number of vacancies $V$ and $N$ follows as $V + L - U$.

### 2.8 Equilibrium

In equilibrium all agents follow their optimal strategy and the steady state flow equations defined above hold. The exogenous components of the model are the number of potential workers $L$, the form of the matching function $M(\cdot,\cdot)$ as well as the arrival rate of shocks $\delta$, the job destruction rate $\xi$, the search intensities for the unemployed $s_0$ and employed workers $s_1$, the discount rate $r$, the value of home production (or leisure) $b(x)$, the cost of posting a vacancy $c$, bargaining power $\beta$, 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. Appendix B provides a simple iterative algorithm that uses these equations and that of the surplus to compute the equilibrium objects.

## 3 Estimation Methodology

In this section, we first present the estimation procedure, then the data. Then we discuss the issue of measurement error on wages that comes with these data. We end this section by describing the parametric specification of the model and the moments used in estimation. An informal discussion of identification is provided in Section 3.5.

### 3.1 Estimation method

To estimate the model we use the Simulated Method of Moments (SMM) as constructing the likelihood function for this model is intractable. The SMM approach works as follows. The data sample allows us to calculate a vector of moments $\widehat{m}_N = \frac{1}{N}\sum_{i=1}^{N} m_i$, where for example, $\widehat{m}_N$ may be the 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 $\theta$ 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 $\widehat{\theta}$ maximizing

$$L_N(\theta) = -\frac{1}{2}\left(\widehat{m}_N - \widehat{m}_S^M(\theta)\right)^{\intercal} \widehat{W}_N^{-1}\left(\widehat{m}_N - \widehat{m}_S^M(\theta)\right),$$

where $\widehat{W}_N$ is taken to be the diagonal of the estimated covariance matrix of $\widehat{m}_N$. To compute the variance of $m_i$ in the actual sample we use the bootstrap.

Because the simulated moments are not necessarily a smooth function of the parameters we use a method developed by Chernozhukov and Hong (2003), which does not require derivatives of the criterion function. Chernozhukov and Hong do not maximize $L_N(\theta)$. Instead, they construct a Markov chain that converges to a stationary process of which the ergodic distribution has a mode that is asymptotically equivalent to the SMM estimator. Appendix C describes this procedure in detail.

### 3.2 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. It contains a nationally representative core 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. The model is estimated separately on each of these subgroups.

Individuals are interviewed once a year and provide retrospective information on their labor market transitions and their earnings. From this we construct histories at a 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, treating 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.[7] 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 three 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 month to month wage changes, we exclude observations where the wage changes by more than a factor

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

of two. This results in variances of wage changes at the annual level that are broadly consistent with PSID estimates (see e.g. Meghir and Pistaferri, 2004).

### 3.3 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.

### 3.4 Parametric 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

$$f(x,y) = f_0 \left( \left[ \exp\left( f_1 \Phi^{-1}(x) \right) \right]^{f_3} + \left[ \exp\left( f_2 \Phi^{-1}(y) \right) \right]^{f_3} \right)^{1/f_3},$$

where $\Phi$ is the standard normal CDF, $f_1$ is the standard deviation of $x$ and $f_2$ 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 $f_0$ determines the scale of wages. Finally the parameter $f_3$ determines the degree of substitutability between $x$ and $y$ with $1/(1-f_3)$ being the elasticity of substitution. When $f_3 = 1$ the production function becomes additive and there are no complementarities between the inputs.

The matching function requires macroeconomic data on vacancies and unemployment to be fully identified. We thus specify an elasticity of 0.5 for the meeting function:

$$M(U + s_1(L-U), V) = \alpha \sqrt{[U + s_1(L-U)]\, V},$$

where the search intensity for the unemployed has been normalized to one (Blanchard and Diamond, 1991, Petrongolo and Pissarides, 2001). 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.

### 3.5 Choice of moments

While using the likelihood function would have made use of all information in the data and the model specification, the use of the method of moments requires a careful choice of moments

Table 1: Moments Used in Estimation

| | |
|---|---|
| (a) Unemployment to Employment transitions (U2E) | $\mathbb{E}\left[\left(1-E_{i,t-1}\right)E_{i,t}\right]$ |
| (b) Employment to Unemployment transitions (E2U) | $\mathbb{E}\left[E_{i,t-1}\left(1-E_{i,t}\right)\right]$ |
| (c) Job-to-Job transitions (J2J) | $\mathbb{E}\left[E_{i,t-1}E_{i,t}J_{i,t}\right]$ |
| (d) Wage levels (zero if unemployed) | $\mathbb{E}\left[E_{i,t}w_{i,t}\right]$ |
| (e) Wage growth within job | $\mathbb{E}\left[E_{i,t-1}E_{i,t}\left(1-J_{i,t}\right)\Delta w_{i,t}\right]$ |
| (f) Wage growth at job change | $\mathbb{E}\left[E_{i,t-1}E_{i,t}J_{i,t}\Delta w_{i,t}\right]$ |
| (g) Variability of wages | $\mathbb{E}\left[E_{i,t}w_{i,t}^2\right]$ |
| (h) Variability of wage growth within job | $\mathbb{E}\left[E_{i,t-1}E_{i,t}\left(1-J_{i,t}\right)\Delta w_{i,t}^2\right]$ |
| (i) Variability of wage growth at job change | $\mathbb{E}\left[E_{i,t-1}E_{i,t}J_{i,t}\Delta w_{i,t}^2\right]$ |
| (j) Employment rate | $\mathbb{E}\left[E_{i,t}\right]$ |

to ensure that our model parameters are identified. We explain our choice in this subsection and we try to justify it by identification considerations. Given the complexity of the model and inference, we can only offer a heuristic argument for identification.

We use a set of unconditional moments, listed in Table 1. In particular, we use moments related to the transitions between employment states and between jobs; the level of wages and their cross sectional variance; and wage growth and its variance within and between jobs. Each of these moments is calculated separately by year in the labor force (year since leaving school for the cohort). Additionally we match the mean vacancy to unemployment rate (based on the mean and standard deviation from Hagedorn and Manovskii, 2008).

Before defining the moments, it is useful to define some notation. Let $E_{i,t}$ be the employment status of individual $i$ in period $t$ since entering the labor force, equal to one if employed and zero if unemployed. Let $J_{i,t}$ indicate that the worker is employed in $t$ and $t-1$, yet in different jobs. Finally, define $w_{it}$ as the log wage of individual $i$ in period $t$ and $E_{it}w_{it}$ is equal to zero if the worker is not employed at $t$. We plot the moments computed on the data, along with standard error bands and the simulated moments from the model in Figures 1, 2, and 3 below (we comment on these plots when discussing model fit in Section 4.1).

These moments allow us to identify the model parameters as follows. The transitions in and out of work and between jobs play a key role in identifying the job destruction rate $\xi$, the intensity of on-the job search $s_1$, and the overall level of matches $\alpha$. In particular $\alpha$ is identified by unemployment to employment transitions, $s_1$ is identified by the job changing rate (relative to transitions from unemployment) and $\xi$ is identified by the rate of separations into unemployment.

The variance of wages is directly linked to the variance of unobserved heterogeneity (or equivalently its weight in the production function) and hence both $f_1$ and $f_2$ are directly linked to the cross-sectional variance of wages. Through its effect on productivity shocks and job mobility, $y$ is also linked to the variability of wage growth within and between jobs. Hence the identification of the weight of $y$ in output, $f_2$, is also driven by the variance of wage growth, conditional on all types of transition. Similarly the within and between-job variance of wage growth is informative about the rate of arrival of productivity shocks $\delta$.

The remainder of the parameters are identified as following. The bargaining parameter $\beta$ is primarily identified by wages following a job change. So the key moment for this is wage growth at job change. The parameter $b$ in $b(x) = f(x, b)$, which reflects the income flow while out of work, is identified by the starting wage in the new jobs following unemployment. The flow cost of vacancy posting $c$ governs the profitability of posting vacancies, and is identified by matching the vacancy to unemployment ratio, which is observed from macroeconomic data.

We use the evolution of the cohort moments over time to identify the degree of complementarity. In the absence of search frictions, any degree of complementarity (supermodularity) will lead to perfect assortative matching (and no dynamics). With search frictions, there is a tradeoff between waiting for an improved match to take advantage of complementarities versus forming a potentially inferior match. The frictions preclude perfect sorting, but on-the-job search implies that workers will tend to move to better matches the longer they are employed. There is no explicit notion of age in our stationary model. We thus measure time in the labor market from the point when a cohort first enters unemployment. From then on an individual may have further non-work spells, but time continues to increase. For brevity we refer to this time as age. The resulting profile is then compared to the lifecycle evolution of the NLSY cohort from the point it completes full time education.

If we consider such a cohort of workers who all initially start out unemployed and follow them forward as they become employed and change jobs, the degree to which their wages spread out is directly related to the degree of sorting in equilibrium. Additionally, if sorting is an important feature of the labor market, the average gains to changing jobs will decline with time in the labor market, as the cohort of workers move toward their ideal matches. Similarly, the variance of wage changes should decline as there are fewer offers that improve matches substantially and workers make fewer transitions. Due to search frictions, the cohort of workers entering the labor market following a spell out of work will start off mismatched, but if there are production complementarities will become better and better matched through on-the-job search, and wages will spread out. Alternatively, if there are no (or weak) production complementariness then sorting is not a feature of the equilibrium, and wages will not spread out with time in the labor force as there is no tendency for the economy to reallocate them across jobs. Thus the degree of complementarity in production can be identified using the profile of wage growth and its variance between jobs as well as within jobs.[8]

Finally the measurement error variance and persistence ($\sigma^2$, $\rho$) are primarily identified by the cross sectional variance of wages and variance of wage growth. Specifically any variance and persistence not explained by the job mobility process will be interpreted as measurement error.

[8] The degree of complementarity will affect the variance of wages within jobs because this is in part due to responses to outside offers. For the same reasons described above the offers that will be able to move wages will decline in frequency with time in the labor market.

Table 2: Model Fit

| Moments Compared | Less than High School Model | Less than High School Data | High School Model | High School Data | College Model | College Data |
|---|---|---|---|---|---|---|
| Employment rate | 0.76 | 0.75 | 0.85 | 0.85 | 0.95 | 0.94 |
| Change jobs given employment | 0.03 | 0.04 | 0.019 | 0.024 | 0.009 | 0.013 |
| Within Job Wage Growth | 0.001 | 0.0026 | 0.0024 | 0.0025 | 0.0021 | 0.0021 |
| Var of within job wage growth | 0.0042 | 0.0049 | 0.0033 | 0.0038 | 0.003 | 0.003 |
| Between Job Wage Growth | 0.083 | 0.082 | 0.14 | 0.080 | 0.24 | 0.10 |
| Var of between job wage growth | 0.057 | 0.064 | 0.040 | 0.062 | 0.062 | 0.058 |

Note: Unit of time is a month. Wage growth is conditional on employment in both periods.

# 4 Estimation Results

The estimation results will be presented in five steps. First, we discuss how well the model fits the moments. Then, we present parameter estimates. In the third subsection, we discuss the implications for wage dynamics. The last two sections deal with sorting. First, we characterize the implications of the model for sorting with respect to unobservable worker and firm characteristics $x$ and $y$. Second, we reconsider the information on sorting contained in the correlation between worker and firm fixed effects estimated from matched employer-employee data.

## 4.1 The fit of the moments

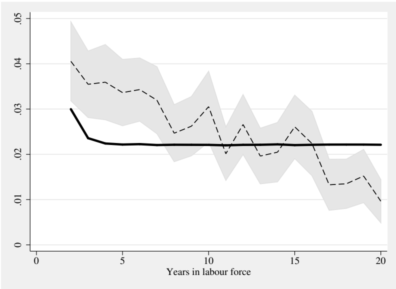
In Table 2 we present average (over age) unconditional moments related to employment, job changes, and wage growth, both within and between jobs. In general the model does well in fitting the average value of these moments, the only exception being between job wage growth, which is overestimated for the two higher educated groups.

In Figures 1, 2 and 3 we compare the full age profiles of the simulated and actual moments, along with plus and minus two standard errors of the data moment. Moments with narrower the standard error bands have larger weight in estimation.[9] Fitting the evolution of the moments poses a greater challenge for the model. By using the age profiles we require the model to fit the individual level dynamics.

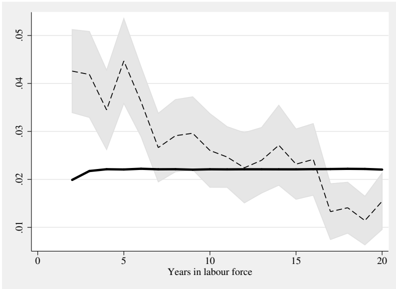
The degree to which the model fits the entire age profile varies from moment to moment and across the education groups, but the model produces qualitatively the correct profiles. It is reasonably successful at fitting job to job transitions, whose pattern may reflect the presence of assortative matching. The decline in the rate of job to job transitions for the two higher education groups and the relative constant rate for the lower skill group reflects the importance of complementarities for the two top groups: their rate of job changing declines as they approach the better suited matches. However for the lower skill group transition rates remain flat.

While the model is reasonably successful at fitting the age profiles for college educated workers, it does less well in tracking the trends for the two lower educated groups of workers, especially the the age profiles of employment transitions. It underestimates the continual decline

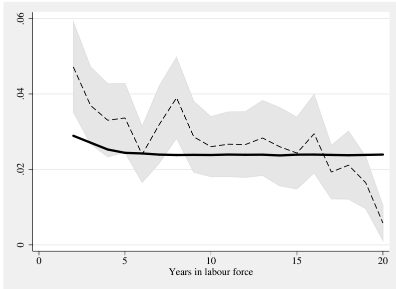
[9]This comparison does not reflect estimation error and hence provides a narrower confidence interval for evaluating the difference between the model and the data moment.



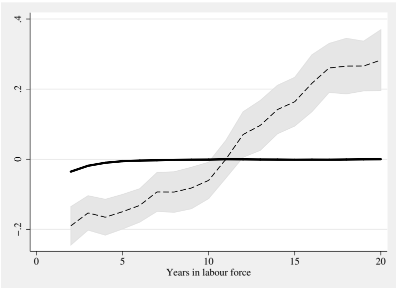

(a) $\mathbb{E}\left[\left(1-E_{i,t-1}\right)E_{i,t}\right]$



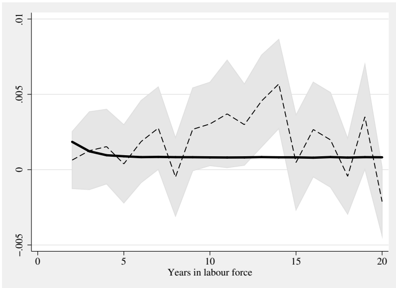

(b) $\mathbb{E}\left[E_{i,t-1}\left(1-E_{i,t}\right)\right]$



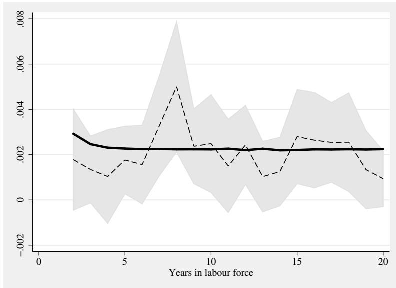

(c) $\mathbb{E}\left[E_{i,t-1}E_{i,t}J_{i,t}\right]$



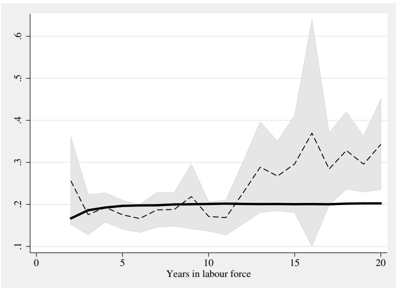

(d) $\mathbb{E}\left[E_{i,t}w_{i,t}\right]$



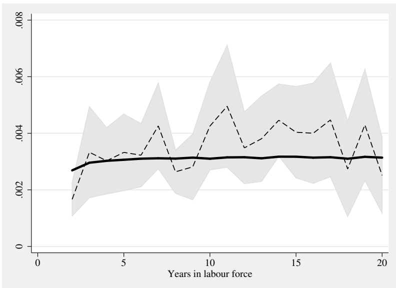

(e) $\mathbb{E}\left[E_{i,t-1}E_{i,t}\left(1-J_{i,t}\right)\Delta w_{i,t}\right]$



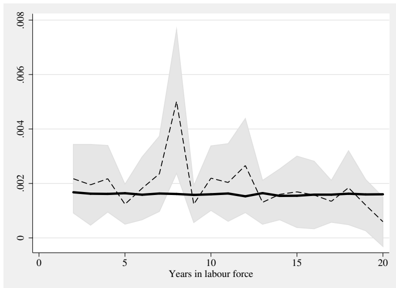

(f) $\mathbb{E}\left[E_{i,t-1}E_{i,t}J_{i,t}\Delta w_{i,t}\right]$



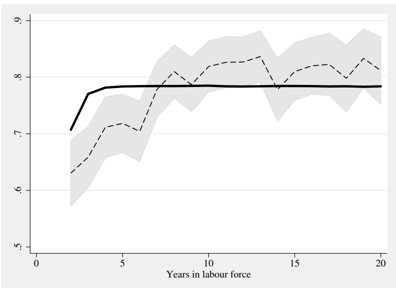

(g) $\mathbb{E}\left[E_{i,t}w_{i,t}^2\right]$




(h) $\mathbb{E}\left[E_{i,t-1}E_{i,t}\left(1-J_{i,t}\right)\Delta w_{i,t}^2\right]$

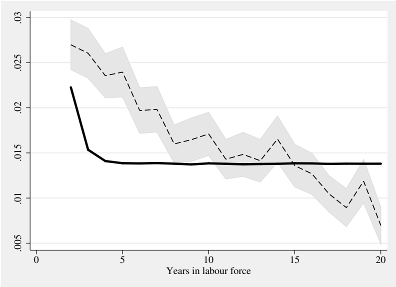



(i) $\mathbb{E}\left[E_{i,t-1}E_{i,t}J_{i,t}\Delta w_{i,t}^2\right]$

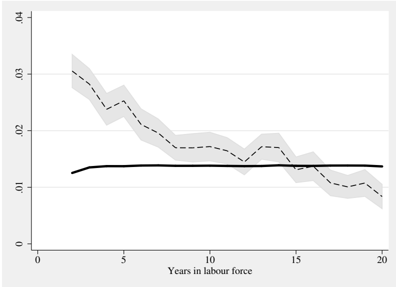



(j) $\mathbb{E}\left[E_{i,t}\right]$

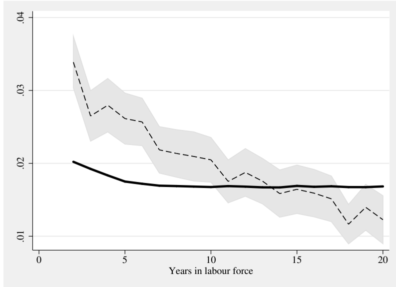
Figure 1: Model fit: Less than high school



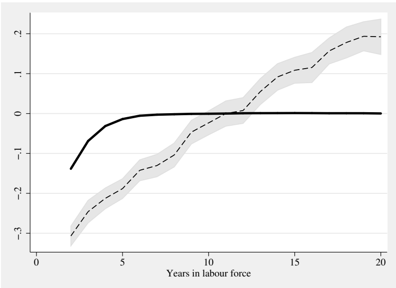

(a) $\mathbb{E}\left[\left(1-E_{i,t-1}\right)E_{i,t}\right]$



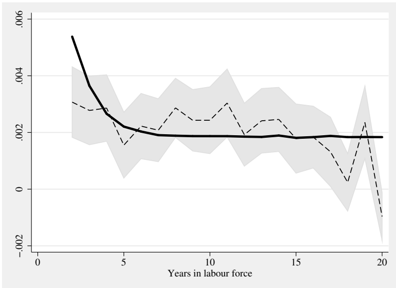

(b) $\mathbb{E}\left[E_{i,t-1}\left(1-E_{i,t}\right)\right]$



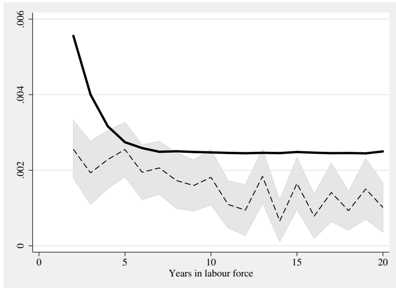

(c) $\mathbb{E}\left[E_{i,t-1}E_{i,t}J_{i,t}\right]$



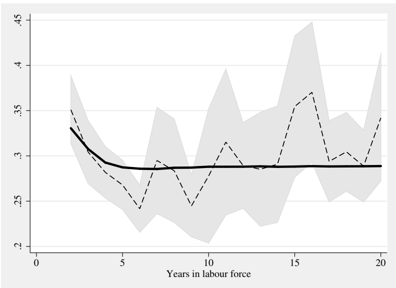

(d) $\mathbb{E}\left[E_{i,t}w_{i,t}\right]$



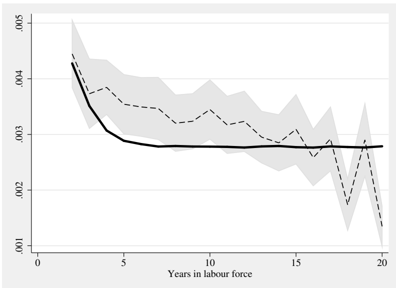

(e) $\mathbb{E}\left[E_{i,t-1}E_{i,t}\left(1-J_{i,t}\right)\Delta w_{i,t}\right]$



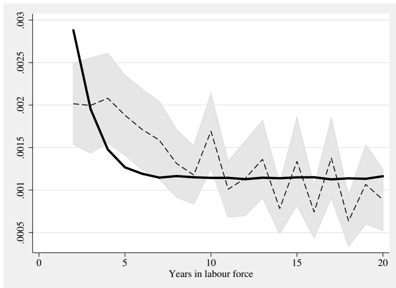

(f) $\mathbb{E}\left[E_{i,t-1}E_{i,t}J_{i,t}\Delta w_{i,t}\right]$



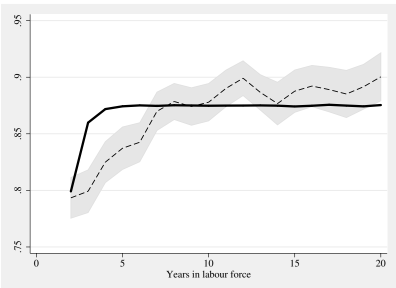

(g) $\mathbb{E}\left[E_{i,t}w_{i,t}^2\right]$




(h) $\mathbb{E}\left[E_{i,t-1}E_{i,t}\left(1-J_{i,t}\right)\Delta w_{i,t}^2\right]$

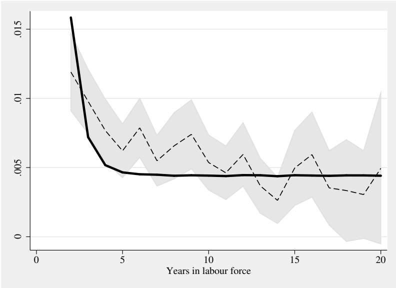



(i) $\mathbb{E}\left[E_{i,t-1}E_{i,t}J_{i,t}\Delta w_{i,t}^2\right]$

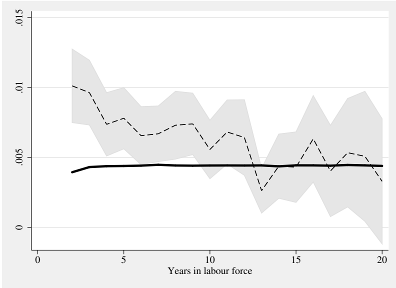



(j) $\mathbb{E}\left[E_{i,t}\right]$

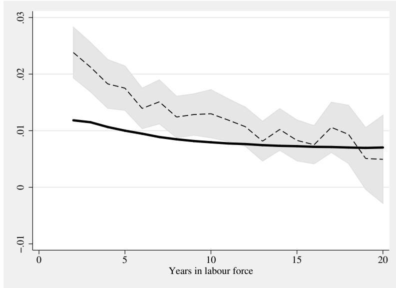
Figure 2: Model fit: High school graduate



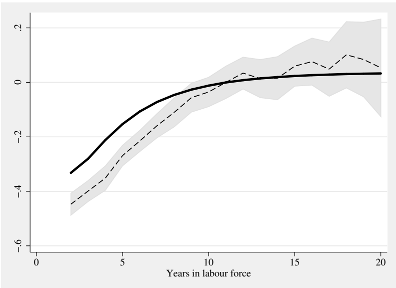

(a) $\mathbb{E}\left[\left(1-E_{i,t-1}\right)E_{i,t}\right]$



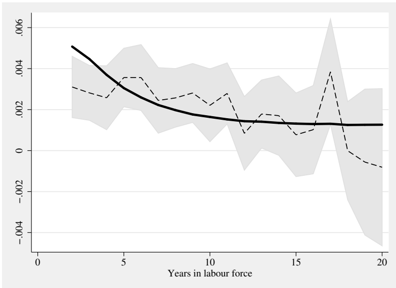

(b) $\mathbb{E}\left[E_{i,t-1}\left(1-E_{i,t}\right)\right]$



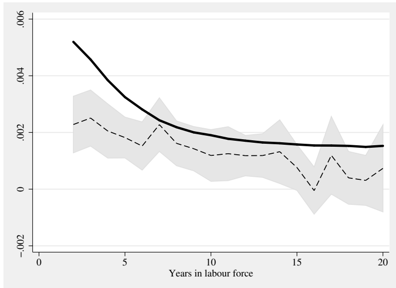

(c) $\mathbb{E}\left[E_{i,t-1}E_{i,t}J_{i,t}\right]$



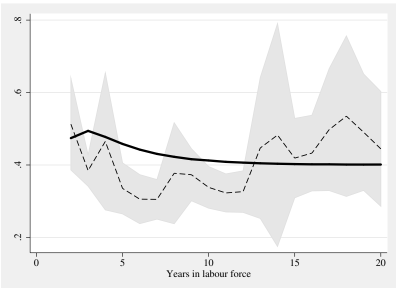

(d) $\mathbb{E}\left[E_{i,t}w_{i,t}\right]$



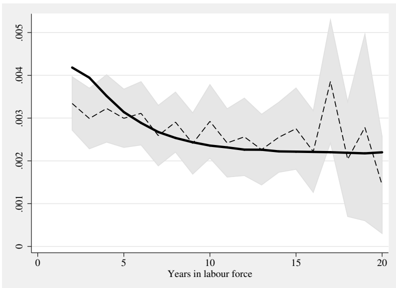

(e) $\mathbb{E}\left[E_{i,t-1}E_{i,t}\left(1-J_{i,t}\right)\Delta w_{i,t}\right]$



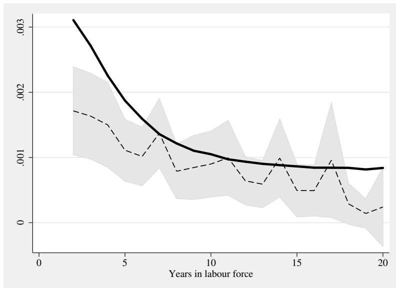

(f) $\mathbb{E}\left[E_{i,t-1}E_{i,t}J_{i,t}\Delta w_{i,t}\right]$



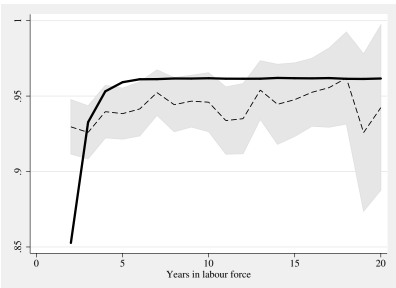

(g) $\mathbb{E}\left[E_{i,t}w_{i,t}^2\right]$




(h) $\mathbb{E}\left[E_{i,t-1}E_{i,t}\left(1-J_{i,t}\right)\Delta w_{i,t}^2\right]$




(i) $\mathbb{E}\left[E_{i,t-1}E_{i,t}J_{i,t}\Delta w_{i,t}^2\right]$




(j) $\mathbb{E}\left[E_{i,t}\right]$

Figure 3: Model fit: College graduate

Table 3: Autocovariances (autocorrelations) of within-job wage growth

| | less than HS | | HS grad | | college grad | |
|---|---|---|---|---|---|---|
| Order | Cov | Corr | Cov | Corr | Cov | Corr |
| 0 | 0.429 | 1.000 | 0.246 | 1.000 | 0.089 | 1.000 |
| 1 | -0.140 | -0.335 | -0.075 | -0.312 | -0.020 | -0.225 |
| 2 | -0.052 | -0.123 | -0.032 | -0.135 | -0.011 | -0.123 |
| 3 | -0.004 | -0.010 | -0.005 | -0.020 | -0.003 | -0.034 |
| 4 | -0.004 | -0.009 | -0.003 | -0.011 | -0.002 | -0.025 |
| 5 | -0.001 | -0.002 | -0.001 | -0.004 | -0.002 | -0.020 |
| 6 | -0.002 | -0.005 | 0.000 | -0.001 | -0.001 | -0.016 |
| 7 | 0.000 | -0.001 | 0.000 | -0.002 | -0.001 | -0.011 |
| 8 | -0.001 | -0.002 | 0.000 | -0.001 | -0.001 | -0.009 |
| 9 | -0.001 | -0.002 | -0.001 | -0.003 | 0.000 | -0.006 |
| 10 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | -0.004 |

Note: Autocovariances (autocorrelations) of the change in residual annual income, with zero mean by construction. The simulated moments present the reduced form dynamics of annual income implied by the model (including measurement error).

in the probability of being in different employment states in two consecutive periods, becoming relatively flat beyond the fifth year. This is true not only for the job to job transitions but also for the unemployment to employment and the employment to unemployment transitions (subplots a, b, c). Similarly, the model misses the continuous positive trend in earnings (subplot d) for the two low educated groups, again flattening out after the first five year. The age profiles of the moments related to wage growth and the variability of wage growth, however, are better reproduced.

One obvious reason for the difficulty in capturing the lifecycle profiles is that we have not allowed for human capital accumulation in the model. Hence all wage growth is due to the search process and the associated counteroffers and wage mobility. A possible explanation for the fact that the fit is better for college graduates may be that on-the-job learning is less important for this group relative to search induced wage growth; that is, for college graduates, wage growth is better explained by on-the-job search and the associated offers and counteroffers that it generates.[10]

## 4.2 Fit of individual wage dynamics

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) and 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.

Because many of these shocks lead to permanent or in any case long-lasting wage gains or

[10]Incorporating human capital accumulation and stochastic shocks to individual ability into this model is clearly outside the scope of this paper but is a part of our research agenda and ultimately an important element of a labor market model for earnings.

losses, they may well lead to the dynamics we observe in the data. To see whether this dynamic structure leads to similar patterns as those we see in the data Table 3 presents the implied auto-covariance structure, $\mathrm{Cov}(\Delta u_t, \Delta u_{t-s})$, where $u_t$ is log annual earnings from the model with measurement error added to match the observable data. We also remove the effects of age from growth, before constructing the autocovariances. Very much like in the results from the actual data we find that autocovariances of order three or above are very low and effectively zero. Moreover, assuming $u_t$ to be the sum of a random walk plus an MA(1) shock, with stationary innovations, implies a permanent shock whose variance is 0.045, 0.032 and 0.027 for the unskilled, the high school graduates and the college graduates respectively.[11] These numbers are similar in magnitude to those calculated from the PSID (different cohorts, different time period) by Meghir and Pistaferri (2004), who find 0.03, 0.03 and 0.04, respectively. This confirms Postel-Vinay and Turon's observation that the observed autocovariances can be generated by the behavior of a search model such as ours.

### 4.3 Parameter estimates

The parameters implied by our estimation are presented in Table 4. Turning first to the production function we see that for the least skilled workers $f_3 = 0.72$, implying approximately an additive production function (elasticity of substitution equal to 3.6) and no real gains from sorting in that labor market. Interestingly, because of search frictions there are still gains from on-the-job search. Indeed, incumbent firms always try to retain employees subject to poaching by increasing their wages. Note that poaching may still be successful even without complementarity in the production function. This is because the surplus function is in this case roughly constant with respect to $x$ but increasing in $y$. This happens because $b(x) = f(x, b)$ increases with $x$ like $f(x, y)$, whereas the cost of a vacancy is independent of $y$ (see equation 12). For the higher skill groups the substitution elasticities are (in ascending order of skill) 0.70 and 0.62 respectively: in both cases there are strong complementarities and gains from sorting. We look at the implications below.

Turning now to the remaining parameters of the production function it is interesting to note that for the college group job heterogeneity is much more important for production than individual heterogeneity as shown by the relative size of $f_1$ and $f_2$. These parameters reflect both the extent of heterogeneity among workers and jobs as well as the relative importance of such characteristics in production. For the lowest skill group the contribution of jobs and individual characteristics are about the same, while for the high school graduates individual characteristics are more important, but job heterogeneity still plays an important role. For college graduates however, job heterogeneity seems to be the dominant factor by far in determining output, relative to individual heterogeneity $x$.

The exogenous monthly job destruction rate $\xi$ is higher for the lower skill groups and implies exogenous separations every 3, 8, and 21 years for the high school dropouts, high school graduates and college graduates respectively. Endogenous separation can occur following a productivity

[11] We estimate the permanent variance as $\mathrm{Var}(\Delta u) + 2\,\mathrm{Cov}(\Delta u_t \Delta u_{t-1}) + 2\,\mathrm{Cov}(\Delta u_t \Delta u_{t-2})$.

Table 4: Parameter Estimates

| | | Less than HS | High school | College |
|---|---|---|---|---|
| Production function parameters | $f_1$ | 3.6423 | 2.5111 | 0.4478 |
| | | (0.1916) | (0.0488) | (0.0246) |
| | $f_2$ | 4.3058 | 1.6655 | 5.1736 |
| | | (0.2340) | (0.0775) | (0.2972) |
| | $f_3$ | 0.7203 | -0.4252 | -0.6125 |
| | | (0.0874) | (0.0421) | (0.0805) |
| Elasticity of substitution | $(1-f_3)^{-1}$ | 3.576 | 0.702 | 0.620 |
| Matching function parameters | $\alpha$ | 0.2725 | 0.5597 | 0.4865 |
| | | (0.0162) | (0.0161) | (0.0144) |
| Search intensity | $s_1/s_0$ | 1.1994 | 0.5193 | 0.2499 |
| | | (0.1065) | (0.0251) | (0.0142) |
| Probability of exogenous job destruction | $\xi$ | 0.0280 | 0.0099 | 0.0039 |
| | | (0.0008) | (0.0005) | (0.0001) |
| Probability of shock to $y$ | $\delta$ | 0.0086 | 0.0178 | 0.0028 |
| | | (0.0037) | (0.0014) | (0.0002) |
| Home production parameter | $b$ | 0.3975 | 0.7102 | 0.1245 |
| | | (0.1750) | (0.0070) | (0.0490) |
| Vacancy posting cost | $c$ | 0.6701 | 5.0198 | 4.3047 |
| | | (0.0756) | (0.0685) | (0.0769) |
| Worker bargaining power | $\beta$ | 0.6690 | 0.0306 | 0.0721 |
| | | (0.0340) | (0.0030) | (0.0030) |
| SD of measurement error | $\sigma$ | 0.0490 | 0.0380 | 0.0323 |
| | | (0.0036) | (0.0011) | (0.0019) |
| Autocorrelation of measurement error | $\rho$ | 0.5487 | 0.6403 | 0.5413 |
| | | (0.1822) | (0.0857) | (0.2012) |

Note: standard errors in parenthesis. Rates are per month.

Table 5: Implications for wage growth and job mobility

| | Less than high school | High school graduate | College graduate |
|---|---|---|---|
| Probability of endogenous job destruction | 0.000 | 0.0061 | 0.0006 |
| Mean separation rate conditional on shock to $y$ | 0.000 | 0.342 | 0.207 |
| Probability of job contact when unemployed | 0.102 | 0.223 | 0.182 |
| Mean matching rate given a contact | 1.000 | 0.461 | 0.648 |
| Probability of job contact when employed | 0.123 | 0.116 | 0.045 |
| Mean matching rate given a contact | 0.256 | 0.161 | 0.161 |
| Share of total wage changes occurring | | | |
| at the same job | 0.4082 | 0.5372 | 0.5064 |
| at a job change | 0.5918 | 0.4628 | 0.4936 |
| Share of wage changes at same job due to | | | |
| shocks | 0.2093 | 0.2508 | 0.1764 |
| counter offers | 0.7907 | 0.7492 | 0.8236 |
| Fraction of wage changes that are positive | | | |
| at the same job | 0.7934 | 0.7558 | 0.828 |
| at a job change | 0.7084 | 0.4788 | 0.587 |

Note: Rates are per month.

shock if the surplus becomes negative. This can occur either because of a resulting mismatch (when sorting is important) or because the productivity shock is so negative that the job cannot pay the worker their non-work value and still cover it own outside option. However, given the CES production function, $f(x, b) < f(x, y)$ if and only $b < y$ whatever $x$: the second source of sorting is therefore not expected to be important. In Table 5 we have calculated the probability of an endogenous separation and the separation rate after a productivity shock. For the unskilled, since sorting plays no role, we are not surprised to find that separations into unemployment can only occur through exogenous job destruction. The role of such exogenous job destruction is much reduced for the higher skill groups both because the overall transition into unemployment is lower and because endogenous separations are more important at least for the high school graduates: based on the estimate of $\delta$, the jobs they work for receive a shock every 4.7 years and in 34% of cases this leads to a transition into unemployment. Hence they are hit by an endogenous job destruction shock every 13.7 years. For the low skill and the the college groups the shocks are rare, which means they are not an important source of either transitions or of wage renegotiations.

The flow value of home time (home production) $b(x)$ is estimated to be equivalent to production with a firm between the 12th and 71st percentile, with no monotonic relation to education group. The monthly cost of posting a vacancy is much lower for recruiting unskilled workers than for the other two groups. The other key parameter for our model is the bargaining parameter $\beta$, with the results implying that unskilled workers obtain 67% of the surplus, while the two higher skill groups only obtain 3.1% and 7.25% respectively. In other words their pay is determined mainly from the sequential auction process. There is no evidence that an additional rent sharing mechanism, besides Bertrand competition, is required to explain wage determination for skilled workers.

The parameter estimates for the search technology ($\alpha$ and $s_1/s_0$) are better understood by calculating the probabilities of job contact when employed and unemployed, which are displayed in Table 5. The contact rates for the unskilled when unemployed are substantially lower (10.2% a month) than for the two higher skill groups (22% and 18%). However the unskilled accept all offers when unemployed. Both the contact rate when working and the rate at which alternative offers are accepted declines with skill.

The sources of wage dynamics are shocks to productivity $y$ that trigger wage renegotiation, and outside offers that lead to a counteroffer and/or job-to-job mobility. Table 5 decomposes these different sources over and above measurement error. For the unskilled 59% of wage changes happen at job change. For the high school and college graduates these are split roughly evenly. For all groups the predominant reason for a wage change at the same job is in response to an outside offer: 79%, 75% and 82% percent for the low, medium and high educated groups respectively.

The last two parameters relate to measurement error. The estimated standard deviation of monthly measurement error is low for all groups, between 0.03 and 0.05. The correlation between any two months within a reporting year is estimated to be between 0.54 and 0.64, although without much precision.

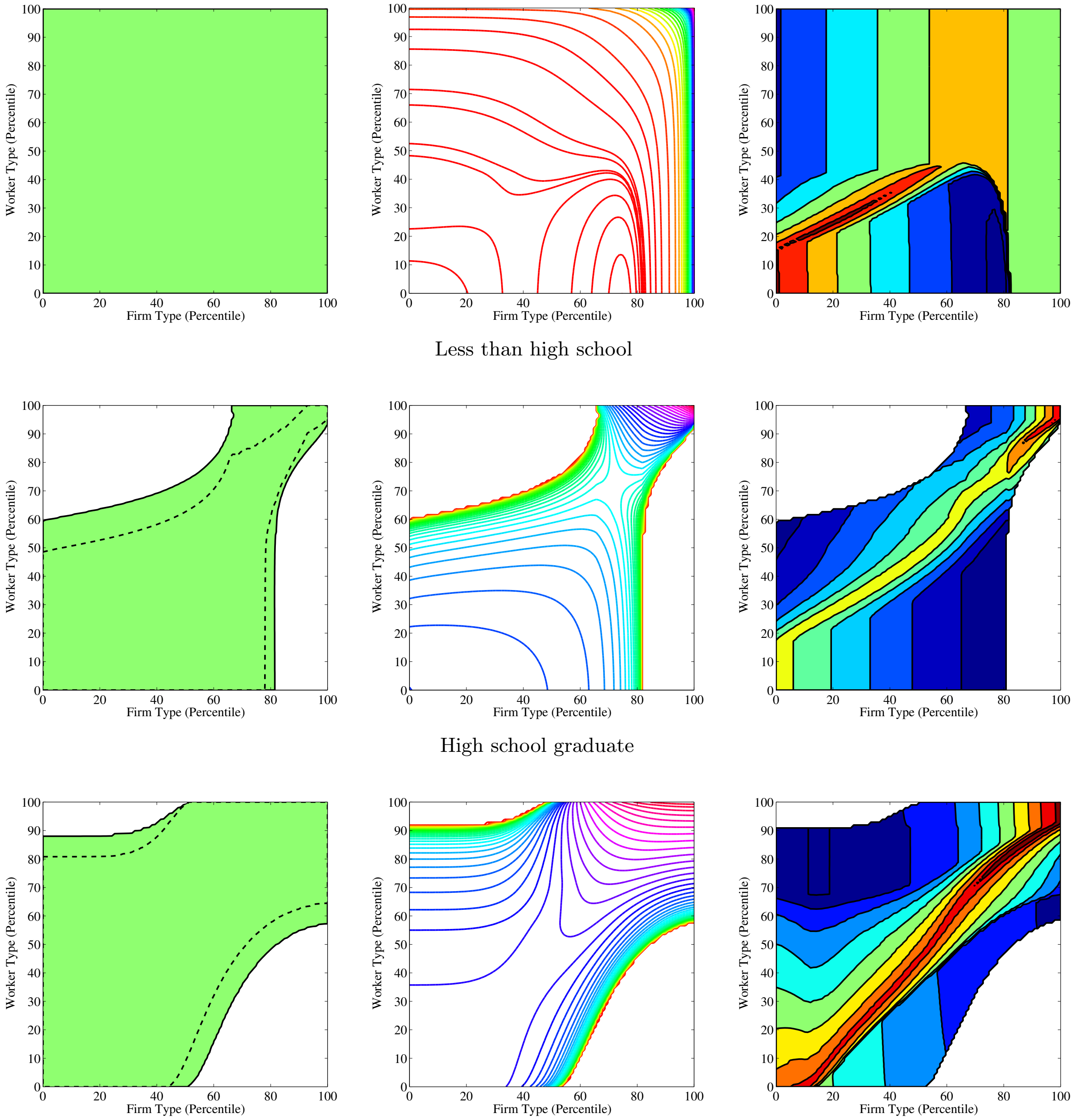
### 4.4 Sorting with respect to unobserved characteristics

The key feature of our model is that it allows for the possibility of sorting, which has important implications for the cost of mismatch. We have already demonstrated that our estimates imply production complementarities at least for the two higher skill groups. This in itself would imply perfect assortative matching in a first best world without search frictions. We now show the extent of actual sorting in steady state, which in turn allows us to quantify an important welfare loss due to mismatch in the presence of search frictions.

In Figure 4 we illustrate several aspects of the equilibrium allocation of workers to jobs. In the the left hand panels (Figure 4.a) we plot the set of possible matches for the three skill groups: $\{x, y|S(x, y) \geq 0\}$. In the unskilled market any firm will match with any worker: the small gains from matching with someone of a similar rank (implied by the high substitution elasticity of 3.6) are not enough to make it worthwhile remaining unemployed (or vacant) and waiting for a better match. However, in the higher skill groups, matches between high type workers and low type firms and between low type workers and high type firms never occur, which is a manifestation of the production complementarities that induce sorting: when worker and job pairs in that region meet they prefer to wait. Of course, even if all matches are possible there can be positive sorting if the surplus is not strictly increasing in $x$ and $y$, leading to the density of matches concentrated on the diagonal from the origin to the top right through the process of on-the-job-search. We now look at sorting in greater detail.

In the centre panels (Figure 4.b) we plot the contour lines of the surplus function $S(x, y)$. Turning first to the college graduates, we can clearly see the effect that on-the-job search will have on sorting. For example, consider a college educated worker at the 60th percentile. The surplus created with any job is positive for this worker, however, the surplus initially increases in the type of the job, is maximized when matched to a job at the 60th percentile, and then declines again. This worker will initially match with any job, but will always move to a job that is closer to the 60th percentile, which may involve moving up or down the quantiles of $y$. The surplus functions for the two higher educated groups are both highest close to the 45 degree line: both workers and firms would prefer match with a partner of a similar type. The surplus function for the low skilled (panel (a)) is quite different. For this group the surplus is always highest when a worker (job) is matched with the best job (worker). At the same time, there is small amount of non-monotonicity in the surplus when considering matches with workers below the median. In this range, the small degree of complementarity in production induces a local maximum near the 45 degree line, which is sufficient to induce workers below the median to move to lower $y$ matches in this region, although they would always move to the very top jobs when the opportunity arises. Low educated workers above the median always prefer to move to higher ranked jobs.

In the right hand panels (Figure 4.c) we illustrate the implication of complementarity combined with on-the-job search by plotting the contours of the joint distribution of matches, $h(x, y)$.




(a) Matching set (b) Surplus function (c) Distribution of matches

Figure 4: Equilibrium matching sets, surplus function and distribution of matches

Note: In panels (a) we plot the set of matches that are formed between unemployed workers and vacant jobs: $\{x, y | S(s, y) \geq 0\}$. The green shaded area represents the decentralized matching set, the dashed lines are the boundaries of the constrained social planner's matching set, approximated by $\mathcal{M}^{SP}(x, y) = \left\{x, y | x > \sum_{i=1}^{I} \tau_i y^{i-1}, y > \sum_{j=1}^{J} \tau_j x^{j-1}\right\}$. In panels (b) we plot the contours of the Surplus function $S(x, y)$. In panels (c) we plot the distribution of matches $h(x, y)$.

Looking first at the two higher educated group of workers, see that the distribution of matches is highest close to the 45 degree line. Indeed, although workers initially form matches according to the large matching set presented in panel (a), they leave these low surplus matches for higher surplus matches with jobs at a similar quantile whenever the opportunity arises. In the stationary equilibrium there are very few matches that are very far from the diagonal.

The distribution of matches for the lowest education group is a little more involved. There is some sorting among the below median matches since the production function is not quite additive and complementarities outweigh individual contributions in this range. The local maximum in the surplus function that occurs for workers below the median implies that workers prefer to move to a job of a similar quantile to their own ability, unless it is substantially higher (above the 75th–80th percentile). For workers above the median they always prefer to move to a higher ranked job when given the opportunity. The result is that there is a high concentration of matches along the 45 degree line below the the median. Above the median things look different. Within a job-type, the distribution of workers above the median is uniform (firms accept all workers and these workers are all moving in and out of job-types at the same rate). Additionally, for firms above the 80th percentile, all workers are equally represented, as they all agree on the desirability of these matches. Finally, it is interesting to note that for workers above the median, their distribution across firm types in humped shape, with the mode around the 75th percentile. This hump shape is the result of the effect that the job-changing decisions of the workers below the median has on the equilibrium distribution of vacancies. Since workers below the median will leave their current job for either a job closer to their own type, or a job above the 80th percentile, there is additional competition for vacancies above the 80th percentile from the point of view of above median workers.

In conclusion, there is a lot of sorting in the economy and plenty of mismatch, which leads to important welfare losses relative to a world with no search frictions or no congestion externalities. We characterize these losses in Section 5.

### 4.5 Measuring sorting with wage data only

The availability of matched employer employee data has motivated empirical research on the distribution of wages across workers and firms. Indeed the availability of such data allows to identify worker and firm fixed effects in a standard Mincer wage equation. Abowd, Kramarz, and Margolis (1999) and a few other studies (see the introduction for references) implemented this idea and concluded that there is very little evidence of sorting, based on the small estimated correlation between worker and firm effects. However, this approach to measuring sorting was recently disputed first theoretically (Eeckhout and Kircher, 2011) and then empirically (Lopes de Melo, 2009, Bagger and Lentz, 2012 and Hagedorn, Law, and Manovskii, 2012).

The key difficulty lies in the fact that in the presence of productive complementarities there is a non-monotonic relationship, in general, between match surplus and the firm type (worker type). Since the maximum possible wage a worker can earn in a match is directly controlled by the match surplus, this implies that there will also be a non-monotonic relationship between wages and the firm type (worker type) when production complementarities are important. This theoretical point was made by Eeckhout and Kircher (2011) and illustrated using a simple search model. In the model we estimate, there are additional forces at work that may amplify or mute this effect, such as on-the-job search, shocks to the productivity of the job, and the fact that wages reflect both the surplus in the current match and the surplus at the worker's outside option.

To illustrate the relationship between wages, surplus, and firm types we calculate the range of potential wage paid to a worker in each possible $(x, y)$ match. The highest possible wage is the wage that gives the entire match surplus to the worker and is implicitly defined by

$$W_1(\overline{w}, x, y) - W_0(x) = S(x, y). \tag{17}$$

The lowest possible wage paid to a worker in an $(x, y)$ match gives the worker a share $\beta$ of the match surplus and is implicitly defined by

$$W_1(\underline{w}, x, y) - W_0(x) = \beta S(x, y). \tag{18}$$

In Figure 5 we present information on wage setting by worker and firm characteristics. We plot, separately by decile of the worker type distribution, the maximum wage against the match surplus (left hand panels) and against the firm type in the match (right hand panel).[12] While maximum wages are strictly increasing in match surplus for all education groups, the non-monotonicity of wages by either firm-type or worker-type is clear for the two higher eduction groups.

It is then clear that any empirical strategy designed to measure the degree of sorting of workers across firms that relies on a monotonic relation between wages and firm types is likely to do poorly in this environment. As an illustration, we apply the fixed effects estimation strategy of Abowd, Kramarz, and Margolis (1999) to data simulated from our model at our estimated parameters. Specifically, we simulate panel data and for each worker we keep track of their wage, their id, and the id of the firm they are employed at. To construct a firm id we assign them the firm type at the start of the match. We estimate the following wage equation:

$$w_{ijt} = \phi_i + \psi_{j(i,t)} + u_{ijt}, \tag{19}$$

where the subscript $j(i, t)$ denotes the firm $j$ in which individual $(i)$ is working in period $t$. The reduced form sorting measure is the the correlation between the estimated worker fixed effect $\phi_i$ and the firm fixed effect $\psi_{j(i,t)}$ of the firms at which they worked. This can be compared to the actual correlation implied by solving the model

Table 6 presents results from this comparison. The correlations between worker and firm

[12]To keep the figures as clear as possible we do not plot the minimum possible wage. For for two highest eduction groups $\beta$ is estimated to be close to zero, implying that the lowest possible wage is approximately constant across all firms for a given worker type, and equal to the maximum possible wage at the match where $S(x, y) = 0$. For the lowest education group the lowest possible wage is approximately a shift down from the maximum possible wage.

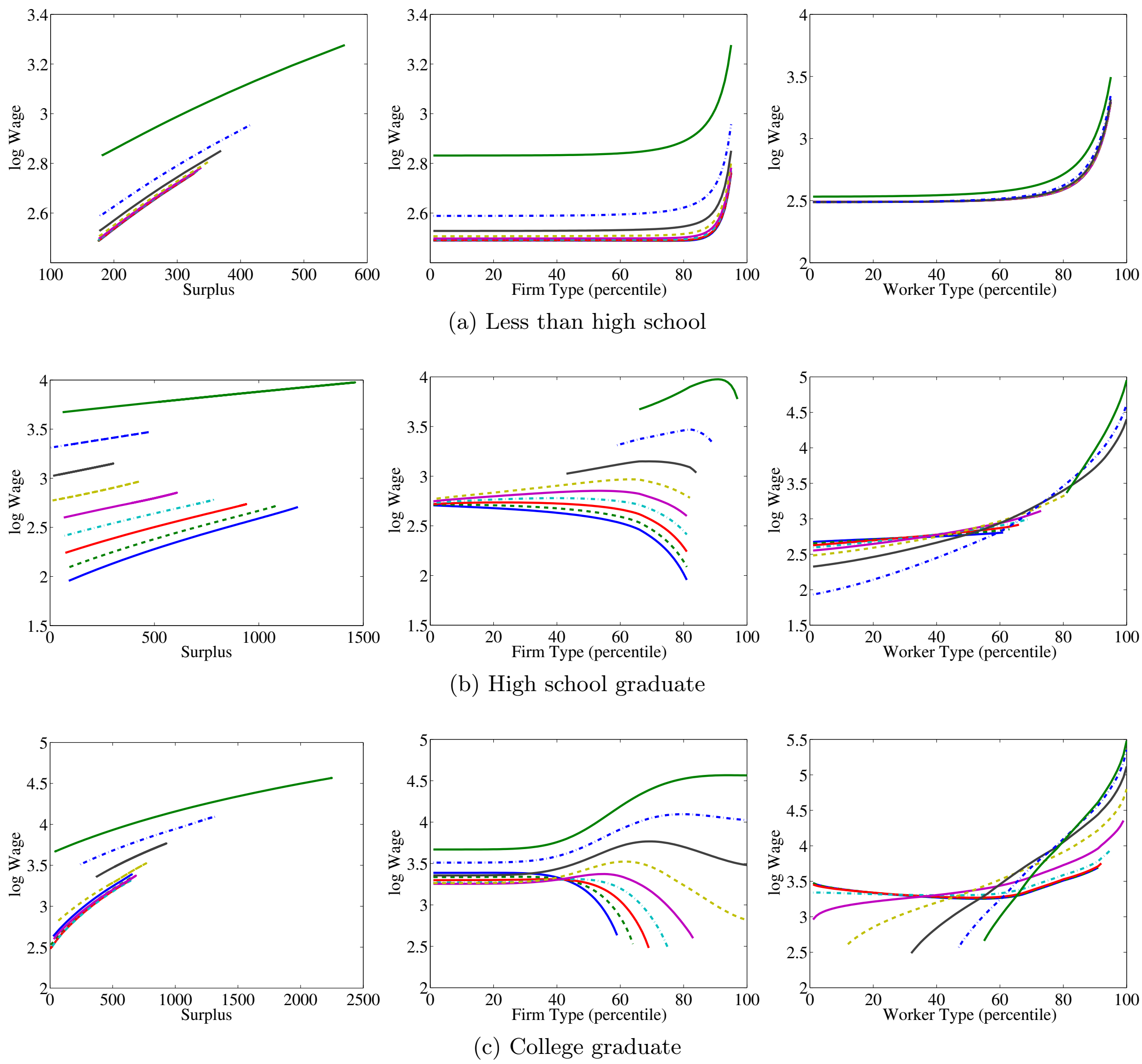



Figure 5: Maximum possible wage

Note: We plot the maximum possible wage in a match by: decile of worker type against match surplus (left hand panels); decile of worker type against firm type (centre panels); and decile of firm type against worker type (right hand panels).

Table 6: Actual and estimated sorting patterns

| | Corr$(x, y)$ | Corr$(\hat{\phi}_i, \hat{\psi}_j)$ | |
|---|---|---|---|
| | | (i) | (ii) |
| Less than high school | 0.338 | -0.008 | 0.007 |
| High school graduate | 0.846 | -0.236 | -0.248 |
| College graduate | 0.858 | 0.644 | 0.315 |

Note: The column labeled Corr$(x, y)$ presents the correlation between worker type and firm type based on the equilibrium distribution of matches in the model, $h(x, y)$. The columns under Corr$(\hat{\phi}_i, \hat{\psi}_j)$ present the correlations between estimated worker and firm fixed effects, based on two different samples: (i) is based on a simulation from the stationary version of the model; (ii) is based on the first 10 years of labor market experience for a cohort who start unemployed.

types implied by the model are positive for all education groups. It is a moderate 0.34 for the high school dropouts, and strong for both high school and college graduates at 0.85 and 0.86 respectively. We present two different estimates for the correlation of estimated fixed effects. The first (i) uses a ten year sample drawn from the stationary equilibrium of the model. The second (ii) uses a sample of with a lot of mobility, corresponding to the first 10 years in the labor force for a given cohort. Looking at column (i), the correlation between estimated fixed effects is effectively zero for the low education group, despite the model correlation of 0.34. For the high school graduates, the correlation in estimated fixed effects is $-0.24$ compared to the model correlation of 0.85. We do not even obtain the correct sign for the correlation. Finally, the correlation between estimated fixed effects for the college graduates is 0.64 compared to that of 0.86 implied by the model.

The correlation between estimated fixed effects is clearly not a useful measure for the degree or sorting or the degree of complementary in this environment. Indeed, while the true model correlations are almost equal for high school and college graduates at 0.85 and 0.86, the fixed effect correlations are of opposite sign -0.24 and 0.64. This difference can be traced to several other important differences between labor markets for the high school and college workers. Considering again the parameter estimates in Table 4, we see that there is somewhat more job turnover among the high school graduates; the exogenous job destruction rate is 2.5 times as high as for college graduates, and the probability of a shock to the match productivity is six times as likely. The implication is that college graduates tend to have a relatively stable career, moving to better matches when the opportunity arises, with very little variability in the quality for any given match. High school graduates on the other hand tend to be separated to unemployment much more often, restarting the process of finding a good match. In addition, the productivity at their current match is subject to relatively high variability. A second big difference between these two labor markets are the estimates for the relative importance of firm to worker heterogeneity, measured as the ratio of $f_2/f_1$, which is 0.66 for high school graduates and 11.6 for college graduates. College graduates face much larger differences in productivity across jobs than high school graduates. The effect of mobility on the correlation of estimated fixed effects can be seen in column (ii) of Table 6, where the estimates are based on a sample of workers in their first 10 years in the labor force. While the correlation does not change much for the high school graduates, the correlation falls from 0.64 to 0.32 for the college graduates, reflecting the fact that for this sample we observe workers moving from originally mismatches jobs (which could have high $y$) to better jobs (possibly at lower $y$).

Summing up, it is clear that simply looking a the correlations between fixed effects in wages is insufficient to inform us about the degree of sorting in a labor market. Some additional information about how workers move between employment states and between jobs, combined with the implications for wages of these moves needs to be incorporated. Since sorting is by definition a statement about an equilibrium outcome, it is necessary to have a model in which the allocation of workers across jobs is an equilibrium outcome, and at the same time can account for the observation that workers continue to move between states of unemployment and employment, as well as between jobs.

Finally, note that recent work by Lentz (2010) and Bagger and Lentz (2012) also builds on the sequential auction model with bargaining that they extend to allow for endogenous search intensity. There, sorting stems both from the value of a job not being necessarily greater than the value of unemployment and also because more able workers search more intensely. In their model, however, there are no productivity shocks, vacancy creation is not modeled and the value of a vacancy is always zero, which makes welfare analysis and policy evaluation less pertinent. Still, their work delivers a rich set of interesting results. Confirming Atakan's (2006) insight,[13] they show that positive (negative) sorting holds if the match production function is supermodular (submodular); using Danish matched employer-employee data they find evidence of modest positive sorting between worker skill and firm technology; lastly, they confirm Eeckhout and Kircher's point about the non-identification of sorting from wage data only. The correlation between wage equation worker and firm fixed effects can be positive when the production function is submodular if workers' bargaining power is weak. They also show that the estimators of sorting proposed by Lopes de Melo (2009) and Eeckhout and Kircher (2011) do not work as they are always positive whether the production function is sub- or supermodular.[14] Bagger and Lentz identify sorting from the correlation between unemployment duration and wages for workers hired by the most productive firms; most productive firms can be identified because they are the most likely to successfully poach.[15]

[13]Shimer and Smith show that complementarity in production is not a sufficient condition for assortative matching in equilibrium. Atakan shows that it is in presence of search costs.

[14]Lopes de Melo proposes to use the correlation between the worker fixed effect and the average worker fixed effect of the co-workers within the firm. Eeckhout and Kircher suggest to use the population variance relative to the average within firm worker fixed effect variance.

[15]Another set of studies use firm data to estimate firm-specific total factor productivity (TFP) and correlate it with the distribution of skills or wages within the firms. See Haltiwanger, Lane, and Spletzer (1999), Abowd, Haltiwanger, Lane, McKinney, and Sandusky (2007), van den Berg and van Vuuren (2010), Mendes, van den Berg, and Lindeboom (2010).

## 5 Welfare Analysis

### 5.1 The impact of search frictions

The potential for welfare-enhancing labor market regulation arises from the job search frictions and the externalities they cause during the job allocation process. The externalities arise 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, or among vacancies, 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).[16] 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 can 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 what they could produce at home, allowing also 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 a better outside option so as to extract surplus via poaching firms.

Any regulatory intervention in the labor market will improve welfare only to the extent that it can address the externalities discussed above, and to the extent to which they are significant. Thus, to provide a measure of the potential welfare gains from labor market regulation (such as in work benefits, unemployment insurance, minimum wages, severance pay etc.) we solve the planners problem respecting the constraints arising from search frictions. Specifically the planner maximizes total output and home production subject to the flow constraints implied implied by the frictions and the costs of vacancies, i.e.

$$\max_{h^{SP},u^{SP},V^{SP}} \int f(x,y)h^{SP}(x,y)\,dx\,dy + \int b(x)u^{SP}(x)\,dx - cV^{SP} \tag{20}$$

subject to the equations (13), (14), (15), (16) and $V = \int v(y)dy$. 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 $\mathcal{M}^{\mathcal{SP}}(x, y)$ and the measure of jobs $N^{SP}$. We approximate the solution to the planner's problem by specifying the boundaries of the matching set as polynomials: $\mathcal{M}^{SP}(x, y) = \left\{x, y|x > \sum_{i=1}^{I} \tau_i y^{i-1}, y > \sum_{j=1}^{J} \tau_j x^{j-1}\right\}$. We find that polynomials of order $I, J = 4$ provide a good approximation as no increase in steady state output is found by further increasing the degree of polynomial.

Table 7 shows the breakdown of contributions to total welfare under different scenarios. The

[16] see also Sattinger (1995)

Table 7: Output and Employment

| | Decentralized Frictional | Constrained Planner | Frictionless with Actual Employment | Frictionless with Full Employment | Optimal Unemployment Insurance |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | | | Less than high school | | |
| Welfare | 100.00 | 100.00 | 108.91 | 140.79 | 100.00 |
| Match Output | 108.17 | 108.17 | 117.08 | 140.78 | 108.17 |
| Home Production | 5.22 | 5.22 | 5.22 | 0.01 | 5.22 |
| Recruiting Costs | -13.39 | -13.39 | -13.39 | 0.00 | -13.39 |
| $E/L$ | 78.52 | 78.52 | 78.52 | 94.87 | 78.52 |
| $N/L$ | 94.87 | 94.87 | 94.87 | 94.87 | 94.87 |
| $V/U$ | 76.13 | 76.13 | 76.13 | - | 76.13 |
| Match quality | 100.00 | 100.00 | 108.24 | 107.72 | 100.00 |
| Corr$(x, y)$ | 0.32 | 0.32 | 1.00 | 1.00 | 0.32 |
| | | | High school graduate | | |
| Welfare | 100.00 | 109.71 | 124.06 | 164.02 | 105.92 |
| Match Output | 115.69 | 92.27 | 139.75 | 163.98 | 88.46 |
| Home Production | 12.83 | 29.71 | 12.83 | 0.04 | 33.69 |
| Recruiting Costs | -28.52 | -12.27 | -28.52 | 0.00 | -16.23 |
| $E/L$ | 86.57 | 67.67 | 86.57 | 95.85 | 66.30 |
| $N/L$ | 95.85 | 71.66 | 95.85 | 95.85 | 71.58 |
| $V/U$ | 69.10 | 12.36 | 69.10 | - | 15.67 |
| Match quality | 100.00 | 102.24 | 121.05 | 128.31 | 100.04 |
| Corr$(x, y)$ | 0.81 | 0.82 | 1.00 | 1.00 | 0.80 |
| | | | College graduate | | |
| Welfare | 100.00 | 101.92 | 107.75 | 125.24 | 100.00 |
| Match Output | 114.98 | 112.40 | 122.74 | 125.24 | 114.98 |
| Home Production | 0.02 | 0.03 | 0.02 | 0.00 | 0.02 |
| Recruiting Costs | -15.00 | -10.51 | -15.00 | 0.00 | -15.00 |
| $E/L$ | 96.35 | 93.16 | 96.35 | 100.00 | 96.35 |
| $N/L$ | 100.22 | 95.88 | 100.22 | 100.22 | 100.22 |
| $V/U$ | 106.01 | 40.00 | 106.01 | - | 106.01 |
| Match quality | 100.00 | 101.10 | 106.75 | 104.94 | 100.00 |
| Corr$(x, y)$ | 0.83 | 0.85 | 1.00 | 1.00 | 0.83 |

Notes: Column (1) is the estimated economy. In column (2) the constrained planner chooses admissible matches to maximize steady state output. Here we restrict the planner to choosing a reservation type for $x$ and $y$, approximated by a fourth order polynomial: $\mathcal{M}^{SP}(x,y) = \left\{x, y | x > \sum_{i=1}^{4} \tau_i y^{i-1}, y > \sum_{j=1}^{4} \tau_j x^{j-1}\right\}$, and the measure of firms active in the market $N$. The frictionless benchmarks of columns (3) and (4) hold $N/L$ at the estimated decentralized level and force the positive assortative allocation $\{x, y(x)\}$. Column (3) keeps the same unemployment and vacancy distributions $u(x)$ and $v(y)$ as in the benchmark. In column (4) we optimally reallocate workers across firms and employment states, filling all possible jobs with a worker. In column (5), optimal unemployment insurance policy is modeled as proportional to home production: $b_0 b(x)$ and we impose the balanced budget $b_0 \int b(x)u(x)\,dx = \tau \int f(x,y)h(x,y)\,dx\,dy$. The implied policy parameters are less than high school: $\tau = 0$, $b_0 = 0$; high school: $\tau = 0.09$, $b_0 = 0.24$; and college: $\tau = 0$, $b_0 = 0$. Match quality is defined as average match output, normalized to 100 for the decentralized economy.

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 (20). For the lowest education group the constrained planner is not able to improve on the decentralized outcome. For the two higher education groups the planner can attain an increase in welfare of 9.7% and 1.9% respectively. The planner wants to improve average match quality (output per match), increase unemployment and reduce the number of jobs at the same time. This is more dramatically true for high school graduates. To understand what is going on, note that increased match efficiency contributes very little to the welfare increase. Output per match increases by a moderate 2.0% and 1.1%, respectively, for the two skilled groups. Hence, the main source of inefficiency is that vacancy costs are too big. With fewer jobs, there are fewer vacancies. This creates unemployment, but for high school graduates this is acceptable because they are very good at producing at home. For college graduates, $b$ is estimated much lower, so the job reduction is much smaller.

For the lowest education group the constrained planner is not able to improve on the decentralized outcome. 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 this is rarely the focus of policy.[17] Thus labor market regulation cannot be effective as an antipoverty measure. In particular the minimum wage here will be welfare reducing; severance pay on its own would be neutral (as in Lazear, 1990) or welfare reducing if combined with minimum wages. Of course, different individual weights in the welfare function, or a different way of valuing inequality, could obviously change that result.

Anti poverty measures need to address search frictions: for the lowest skill group policies that improve search technology could improve welfare by up to 46%. Part of this increase comes from reducing unemployment. But part also comes from improving sorting since the elasticity of substitution, albeit large is not infinite (it is 3.6). This can be seen from the fourth column of Table 7 where unemployment is kept equal to the level in the benchmark economy.

### 5.2 The welfare effects of mismatch

Search frictions prevent positive assortative matching and cause unemployment. Indeed there are varying degrees of complementarity in production for all three education groups and the extent to which mismatch causes welfare losses will vary. One way of gauging this is to ignore search frictions and sort workers and jobs optimally imposing perfect assortative matching as in Becker (1973).[18] Column (3) runs the counterfactual experiment where $u(x)$ and $v(y)$ are kept the same as estimated but $h(x, y)$ is changed into $[L - u(x)]\mathbf{1}\{y = y(x)\}$, which is the distribution of matches that is obtained by imposing the optimal job match to each employed worker ($y(x) = \arg\max_y S(x, y)$). Match quality increases by 8.24 percentage points for unskilled

[17] Notwithstanding the suspicious aim of the policy: improving output per match mainly by reducing the number of jobs.

[18] See Gautier and Teulings (2012) for an interesting approach to measuring the cost of search frictions in the presence of sorting, albeit in the context of a simpler model

workers and by 21.05 and 6.75, respectively, for the two skilled groups. The increases in welfare are quite similar at 8.91, 24.06 and 7.75, percentage points.

In column (4), we run a similar counterfactual in which we assign all jobs to workers. Unemployment is drastically reduced and welfare increases a lot, largely because there are no recruiting costs, but not necessarily because of reduced unemployment, as imposing full employment implies slightly lower match quality for the lowest and highest education groups (see also the discussion in Subsection 6.2).

# 6 Optimal Unemployment Insurance

The final column of Table 7 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: $(1 + b_0)\, b(x)$. Benefits are funded by a proportional tax on match output: $b_0 \int b(x)u(x)\, dx = \tau \int f(x, y)h(x, y)\, dx\, dy$.

## 6.1 Overall efficiency gains

Since the constrained planner is unable to improve over the decentralized outcome for the lowest education group it is optimal to have no unemployment insurance for this group. In interpreting this result note that utility here is linear and workers are thus assumed risk neutral; hence UI has no insurance benefit per se but could improve welfare by altering the job acceptance probabilities and indirectly eliminating very low value matches. Allowing for risk aversion and/or for liquidity constraints would add an additional channel to make UI welfare improving. We also find that it is optimal to have no unemployment insurance for the college graduates, despite the potential gain of 1.9 percent obtained by the constrained planner: for this group the relatively crude policy tool of unemployment insurance is too far from what the planner would do to produce any benefits. Finally, for the high school graduates, we find that optimal unemployment insurance can deliver 5.9% of improved welfare, corresponding to 60 percent of the potential gains attainable by the constrained planner. This involves increasing the baseline flow utility of being out of work (home production) for each individual by 24 percent and taxing output at 9 percent.

It is worth noting again that the improvement in steady state output comes from very different sources when comparing the elimination of frictions to the constrained planner or optimal unemployment insurance scheme. With the removal of frictions there is a direct increase in market production and a gain when netting out the costs of vacancy creation from home production. For the groups where the constrained planner can improve steady state output, this is implemented largely by raising unemployment and reducing the number of vacancies, resulting in lower vacancy creation costs, high levels of home production, and improving the average quality of productive matches.

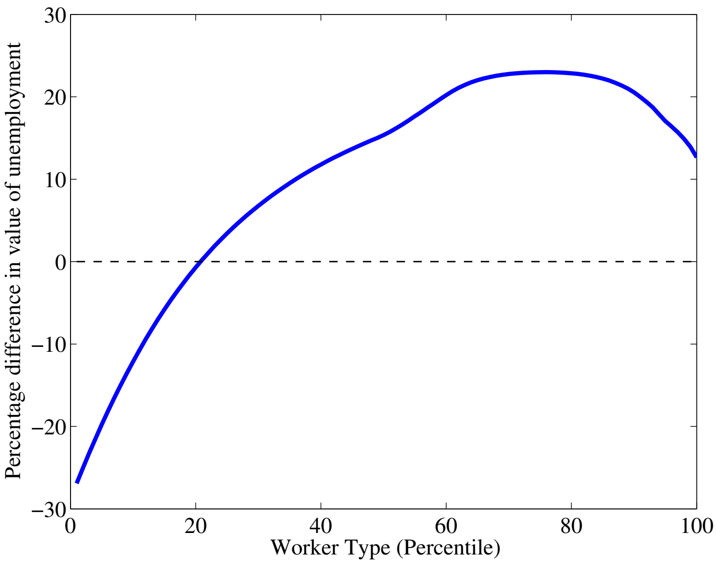



Figure 6: Welfare difference in value of unemployment by worker type

Note: We calculate the distributional impact for the high school educated group, where the optimal unemployment insurance scheme implies a five percent gain in steady state output. This is calculated as $100 \times \frac{W_0^{UI}(x) - W_0(x)}{W_0(x)}$.

## 6.2 The redistributive effects of the policy

As discussed above, we find that there is a potential gain from an optimal unemployment insurance scheme for the high school graduates, but not for either the lower or higher education groups. For the high school drop-outs there are simply no efficiency gains to realize (short of eliminating frictions) while for the college graduates the cost of increasing the duration of unemployment spells outweighs the benefit from improved matches. In addition to overall efficiency gains, we are also interested in the redistributive effects of policy. In an environment with heterogeneous workers it is not necessarily the case that an increase in steady state output will benefit all workers the same, indeed it may harm some. In Figure 6 we plot the difference in value, by worker type $x$, between being unemployed in an economy with and without the optimal unemployment insurance scheme. While the value of unemployment is higher for all worker types above the first quintile in the environment with UI, the value of unemployment is substantially lower for the bottom twenty percent of worker types. The increase in unemployment duration that improves match quality is born disproportionately by the low type workers who now have many fewer jobs to match with; they also receive lower net wages when employed due to the tax on output, and are not fully compensated by the marginal increase in benefits when unemployed. The unemployment insurance scheme, while improving overall steady state output, also has the side effect of transfer from the lowest skilled to the higher skilled. As a result the group that is usually the target of such policies ends up worse off.

# 7 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 complementary 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 higher 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 unemployment insurance can improve efficiency to the extent that they address the externalities induced by search frictions. We establish that these are very small. We then show that the optimal unemployment insurance benefits are zero for the lowest and highest education group, and moderate for the high school graduates. Moreover we show that even for this group where there are efficiency gains, the policy would redistribute wealth from low skilled workers to workers above the first quintile of the skill distribution.

Our model opens up an empirical research agenda on which to build and address important issues. We demonstrate the importance of heterogeneity, sorting and search frictions. Among these are the welfare and labor 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, and on the distribution of worker types employed at the same firm type. 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.

# A Mathematical Appendix

## A.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 jobs can also be destroyed at an exogenous rate $\xi$. Then we have that

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

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

$$\begin{aligned}
(r+\xi+\delta)\,S(x,y) &= f(x,y) - b(x) + c \\
&\quad - s_0\kappa\beta \int S(x,y)^+ v(y)\,dy - s_0\kappa\,(1-\beta) \int S(x,y)^+ u(x)\,dx \\
&\quad - s_1\kappa\,(1-\beta) \iint \left[S(x,y) - S(x,y')\right]^+ h(x,y')\,dx\,dy \\
&\quad + s_1\kappa\beta \int \left[S(x,y') - S(x,y)\right]^+ v(y')\,dy' + \delta \int S(x,y')^+\,dy'. \qquad (21)
\end{aligned}$$

Note that

$$\left[r + \delta + \xi + s_1\kappa v(\overline{B}(x,y))\right] \frac{\partial S(x,y)}{\partial y} = \frac{\partial f(x,y)}{\partial y} - (r+\delta)\Pi_0'(y), \qquad (22)$$

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 $\mu$. Therefore, for any $x$, the set of $y$'s maximizing the surplus $S(x,y)$ is the set of $y$'s maximizing the surplus flow $f(x,y) - rW_0(x) - r\Pi_0(y)$.

## A.2 The expected profits of a vacant job increase with productivity

From the previous expression note also that

$$(r+\delta)\Pi_0'(y) = s_0\kappa\,(1-\beta) \int_{S(x,y)\geq 0} \frac{\partial S(x,y)}{\partial y} u(x)\,dx + s_1\kappa\,(1-\beta) \int h(x, B(x,y)) \frac{\partial S(x,y)}{\partial y}\,dx, \qquad (23)$$

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

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

Hence, plugging the above expression for $\partial S(x, y)/\partial y$ in equation (23) shows that $\Pi_0'(y)$ is positive if $\partial f(x, y)/\partial y$ is positive.

# B Computing the Equilibrium

The equilibrium is characterized by knowledge of the number of jobs $N$, the labor 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 $(\kappa, h, S)$ can be constructed as follows.

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

1. Calculate $u(x)$ using (15) 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 (16).

Second,

1. Update $h$ using equation (14) as
$$h(x, y) \leftarrow \frac{\delta \int h(x, y')\, dy' + \left[s_0 u(x) + s_1 h\left(x, B(x, y)\right)\right] \kappa v(y)}{\delta + \xi + s_1 \kappa v\left(\overline{B}(x, y)\right)} \mathbf{1}\{S(x, y) \geq 0\}. \tag{24}$$
2. Update $S$ using equation (12).
3. Update $\kappa$ using the free entry equation (10)
$$\kappa \leftarrow \frac{(c - \delta \overline{\Pi}_0)/(1 - \beta)}{s_0 \int S\left(x, 0\right)^+ u(x)\, dx + s_1 \int \left[S\left(x, 0\right) - S(x, y')\right]^+ h(x, y')\, dx\, dy'}. \tag{25}$$

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

# C Chernozhukov and Hong's Algorithm for SMM

The estimation procedure of Chernozhukov and Hong (2003) consists of simulating a chain of parameters that (once converged) has the quasi-posterior density

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

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 $\theta^j$. 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 $\theta^0$. Next, generate $\psi$ from a proposal density $q(\psi|\theta^j)$ and update $\theta^{j+1}$ from $\theta^j$ for $j = 1, 2, ...$ using

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

where

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

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

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

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 if the optimal weighting matrix is used.[19] 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 1,000 elements (pooled over the 100 chains) to obtain parameter estimates and the standard errors.[20]

[19] To avoid the problem similar to that pointed out by Altonji and Segal (1996) we decided against using the optimal weight matrix. Indeed using it did not give sensible results. In this case the chain converges to a stationary process where the variance is a consistent estimator of the inverse of the Hessian $J^{-1}$ (Theorem 1 of Chernozhukov and Hong, 2003) and the sandwich estimator $(J^{-1}IJ^{-1})$ has to be used to calculate standard errors appropriately. Because of the lack of precision in the numerical approximation of the gradient $G(\theta) = \nabla_\theta \widehat{m}_S^M(\theta)$ (with $\widehat{I} = G(\widehat{\theta})^\intercal \widehat{W}_N^{-1} G(\widehat{\theta})$), we report the variance of the MC chain $(\widehat{J}^{-1})$. This is still informative. (Discussions with Han Hong and Ron Gallant helped us to understand this.)

[20] 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).

## References


Abowd, J. M., and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–45.

Abowd, J. M., J. Haltiwanger, J. Lane, K. L. McKinney, and K. Sandusky (2007): "Technology and the Demand for Skill:An Analysis of Within and Between Firm Differences," NBER Working Papers 13043, National Bureau of Economic Research, Inc.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67(2), 251–334.

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. G., and L. M. Segal (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business & Economic Statistics*, 14(3), 353–66.

Altonji, J. G., and R. A. Shakotko (1987): "Do Wages Rise with Job Seniority?," *Review of Economic Studies*, 54(3), 437–59.

Altonji, J. G., A. Smith, and I. Vidangos (2009): "Modeling Earnings Dynamics," NBER Working Papers 14743, National Bureau of Economic Research, Inc.

Altonji, J. G., and N. Williams (2005): "Do wages rise with job seniority? A reassessment," *Industrial and Labor Relations Review*, 58(3), 370–397.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008): "High wage workers and low wage firms: negative assortative matching or limited mobility bias?," *Journal of the Royal Statistical Society Series A*, 171(3), 673–697.

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 (2012): "An Equilibrium Model of Wage Dispersion and Sorting," Unpublished manuscript.

Barlevy, G. (2008): "Identification of Search Models using Record Statistics," *Review of Economic Studies*, 75(1), 29–64.

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

Blanchard, O. J., and P. A. Diamond (1991): "The Aggregate Matching Function," NBER Working Papers 3175, National Bureau of Economic Research, Inc.

Bonhomme, S., and J.-M. Robin (2009): "Assessing the Equalizing Force of Mobility Using Short Panels: France, 1990-2000," *Review of Economic Studies*, 76(1), 63–92.

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.

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

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

Eeckhout, J., and P. Kircher (2011): "Identifying Sorting–In Theory," *Review of Economic Studies*, 78, 872–906.

Gautier, P. A., and C. N. Teulings (2012): "Sorting and the Output Loss due to Search Frictions," Discussion Paper TI 2011-010/3, Tinbergen Institute.

Gautier, P. A., C. N. Teulings, and A. van Vuuren (2010): "On-the-Job Search, Mismatch and Efficiency," *Review of Economic Studies*, 77, 245–272.

Goux, D., and E. Maurin (1999): "Persistence of Interindustry Wage Differentials: A Reexamination Using Matched Worker-Firm Panel Data," *Journal of Labor Economics*, 17(3), 492–533.

Gruetter, M., and R. Lalive (2009): "The importance of firms in wage determination," *Labour Economics*, 16(2), 149–160.

Guvenen, F. (2007): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *American Economic Review*, 97(3), 687–712.

——— (2009): "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, 12(1), 58–79.

Hagedorn, M., T. H. Law, and I. Manovskii (2012): "Identifying Equilibrium Models of Labor Market Sorting," NBER Working Paper 18661.

Hagedorn, M., and I. Manovskii (2008): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," *American Economic Review*, 98(4), 1692–1706.

Haltiwanger, J. C., J. I. Lane, and J. R. Spletzer (1999): "Productivity Differences across Employers: The Roles of Employer Size, Age, and Human Capital," *American Economic Review*, 89(2), 94–98.

Lazear, E. (1990): "Job Security Provisions and Employment," *The Quarterly Journal of Economics*, 105(3), 699–726.

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

Lise, J. (2012): "On-the-job search and precautionary savings," *Review of Economic Studies*, forthcoming.

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

Low, H., C. Meghir, and L. Pistaferri (2010): "Wage Risk and Employment Risk over the Life Cycle," *American Economic Review*, 100(4), 1432–67.

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.

MENDES, R., G. J. VAN DEN BERG, AND M. LINDEBOOM (2010): "An empirical assessment of assortative matching in the labor market," *Labour Economics*, 17(6), 919–929.

MORTENSEN, D., AND C. PISSARIDES (1994): "Job Creation and Job Destruction in the Theory of Unemployment," *The Review of Economic Studies*, 61, 397–415.

MOSCARINI, G. (2005): "Job Matching and the Wage Distribution," *Econometrica*, 73(2), 481–516.

PETRONGOLO, B., AND C. A. PISSARIDES (2001): "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39(2), 390–431.

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.

SHI, S. (2001): "Frictional Assignment. I. Efficiency," *Journal of Economic Theory*, 98(2), 232–260.

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.

TEULINGS, C. N., AND P. A. GAUTIER (2004): "The Right Man for the Job," *Review of Economic Studies*, 71(2), 553–580.

TOPEL, R. H. (1991): "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority," *Journal of Political Economy*, 99(1), 145–76.

TOPEL, R. H., AND M. P. WARD (1992): "Job Mobility and the Careers of Young Men," *The Quarterly Journal of Economics*, 107(2), 439–79.

VAN DEN BERG, G. J., AND A. VAN VUUREN (2010): "The effect of search frictions on wages," *Labour Economics*, 17(6), 875–885.

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.