Wage dispersion and Recruiting Selection *

Benjamín Villena-Roldán
Centre for Applied Economics, University of Chile

April 2009

Abstract

In this paper I introduce a novel source of residual wage dispersion. In the model, workers are heterogenous in productivity and randomly apply to *ex ante* identical posted vacancies. Each employer simultaneously meets several applicants, offers the position to the best candidate and bargains with her about the wage. Since the outside option of the employer is to hire the second-best worker, the wage paid to the best applicant decreases in the productivity of her closest competitor. Because the assignment of workers to vacancies is random in equilibrium, each worker faces a nondegenerate distribution of wages given her productivity before applying to a job. The framework suggests that the capability of search models to generate residual wage dispersion must be restricted to match-specific shocks.

The model also predicts (i) residual wage dispersion of level wages increasing in productivity; (ii) residual wage dispersion of log wages decreasing in productivity; (iii) a negative relation between unemployment and residual wage dispersion and (iv) positive relation between productivity dispersion and residual wage dispersion. To assess these empirical predictions, I calibrate the model to match the mean and variance of the log wages in CPS data 1985-2006. The model's predictions are strongly supported in the data.

Keywords: Wage Dispersion, Recruiting Selection, Nonsequential Search.

*This paper is based on the second chapter of my dissertation at University of Rochester. I am very grateful to Mark Aguiar and William Hawkins for their advice and numerous discussions. I also thank José Mustre-del-Río for very useful comments. Usual disclaimers apply. Author's email: bvillena@dii.uchile.cl

1 Introduction

In this paper I introduce a theory of pure wage dispersion different from those relying on on-the-job search (Burdett and Mortensen 1998; Mortensen 2005) or exogenous reservation wage heterogeneity (Albrecht and Axell 1984). As in the recruiting selection model of Villena-Roldan (2008), workers are heterogenous in time-invariant productivity and apply to only one posted vacancies per period. *Ex ante* identical employers simultaneously meet a random number of workers and make an offer to the best applicant. Then, both sides bargain over the wage considering that the employer's outside option is to hire the second-best applicant. The closer the productivity of the second-best worker, the lower the wage paid to the top applicant and viceversa. Since the productivity of the second-best applicant is random in equilibrium, each worker faces a nondegenerate distribution of wages given her productivity before applying to a job.

Alternatively, the proposed mechanism can be understood as a second-price auction for a job among all the applicants in a particular vacancy. The workers "bid" by accepting lower wages to obtain the job. Under some plausible conditions, the highest productivity worker can offer the highest profits to the firm among all applicants. Thus, the top worker obtains the job, but how much of the surplus she gets depends on the productivity of other workers in the same application pool. The model explains why *ex ante* identical matches split the generated surplus differently *ex post*. Therefore, in this paper I refer to residual wage dispersion as the component of wages that is independent from worker and firm productivity. In order words, it is match-specific.

Even though most times puzzling empirical patterns motivate theoretical contributions that can account for them, sometimes the theory suggests what kind of empirical evidence we should look into. In this case, the recruiting selection model coupled with a wage-setting mechanism that allows for competing offers of applicants generates predictions that have not seemingly been previously explored in the literature. In particular, the model predicts that (i) residual wage dispersion is negatively correlated with the unemployment rate in a labor market and (ii) residual wage dispersion is positively correlated with the productivity dispersion of the labor force. I find strong support for both predictions in the US using CPS-ORG and CPS-March data for 1985-2006. These findings are robust to a large number of alternative specifications. The intuition behind these results is simple. For the first fact, consider that there are more applicants per vacancy on average when the unemployment rate is high. Therefore, the second-best applicant is expected to be closer to the top applicant. Under the described wage-setting mechanism, a large unemployment rate is negatively correlated with both the level and the volatility of paid wages. For the second fact, consider that when the dispersion of workers' productivity distribution increases, so does the productivity dispersion of the second-best applicants.

Other studies also show that residual wage dispersion and wage inequality are positively related. According to Lemieux (2006) the increasing share of groups with highly volatile wages, such as more educated and older workers in the US labor force, explains to some extent the aggregate increase of residual wage dispersion. However, his contribution explains how to correctly measure trends of residual wage dispersion, but it is silent about the underlying causes that make skilled and older workers have more wage dispersion. The explanation for the positive relation between productivity dispersion and residual wage dispersion is absent.

This paper also attempts to explain the relation between human capital literature, labor status transitions and wage dispersion. There are models explaining endogenous wage dispersion by worker heterogeneity (Albrecht and Axell 1984) in exogenous reservation wages heterogeneity. Wages of similar workers differ because some of them are paid the reservation wage of a higher type by chance. However, this source of residual wage dispersion comes from a second-order importance heterogeneity for labor market outcomes (Bontemps, Robin, and van den Berg 1999). In contrast, the recruiting selection model generates residual wage dispersion from productivity, the first-order importance source of workers' heterogeneity. Moreover, models based on reservation wage heterogeneity have a fragile equilibrium because the wage dispersion shuts down for any positive searching cost (Gaumont, Schindler, and Wright 2006). In contrast, the residual wage dispersion in this paper is robust to the introduction of search costs.

Additionally, the model generates interesting predictions about the impact of cyclical shocks in residual wage dispersion. First, the recruiting selection model provides a theoretical support for the empirical finding that the productivity of recently hired workers is countercyclical[1]. The model also predicts that a negative aggregate shock that raises unemployment will cause a reduction of the residual wage dispersion in the short-run. Although the long-run effect is theoretically ambiguous, the sign of the short-run effect tend to prevail. Some new compiled data on new hirings by Haefke et al. (2007) can be used to test these predictions.

This paper also contributes to a better understanding of the frictional wage dispersion puzzle discussed by Hornstein, Krusell, and Violante (2007). These authors show that for a wide class of sequential search models unemployment spells become counterfactually long if workers face a realistic amount of wage dispersion. Hornstein et al. (2007) show that even the canonical sequential search with on-the-job search model (i.e. a version of Burdett and Mortensen (1998)) can simultaneously fit the data of separation rates, average tenure and wage dispersion by using unreasonably high discount rates or largely

[1]See Bils (1985) and Solon, Barsky, and Parker (1994)

negative values of unemployment benefits.[2]

Although the model in this paper and the class of models Hornstein et al. (2007) analyze have the same underlying structure *conditional on worker's individual productivity*, the implications of the recruiting selection model are different in three ways. First, what drives long unemployment spells in standard search models is the rent-seeking behavior of workers waiting for good offers. The larger the dispersion in wage offers, the more attractive becomes to wait for a good draw and the longer the expected duration of unemployment. In contrast, in the recruiting selection model workers always accept the job offers they receive. There is no rent-seeking behavior because workers are hired only when their applications arrive to a pool with less productive applicants. All workers would like to accept the job, but only the most profitable applicant gets it.[3]

The traditional "rent-seeking" interpretation of the search model is problematic because empirical studies for the US, Canada and European countries, show that workers who wait more obtain *lower* wages, once controlling for other observables.[4] Moreover, non-structural unemployment duration models show that skilled, male, young and white workers have higher job finding probability; with the exception of age, all of these characteristics are associated with higher earnings.[5] Surveys like Devine and Kiefer (1991) and Eckstein and van den Berg (2007) also show that the acceptance probability of job offers is usually high and homogenous across workers. The fact that low-wage workers tend to be unemployed longer in these models is usually attributed to the fact that they have intrinsically lower job arrival rates. A recruiting selection model offers an alternative interpretation: the low arrival of offers reflects application rejections. Since workers' acceptance of job offers plays a secondary role, it is reasonable to focus on modeling why low productivity workers receive so few offers. Under the recruiting selection approach high-wage workers also experience shorter unemployment spells because they are the top applicant more often than low productivity workers.

Second, in this paper the definition of residual wage dispersion is conditional on both workers' and firms' productivities. An ideal measure of residual wage dispersion should solely reflect match-specific shocks, which is a much smaller fraction of the total variance of wages. This source of variance can only be identified in a matched employer-employee

[2]Other augmentations such as risk-aversion or stochastic wage process mitigate the problems of the canonical model, but ultimately fail to deliver a reasonable fit to unemployment duration and wage dispersion data under plausible parameterizations.

[3]This result comes from the assumption of *ex ante* identical firms. If I would allow for firm's *ex ante* heterogeneity, firms may not find it optimal to make non-rejectable offers to all best applicants and wage offers from heterogenous firms would generate incentives for workers to wait.

[4](Addison and Portugal 1989; Belzil 1995; Christensen 2002)

[5]See (Machin and Manning 1999; Abbring, van den Berg, and van Ours 2001) among many others.

panel dataset. For instance Woodcock (2008) using LEHD[6] estimates that the amount of wage dispersion due to match-specific shocks is about 19% of the total variance of wages[7], while the firm's component accounts for 25% approximately. Under the interpretation of Hornstein et al. (2007), residual wage dispersion includes firm and match specific productivities. Therefore, this paper shows that the basic underlying structure of search models is considerably more compatible with available evidence than previous research suggests.

Third, Hornstein et al. (2007) focus on average values of residual wage dispersion, neglecting the heterogeneity for different types of workers. Put differently, they assume that residual wage dispersion is independent of worker characteristics. In the recruiting selection model under an equilibrium in which more productive workers are preferred by employers, the model predicts clear cross sectional patterns of residual wage dispersion. High productivity workers have greater individual level wage dispersion, whereas low productivity workers face greater dispersion in relative (logarithmic) terms. I develop some empirical methods using CPS-ORG and CPS-March data to investigate these predictions and I find strong evidence supporting them, except for the very left tail of the distribution of estimated productivities. I argue that such a consistent cross-sectional pattern is remarkable empirical support for the model since the constructed measures of productivity are orthogonal to the measured residual wage dispersion.

The model is calibrated to target the mean and variance of the unconditional log earnings distribution of CPS (1985-2006). The numerical results illustrates the theoretical findings of the paper and quantifies them. Level residual wage dispersion increases in productivity , while relative (log) wage dispersion decreases in productivity as it is in the data (except for the very high wage workers). The amount of residual wage dispersion generated in general accounts for about 5% of the total log wage variation although this number increases up to 20% in variation in levels. These numbers are far from replicating the amount of residual wage dispersion measured in an admittedly limited way using CPS data. However, the theoretical framework of the model suggests that this discrepancy is much smaller than suggested by previous studies.

The rest of the paper is organized as follows: In Section 2 I introduce the recruiting selection model[8], emphasizing the wage determination process which is the engine of individual wage dispersion. Section 3 contains the calibration results. In Section 4 I test the empirical predictions of the model using CPS data. Finally, I present some conclusions in Section 5.

[6]US Census Bureau's Longitudinal Employer-Household Dynamics

[7]See Table 2 of Woodcock (2008)

[8]For further details and proofs, see Villena-Roldan (2008)

2 The model

There is no aggregate uncertainty and jobs are *ex ante* identical. Workers are risk-neutral and heterogeneous only in time-invariant productivity θ. They submit only one application per period.

Because all vacant jobs are *ex ante* identical, in a symmetric equilibrium all firms receive workers with the same probability. Employers decide whether to create vacant jobs and, once applications arrive, which applicant to hire, if any. Then, the employer perfectly observes the applicants' productivities and picks the best candidate. Applicants not chosen remain into the unemployed pool.

Applicants are drawn from the distribution of unemployed workers $F(\theta)$. In partial equilibrium, this distribution can be seen as exogenous. By considering that employer's vacancy creation decision affects the composition of the unemployed pool in the economy, the distribution $F(\theta)$ becomes endogenous in general equilibrium, as shown in Villena-Roldan (2008).

Once a worker is chosen, employer and applicant bargain over the wage. The production of the job depends only on the worker's productivity. Given a distribution of unemployed workers $F(\theta)$, there are equilibrium hazard rates out of unemployment $\pi(\theta)$ that partially determine the value of the outside option of the worker. The critical element that generates residual wage dispersion is that the outside option for the employer is to hire the second-best applicant of the pool with productivity θ_2. The employer can perfectly observe the productivity of all the applicants in the pool. This strong assumption eases the exposition, keeps aside informational issues and highlights the role of search frictions in recruiting selection as a source of residual wage dispersion. I discuss below that the informational assumption can be weakened without changing the results of the model.

Matching

The total number of applicants in the economy is A. Employers optimally post V vacant jobs per period. Since all employers are *ex ante* identical, workers randomly sort themselves into vacancies. Thus, the probability that the application of a given worker arrives to a vacancy is $1/V$. For any open vacancy, a random number of applicants K arrive according to a binomial distribution with probability of success $1/V$. Considering that there is a constant ratio V/A (market tightness) and that the number of unemployed workers is arbitrarily large, the number of applications received per vacancy K converges to a Poisson distribution with mean $\lambda = A/V$.

Workers

I denote θ as the productivity of best applicant who arrives to a vacancy and θ_2 as the one of the second-best applicant. For notational purposes it is important to have in mind that every time a density or an expectation is conditioned on some value of θ, it means that $\theta > \theta_2$ by definition.

Workers are wage-maximizers and discount the future at rate $\beta = (1 + r)^{-1}$. If employed, they receive a wage w previously set via bargaining with their current employer each period. An exogenous separation shock destroys the match with probability η at the end of each period. The worker then becomes unemployed and obtains a value of $Q(\theta)$. If the match survives, the worker receives the value of being employed at wage w. Consequently, the worker's lifetime value is

$$\begin{aligned} W(w) &= w + \beta(1-\eta)W(w) + \beta\eta Q(\theta) \\ W(w) &= \frac{w + \beta\eta Q(\theta)}{1-\beta(1-\eta)} \end{aligned} \tag{1}$$

While unemployed workers receive unemployment income $b(\theta)$ that depend on the productivity in an arbitrary way. They face an endogenous hazard rate (probability of finding a job) $\pi(\theta)$ depending on their productivity, which will be explained below. Conditional on her productivity θ, a worker who is the top applicant in a vacancy is hired and receives a wage w depending on the stochastic value of the employer's outside option. If not hired, the worker obtains the value of being unemployed next period. Thus, the value of an unemployed worker equals

$$Q(\theta) = b(\theta) + \beta\pi(\theta)\mathbb{E}[W(w)|\theta] + \beta(1-\pi(\theta))Q(\theta) \tag{2}$$

Employers

Employers are risk-neutral and maximize the present value of profits. With a free-entry condition, there are zero profits due to vacancy posting, so that the value of a vacancy filled with the highest productivity worker arrived is

$$J(\theta) = \frac{\theta - w}{1-\beta(1-\eta)} \tag{3}$$

Given the distribution $F(\theta)$ and A, the free entry condition determines the number of posted vacancies. Denoting the highest lifetime profits obtained from hiring the most profitable applicant $J^1 = \max_i \{J(\theta)\}_{i=1}^k$, the free entry condition becomes

$$0 = -\kappa + \beta\mathbb{E}[\mathbb{E}[J^1|K = k]]$$

In Villena-Roldan (2008) I show there is a unique solution (unique λ) to the Free-Entry problem given a distribution of unemployed workers $F(\theta)$.

Equilibrium hazard rate

To derive an equilibrium hazard rate, the approach is to first conjecture the existence of an equilibrium and then derive sufficient conditions for the equilibrium to hold. Probably the most interesting equilibrium is the "Coincidence Ranking" equilibrium:

Definition 1 *A Coincidence Ranking (CR) equilibrium is a symmetric equilibrium in which the most productive worker for a vacancy is always optimally chosen by any employer.*

Under a CR equilibrium, the probability of being chosen given that k randomly drawn applicants from $F(\theta)$ arrived is just $F(\theta)^{k-1}$. Considering that the number of applicants per vacancy K has a Poisson distribution and the Bayes' law of total probabilities, the endogenous hazard rate is

$$\pi(\theta) = \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} F(\theta)^{k-1} = e^{-\lambda(1-F(\theta))} \tag{4}$$

If there is a stochastic screening cost due to the fact that some workers have noisier signals of their productivity as in Villena-Roldan (2008), the hazard rate takes the form

$$\pi(\theta) = \chi e^{-\lambda(1-F(\theta))} + (1-\chi)\frac{1-e^{-\lambda}}{\lambda} \tag{5}$$

where χ stands for the probability that the firm can freely observe the productivities of its application pool. Thus, $1-\chi$ represents the probability that the screening process is so expensive that firms optimally decide to screen only one applicant.[9] In the latter case, the firm obtains the expected profitability of the average worker in the unemployment pool.

Consequently, in the general case under CR equilibrium, the free entry condition can be written as

$$0 = -\kappa + \beta\chi \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^{k}}{k!} \int_{\underline{\theta}}^{\infty} J(\theta) k f(\theta) F(\theta)^{k-1} d\theta + \beta(1-\chi) \int_{\underline{\theta}}^{\infty} J(\theta) f(\theta) d\theta$$

[9] Alternatively, there are two kinds of employers that cannot be distinguished by workers before applying. A share χ of vacancies have a screening technology while a share $1-\chi$ does not have it, so they just search sequentially.

Wage determination

The employer's outside option is to offer the position to the second-best applicant or to post a vacancy again if only one applicant arrives. Both best and second-best applicant would prefer to get the job as long as the value of being employed at the paid wage w surpasses their value of being unemployed $Q(\theta)$. The wage that makes a worker indifferent between being employed or unemployed is the reservation wage $\underline{w}(\theta)$, which satisfies $W(\underline{w}(\theta)) = Q(\theta)$.

The competition between the two best applicants can be thought as a second-price auction on the employer's profit. This wage-setting mechanism only requires that employers have the capability of observing applicants' productivities. If workers do not observe the productivities of their competitors in a vacancy, it is always a weakly dominant strategy for them to submit their highest bid $\theta - \underline{w}(\theta)$, i.e. to accept their reservation wage.[10] The latter holds regardless of the number of applicants or their specific productivities. Moreover, even the assumption that the employer learns applicant's productivities can be relaxed. This is possible if the employer can write an enforceable contract on the *ex post* realization of the promised worker productivity or on the profit provided to the firm.

Then, the second-best applicant "bids" for the position by pushing her wage down up to the point where she gets her reservation wage. Let us conjecture that a Coincidence Ranking equilibrium holds. Having learned the second-best worker productivity, the top applicant accepts a wage low enough to tie the best offer that her closest competitor can make to the employer. Because of the CR conjecture, the top applicant gets the job because she is able to generate higher profits than the second-best applicant.

Hence, the outside option of the firm $O(\cdot)$ is

$$O(\theta_2) = \mathbb{I}(K \geq 2) \frac{\theta_2 - \underline{w}(\theta_2)}{1 - \beta(1-\eta)}$$

where $\mathbb{I}(K \geq 2)$ is an indicator function taking value 1 if more than one applicant arrives, and 0 otherwise.

The previous approach could be extended by allowing the employer and the chosen worker to bargain about the generated surplus. For this mechanism to work, the employer should convincingly communicate the value of his outside option to the best applicant. The Nash-bargaining problem takes the form

$$\max_{w} \left(W(w) - Q(\theta)\right)^{\alpha} \left(J(\theta) - O(\theta_2)\right)^{1-\alpha}$$

[10] For first-price auctions, the hired worker's payoff will not change in competitor's "bids" by definition if workers do not know the number of competitors nor their productivities. The endogenous wage dispersion doesn't follow in this case.

The second-price auction setup is now a particular case when $\alpha = 1$. Replacing equations (1), (2) and (3), we obtain

$$\max_{w} \left(w - (1-\beta)Q(\theta)\right)^{\alpha} \left(\theta - w - O(\theta_2)\right)^{1-\alpha}$$

The general solution of this problem implies a wage depending on the productivity of the best and second-best applicants.

$$w(\theta, \theta_2) = \alpha(\theta - \mathbb{I}(K \geq 2)(\theta_2 - \underline{w}(\theta_2))) + (1-\alpha)(1-\beta)Q(\theta) \tag{6}$$

The expected value of the wage conditional on the productivity of the top worker, θ, is therefore (using law of total expectations)

$$\begin{aligned}
\mathbb{E}[w(\theta, \theta_2)|\theta] &= \mathbb{E}\left[\mathbb{E}[w(\theta, \theta_2)|\theta, K = k]\right] \\
&= \alpha\theta + (1-\alpha)(1-\beta)Q(\theta) - \alpha\mathbb{E}\left[\int_{\underline{\theta}}^{\theta} (\theta_2 - \underline{w}(\theta_2))F(\theta_2|\theta)^{k-2} f(\theta_2|\theta) d\theta_2\right] \\
&= \alpha(\theta - D(\theta)) + (1-\alpha)(1-\beta)Q(\theta) \qquad (7) \\
\text{with } D(\theta) &= \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} \int_{\underline{\theta}}^{\theta} (\theta_2 - \underline{w}(\theta_2)) \frac{(k-1)F(\theta_2)^{k-2} f(\theta_2)}{F(\theta)^{k-1}} d\theta_2 \qquad (8)
\end{aligned}$$

The last expression is constructed by considering the lack of coordination of the applicants. The arrival of a worker does not affect the probability of arrival for other workers. Consequently, the expression (8) considers that the probability of being the only applicant to a vacancy is $e^{-\lambda}$ and that the expected surplus generated by the second-best applicant is zero. The probability of being two applicants is $e^{-\lambda}\lambda$, and so on.

The value $D(\theta)$ conceptually represents the expected value of the surplus generated by the second-best worker when the top applicant has productivity θ, for all the possible number of applications received in a given vacancy. Substituting expression (7) into (1) yields

$$\mathbb{E}[W(w(\theta, \theta_2)|\theta] = \frac{\alpha(\theta - D(\theta)) + (1-\alpha)(1-\beta)Q(\theta) + \beta\eta Q(\theta)}{1 - \beta(1-\eta)} \tag{9}$$

Replacing the latter into (2) and doing some algebra, we can obtain an expression for $Q(\theta)$ for any θ

$$Q(\theta) = (1-\beta)^{-1}[b(\theta)S(\theta) + (\theta - D(\theta))(1 - S(\theta))] \tag{10}$$

$$\text{with } S(\theta) = \frac{1 - \beta(1-\eta)}{1 - \beta(1-\eta) + \alpha\beta\pi(\theta)} = \frac{r + \eta}{r + \eta + \alpha\pi(\theta)} \tag{11}$$

The reservation wage is the one that makes the worker indifferent between accepting the offer or remaining unemployed, that is $Q(\theta) = W(\underline{w}(\theta))$. Using equations (1) and (10) it follows that

$$\underline{w}(\theta) = b(\theta)S(\theta) + (\theta - D(\theta))(1 - S(\theta)) \tag{12}$$

Moreover, by substituting equation (10) into (6), a more compact expression for the wage arises

$$w(\theta, \theta_2) = \alpha(\theta - \theta_2 + \underline{w}(\theta_2)) + (1 - \alpha)\underline{w}(\theta) \tag{13}$$

The expected wage conditional on productivity type θ is thus

$$\mathbb{E}[w(\theta, \theta_2)|\theta] = \alpha(\theta - D(\theta)) + (1 - \alpha)\underline{w}(\theta)$$

In order to compare the results to the previous literature, I assume that the unemployment income $b(\theta)$ is a fraction $0 < \rho < 1$ of the expected wage conditional on the productivity, i.e. $b(\theta) = \rho\mathbb{E}[w(\theta, \theta_2)|\theta]$. Even though one may think this is an innocuous assumption , it generates an equilibrium in which all workers with positive productivity find optimal to search for a job, i.e. participation rate is 1.This fact has general equilibrium effects in the population of unemployed workers. Under other schemes, some workers may find it optimal to stay out of the labor force. Because every worker wants to participate, the hazard rate out of unemployment is higher for more productive workers under this assumption. In every model fitting the framework of Hornstein et al. (2007) workers facing high hazard rate also have lower relative wage dispersion, i.e. low Mean-min ratio. In conclusion, this particular assumption on the value of the unemployment benefit tends to generate lower residual wage dispersion once one considers the effect of productivity heterogeneity in a recruiting selection model.

Substituting the former expression in (12), I obtain that

$$\begin{aligned}
\underline{w}(\theta) &= \left(\frac{1 - (1 - \alpha\rho)S(\theta)}{1 - (1 - \alpha)\rho S(\theta)}\right)(\theta - D(\theta)) \\
\underline{w}(\theta) &= \left(\frac{1 - (1 - \alpha\rho)S(\theta)}{1 - (1 - \alpha)\rho S(\theta)}\right)\left(\theta - \sum_{k=1}^{\infty}\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}\int_{\underline{\theta}}^{\theta}(v - \underline{w}(v))\frac{(k-1)F(v)^{k-2}f(v)}{F(\theta)^{k-1}}dv\right)
\end{aligned} \tag{14}$$

Equation (14) is a Volterra linear integral equation of the second kind which is easily solved by numerical methods. By a standard application of the Contraction Mapping theorem, we can prove the existence and uniqueness of its solution.

The expected wage can also be more compactly expressed as

$$E[w(\theta, \theta_2)|\theta] = \frac{1 - (1 - \alpha)S(\theta)}{1 - (1 - \alpha)\rho S(\theta)}(\theta - D(\theta)) \tag{15}$$

Individual wage dispersion

The structure of the model provides a mechanism in which any given worker of productivity θ faces an *ex ante* nondegenerate distribution of wages. Several measures of dispersion of that distribution can be computed. Then, the expected value and the variance of the wage of a worker θ is respectively,

$$\mathbb{E}[w(\theta,\theta_2)|\theta] = \alpha(\theta - D(\theta)) + (1-\alpha)\underline{w}(\theta) \quad (16)$$

$$\mathbb{V}[w(\theta,\theta_2)|\theta] = \alpha^2\mathbb{V}[\theta_2 - \underline{w}(\theta_2)|\theta]$$
$$= \alpha^2\left(\sum_{k=2}^{\infty}\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}\int_{\underline{\theta}}^{\theta}(v-\underline{w}(v))^2\frac{(k-1)F(v)^{k-2}f(v)}{F(\theta)^{k-1}}dv - D(\theta)^2\right) \quad (17)$$

Intuitively, individual wage dispersion increases in productivity because the second-best worker is drawn from a distribution censored at a highest point as the productivity of the best worker increases. The result is formally shown

Proposition 2 *The variance of the wage is strictly increasing in θ*

Proof. See Appendix A ■

It is also possible to compute the individual wage dispersion in logarithms

$$\mathbb{V}[\log w|\theta] = \mathbb{E}_K[\mathbb{E}[(\log w)^2|\theta,k]] - (\mathbb{E}_K[\mathbb{E}[\log w|\theta,k]])^2$$

By computing a Taylor expansion, I can approximate the individual dispersion of log wages as[11]

$$\mathbb{V}[\log w|\theta] \approx (\log\mathbb{E}[w|\theta] - \mathbb{E}[\log w|\theta])^2 + \frac{\mathbb{V}[w|\theta]}{\mathbb{E}[w|\theta]^2} \quad (18)$$

This equation gives some intuition about the behavior of residual wage dispersion in relative terms. The variance of the wage is bounded by the variance of the surplus of the second-best applicant, which is finite; however, the expected value of the wage increases with productivity unboundedly because (i) in a Nash bargaining setting more wages always increase in productivity and (ii) in a Coincidence Ranking Equilibrium the outside option of the worker also increases in productivity. At some point, the square of the coefficient of variation $\mathbb{V}[w|\theta]/\mathbb{E}[w|\theta]^2$ must decline.

[11] First, I expand $\log w(\theta,\theta_2)$ around the expected surplus of the second-best applicant conditional on θ, i.e $\widetilde{\theta}_2 - \underline{w}(\widetilde{\theta}_2) = D(\theta)$. Hence,

$$\log w(\theta,\widetilde{\theta}_2) \approx \log(\mathbb{E}[w(\theta)]) + \frac{(\theta_2 - \underline{w}(\theta_2) - D(\theta))}{\mathbb{E}[w(\theta)]}$$

By subtracting $\mathbb{E}[\log w(\theta,\theta_2)]$ on both sides and integrating yields equation (18)

Other measures of wage dispersion that I compute are the average expected variance of (log) wages conditional on the number of applicants

$$\mathbb{E}_K[\mathbb{V}[w|\theta,k]] = \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}\left(\mathbb{E}_{\theta_2}[w^2|K=k] - \mathbb{E}_{\theta_2}[w|K=k]^2\right)$$
$$\mathbb{E}_K[\mathbb{V}[\log w|\theta,k]] = \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}\left(\mathbb{E}_{\theta_2}[(\log w)^2|K=k] - \mathbb{E}_{\theta_2}[\log w|K=k]^2\right)$$

Finally, other measures of wage dispersion that give information about how far away from each other are the mean of the individual wage and the extreme points of the individual distribution's support. The Mean-min ratio conditional on productivity type is

$$\text{Mmin}(\theta) \equiv E[w(\theta)|\theta]/\underline{w}(\theta) = \frac{1-(1-\alpha)S(\theta)}{1-(1-\alpha\rho)S(\theta)} = 1 + \frac{\alpha(1-\rho)(r+\eta)}{\alpha\rho(r+\eta)+\alpha\pi(\theta)} \tag{19}$$

This expression is totally analogous to the one obtained by Hornstein et al. (2007) in a continuous time framework, although the interpretation is somewhat different. Under a Coincidence Ranking Equilibrium, when the productivity and therefore, the hazard rate $\pi(\theta)$ is large, the percentage gap between the average and reservation wages decreases *conditional on productivity*. In this framework, any time a worker becomes the top applicant, the generated surplus is strictly positive so that both sides accept the match. The source of unemployment spells is the worker's application rejection, not rent-seeking behavior. Similarly, the Max-mean ratio can be written as

$$\begin{aligned}\text{maxM}(\theta) &= \frac{\alpha\theta + (1-\alpha)\underline{w}(\theta)}{\mathbb{E}(w(\theta)|\theta)} = \frac{\alpha D(\theta)}{\theta - D(\theta)}\left[\frac{1-(1-\alpha)\rho S(\theta)}{1-(1-\alpha)S(\theta)}\right] \\ &= \frac{D(\theta)}{\theta - D(\theta)}\left[\frac{(r+\eta)(1-\rho(1-\alpha))+\alpha\pi(\theta)}{r+\eta+\alpha\pi(\theta)}\right]\end{aligned} \tag{20}$$

Since the average surplus of the second best applicant is bounded below from the second best applicant of the whole distribution, the ratio $\frac{D(\theta)}{\theta - D(\theta)}$ necessarily declines as θ becomes very large, i.e. for very high productivity workers, the second-best competitor is always very far away so that she generates a low maxM(θ) value.

By integrating these over θ, we can get average individual wage dispersion measures.

Finally, the Law of Total Variance allows to compute the contribution of the second-best applicant productivity and of the number of applicants to the total variance of wages. We can express the unconditional variance of wages (or some other function of wages) as

$$\mathbb{V}[w] = \underbrace{\mathbb{V}_\theta\left[\mathbb{E}[w|\theta]\right]}_{\text{Cross population variation}} + \underbrace{\mathbb{E}_\theta\left[\mathbb{V}[w|\theta]\right]}_{\text{Average individual variation}} \tag{21}$$

The first term in (21) stands for the variation explained by average wages conditional on productivity across the labor force. The second term is the variation of wages generated at individual level. Therefore, the fraction of the total variance of wages explained by individual residual wage dispersion can be computed as

$$\mathcal{R} = \frac{\mathbb{E}_\theta \left[\mathbb{V}[w|\theta]\right]}{\mathbb{V}[w]} \tag{22}$$

There is an analogous decomposition for log wages.

Equilibrium existence

In this section I analyze some conditions needed to have a well-defined CR equilibrium.

First, although unemployment income $b(\theta)$ can freely depend on the productivity in general, some restrictions are necessary so that all workers rationally prefer to work. Using (10), this condition is simply that $Q(\theta) > (1-\beta)^{-1} b(\theta) \quad \forall \theta$, which holds if and only if $\theta > b(\theta) + D(\theta)$. Intuitively, the condition says that the productivity of a participating worker must be larger than the income received by the worker while not searching and the expected value of the outside option of a prospective employer. The latter surplus is also the loss of the surplus a worker expects due to the competition of a second-best worker.[12] When $b(\theta) = \rho \mathbb{E}[w(\theta, \theta_2)|\theta]$ the participation constraint holds if and only if $\rho < 1$.[13]

Second, since employers hire the most profitable applicant, the productivity and profitability ranking must coincide for all productivities to have a CR equilibrium. A necessary and sufficient condition for this is that the value of the hired worker $J(\theta)$ is increasing in θ. This is accomplished if and only if $\underline{w}'(\theta) \leq 1$ because in any application pool the best worker can always defeat the best offer her closest competitor can make. To obtain this condition, we derive expression (12)

$$\underline{w}'(\theta) = b'(\theta)S(\theta) + b(\theta)S'(\theta) + (1 - D'(\theta))(1 - S(\theta)) - (\theta - D(\theta))S'(\theta) \tag{23}$$

By deriving (5) and using the definition of $S(\theta)$ in (11), I obtain

$$S'(\theta) = -\lambda \chi f(\theta) S(\theta)(1 - S(\theta))$$

[12]To see this, note that the value of being employed $Q(\theta)$ must exceed the unemployment income a worker would receive if not searching for a job. Thus, the previous condition follows because $b(\theta) \leq (1-\beta)Q(\theta) = S(\theta)b(\theta) + (\theta - D(\theta))(1 - S(\theta))$

[13]Using (15), if follows that $\theta - D(\theta) > \rho \left(\frac{1-(1-\alpha)S(\theta)}{1-(1-\alpha)\rho S(\theta)}\right)(\theta - D(\theta))$ it is simple to verify that the condition holds if $\rho < 1$.

Differentiating equation (8) yields

$$D'(\theta) = \frac{\lambda f(\theta)}{F(\theta)}(\theta - \underline{w}(\theta) - \widetilde{D}(\theta))$$

$$\text{with} \quad \widetilde{D}(\theta) = \sum_{k=2}^{\infty} \frac{e^{-\lambda}\lambda^{k-2}}{(k-2)!} \int_{\underline{\theta}}^{\theta} (\theta_2 - \underline{w}(\theta_2)) \frac{(k-1)F(\theta_2)^{k-2} f(\theta_2)}{F(\theta)^{k-1}} d\theta_2$$

Substituting these results into (23) and rearranging yields

$$\lambda f(\theta)(1 - S(\theta)) \Big(\chi S(\theta)(\theta - D(\theta) - b(\theta)) - F(\theta)^{-1}(\theta - \underline{w}(\theta) - \widetilde{D}(\theta)) \Big) < S(\theta)(1 - b'(\theta)) \tag{24}$$

By inspection of (24), the partial equilibrium conditions for the existence of a CR equilibrium as in Villena-Roldan (2008) are similar.

- Sufficiently high dispersion of productivities so that the hazard rate does not increase too starkly. For a very spiky density of productivities $f(\theta)$, the value of the worker's outside option may increase so much that the associated increase in productivity cannot compensate the amount the employer has to yield to the worker. If such, the employer may be better off by not hiring the top applicant.

- Relatively low λ. If the number of interviewed applicants is too high, the chances that a very high productivity worker is hired are so close to 1 that very little surplus remains for the employer. Therefore, in case of very high unemployment, employers may obtain higher profits by hiring an applicant of lower productivity. A similar intuition applies for the case of low χ because $\theta - D(\theta) - b(\theta)$ is always positive for individuals in the labor force.

- The derivative of the unemployment benefit respect to the productivity type must be low enough. Otherwise, the worker's outside option cannot be offset by the increase of worker's productivity.

- Moreover, as $F(\theta)$ is very low the left-hand side of the inequality is likely to be negative. Therefore, the condition is more likely to hold for low productivity workers.

Endogenous distribution (General Equilibrium)

So far, I considered the distribution of unemployed workers to be exogenous. In Villena-Roldan (2008) I show how to make this distribution endogenous by taking into account the effects of recruiting selection in the aggregate distribution of worker productivities. Without going into further detail, I briefly outline here how to close the model in general equilibrium.

There is an exogenous distribution of productivity $f_\theta(\theta)$ for the whole population of workers. Since the unconditional probability of being unemployed in steady state is $\frac{\eta}{\eta+\pi(\theta)}$, the density of the unemployed is

$$f(\theta) = \frac{1}{\mathcal{U}} f_\theta(\theta) \frac{\eta}{\eta + \pi(\theta)} \tag{25}$$

where $\mathcal{U}$ stands for the aggregate rate of unemployment. Remembering that $\pi(\theta)$ depends on $F(\theta)$, by integrating we obtain a Volterra equation for $F(\theta)$ which is the endogenous distribution of unemployed workers.

$$F(\theta) = \int_{\underline{\theta}}^{\theta} f_\theta(v) \frac{\eta}{\mathcal{U}(\eta + \pi(v))} dv \tag{26}$$

Using the same rationale, the density of employed is thus

$$f^E(\theta) = \frac{1}{1-\mathcal{U}} f_\theta(\theta) \frac{\pi(\theta)}{\eta + \pi(\theta)} \tag{27}$$

Discussion

The residual wage dispersion generated in this model is match-specific. Even though Hornstein, Krusell, and Violante (2007) make a strong case against the standard sequential search model as a complete explanation for realistic wage dispersion, the introduction of worker heterogeneity that affects the probabilities of labor status transitions clarifies the discussion in some aspects. First, the amount of residual wage dispersion Hornstein et al. (2007) want to explain is excessive since it includes a large component of firm-specific effects, without any specific criticism on the empirical procedures they use to compute this dispersion. Moreover Lemieux (2006) estimates the measurement error accounts for about 10% of the total variance of CPS-ORG log hourly wages, suggesting an even smaller target residual wage dispersion for the standard model to match. Second, worker heterogeneity generates important implications for the cross sectional distribution of the residual wage dispersion that are not analyzed in a simple search and matching framework. In particular, the analysis in the preceding subsection shows that level wage dispersion increases in productivity while the relative (log) wage dispersion decreases in productivity.

Another point is that even in the simplest version of the model -without a stochastic outside option for the employer- the productivity of new hirings is countercyclical as shown by the empirical evidence using panel data for the US (Bils 1985; Solon, Barsky, and Parker 1994). The recruiting selection model can easily explain this fact. Any shock that increases the unemployment rate will lead to a higher number of applications per vacancy. The firms screen more applicants in expectation than they do when unemployment is low.

As a consequence, the productivity of hired workers improves in recessions because the expected productivity of the top applicant increases in the number of applicants received.

Moreover, distinctive implications of the model arise for the residual wage dispersion behavior along business cycles. For a cyclical shock that increases unemployment or decreases vacancies, there are short run or partial equilibrium effects -keeping $F(\theta)$ fixed- that reduce residual wage dispersion[14]. Two forces work together to reduce residual wage dispersion:

- Truncation Effect: The productivity of the second-best increases since she comes from a larger pool of applicants. Hence, the second-best worker distribution shifts right. For this reason, the distribution from which the second-best worker is drawn is "more censored" at any specific level of productivity θ. In other words, keeping θ fixed the mass of workers with lower productivity is reduced. The volatility of the distribution decreases.

- Dispersion Effect: For a higher number of workers selected, the second-best worker distribution becomes less volatile, and naturally residual wage dispersion decreases.

These intuitions are formally proven

Proposition 3 *The variance of the maximum worker productivity in a vacancy decreases in the average number of applicants, other things equal, that is,* $\partial \mathbb{V}[w(\theta)]/\partial\lambda < 0$

Proof. See Appendix. ■

In the long-run, however, the effect of an increase of the unemployment rate on residual wage dispersion is generally ambiguous. Once the number of applicants per vacancy increases, firms' recruiting selection raises the job finding rate of highly productive workers. Two opposing effects determine the long-run changes in residual wage dispersion

- Truncation effect: In the long-run (general equilibrium), more selective recruiting selection reduces the quality of the unemployed. Hence, the distribution of the second-best worker shifts to the left, making the distribution from which the second-best worker is drawn "less censored" at any given point. This increases residual wage dispersion.

- On the other hand, higher selectivity makes the group of unemployed workers more homogenous, which drives the residual wage dispersion down.

[14] The effect may be the opposite if the initial unemployment rate is very low, i.e. λ is low so that the chances that only one applicant per vacancy arrives ($e^{-\lambda}$) are relatively large. For empirical values of unemployment rate and average number of applicants per vacancy, the chance of receiving only one applicant is very small, so this theoretical possibility is not a concern.

Finally, an additional prediction of the model is that an increase in the underlying dispersion of productivities will increase the residual wage dispersion of observed wages, *other things equal*. The productivity distribution of the second-best applicant depends on the underlying distribution of productivities of the unemployed, which in turn, depends on the productivity dispersion of the labor force. Thus, an exogenous increase in the variability of the labor force will generate a greater dispersion in the distribution of the second-best applicant. Such a prediction is in line with previous findings in the literature that show a simultaneous increase of productivity (human capital, ability) and residual wage dispersion (Lemieux 2006). This prediction will also be tested.

3 Calibration and Results

As in Villena-Roldan (2008) I assume the length of a period is two weeks, because this is the median vacancy duration in the National Employer Survey (NES 97). To have a stationary unemployment rate in steady state, $\mathcal{U}$, outflows must equal inflows. The number of workers just hired each period is the average hazard rate $\mathbb{E}[\pi(\theta)]$ multiplied by the mass of unemployed workers. Since separations are exogenous, the amount of workers going into unemployment is simply the separation rate multiplied by the mass of employed workers. Hence, by using a result previously obtained in Villena-Roldan (2008)[15]

$$\begin{aligned} \mathbb{E}[\pi(\theta)]\mathcal{U} &= \eta(1-\mathcal{U}) \\ \frac{1-e^{-\lambda}}{\lambda} &= \eta\frac{1-\mathcal{U}}{\mathcal{U}} \end{aligned} \tag{28}$$

The separation rate is measured as the ratio between the number of workers unemployed for one month, to the number of total employed workers, with a correction suggested by Shimer (2005). I obtain a biweekly rate of 1.21% using CPS data 1985-2006 (see Villena-Roldan (2008)).

It is necessary to have an exogenous distribution of worker productivities. To do so, I assume that the exogenous distribution of worker's productivities $f_\theta(\theta)$ is log normal such that $\log(\theta) \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$. Using this primitive, I endogenously determine the distribution of unemployed workers $F(\theta)$ using equation (26). The mean and variance of the underlying distribution of workers' productivities is chosen to target the mean and variance of the log weekly earnings distribution in CPS-ORG data 1985-2006 (in dollars of 2000). Since the literature is generally confined to understand the residual wage dispersion of hourly wages, I assume all workers have a fixed supply of 40 weekly hours.

[15]The proof is very simple. Just integrate the right-hand side of equation (4) or (5) expressed as an infinite summation.

Two models are calibrated to assess the importance of the specific wage-setting mechanism in generating residual wage dispersion. I analyze two cases: the symmetric Nash bargaining ($\alpha = 0.5$) and the pure second-price auction, which corresponds to the case $\alpha = 1$. The targets and parameters chosen are summarized in Table 1.

Table 1: Targets/Parameters

Targets / Parameters	Model 0	Model 1	Comment
λ: Mean applicants	4.98	4.98	NES 97
$\mathcal{U}$: Unempl. rate	5.72%	5.72%	CPS average 85-06
η: Separation rate	1.21%	1.21%	CPS average 85-06
α : Worker's barg. power	0.5	1	Nash / Second-price
β : Discount factor	$\sqrt[26]{0.95}$	$\sqrt[26]{0.95}$	Standard
ρ: Replacement ratio	0.5	0.5	Hornstein et al. (2007)
μ_θ: Mean $\log(\theta)$	7.844	7.651	Get mean log wages
σ_θ: SD $\log(\theta)$	0.118	0.124	Get SD log wages
κ Vacancy cost	141487	110341	Consistent with free-entry

I show the results of the two models in Table 2. The parameters of the log-productivity distribution are chosen to target the mean and standard deviation of log weekly earnings in CPS data as closely as possible. The detailed computational algorithm is described in the Appendix. In the upper section of Table 2 I also report other unconditional moments of the data. Although the models are calibrated to target the log mean and log standard deviation of wages, do not do very well in replicating the variance of wages in level.

The middle section of Table 2 shows several average measures of individual wage dispersion for employed individuals. These measures are computed by numerical integration of the cross sectional measures of individual wage dispersion. I compare the obtained measures to the estimated productivities $\widehat{\theta}_1$ and $\widehat{\theta}_2$ I obtain from running Mincer regressions to isolate the match-specific component using CPS data in Section 4. In line with Hornstein, Krusell, and Violante (2007), the model is not able to replicate large Mean-min ratios. However, the model does generate Max-mean ratios that are larger than the magnitudes in the data. The model cannot replicate the magnitudes of the variance of residual log wages or the residual variance of level wages (at best 20% of level variance for Model 1).

Because of recruiting selection, the hazard rate of high θ workers is very close to one, which makes the gap between reservation wages and paid wages small. For low productivity workers, the outside option is mainly determined by the unemployment benefit $b(\theta)$, which explains why there may be a wider relative gap between paid and reservation wages in this case. Although the magnitudes of mean-min ratio seem as low as the values reported by Hornstein et al. (2007), the meaning of these statistics are different in the two models. While here it reflects the amount of wage dispersion once the firm and worker heterogeneity are perfectly controlled for, in Hornstein et al. (2007) the residual wage dispersion includes match and firm heterogeneity.

The lower section of the table displays measures of cross-sectional wage dispersion, that is, wage dispersion that is solely attributed to the variation in productivities. This, in turn, includes the variation of equilibrium reservation wages generated by productivity differences. Hence, the model attributes more wage variation to productivity heterogeneity than the data suggests. I compute the contribution of individual wage dispersion to the total variation in wages using the equation (22). In the best case, the model can explain about 20% of the total variation in wages which is number that resembles the magnitude obtained by Woodcock (2008) for match-specific variation using a matched employer-employee panel data. Since the residuals computed in this paper utilize CPS data, the magnitude of residual wage dispersion observed can be interpreted as an upper bound of the match-specific residual wage dispersion we would observe using an ideal dataset.

Cross-sectional results are shown in Figures 1 and 2. The equilibrium reservation wage and expected wages for both models are depicted in Figure 1. The Nash-bargaining model reservation wages and average wages lie below of their counterparts for the second-price auction case. This is a consequence of the greater share of the surplus workers can extract when there is a second-price auction wage-setting mechanism. In Panel D of Figure 1, we can observe that the standard deviation of the individual level wage is clearly increasing for both calibrations, as predicted by Proposition 2. The Nash case generates less individual level wage dispersion because the lower bargaining parameter attenuates the variability induced by the competition of the second-best applicant.

I display cross-sectional patterns of individual wage dispersion in Figure 2. As analyzed above, there is a decreasing pattern in productivity for all of the measures. The fact that the max-mean ratio is higher than the mean-min ratio suggests that workers are usually paid wages relatively close to their reservation wages, and only rarely obtain earnings near the maximum wage conditional on their productivity.

4 Empirical Evaluation

In this section, I explain how I treated the data to empirically assess the model. Since the model generates match-specific volatility in wages, the ideal dataset should contain

Table 2: Results of calibrated models

Unconditional Moments

		Data	Model 0	Model 1
$\mathbb{E}[w]$		14.825	14.668	14.391
$\mathbb{E}[\log w]$		2.531	2.533	2.534
$\mathbb{V}[w]$		99.699	56.969	50.840
$\mathbb{V}[\log w]$		0.322	0.328	0.325
$\mathbb{CV}[w]$		0.674	0.515	0.495

Individual wage dispersion

	$\widehat{\theta}_1$	$\widehat{\theta}_2$	Model 0	Model 1
$\mathbb{E}_\theta\left[Mmin[w\|\theta]\right]$	1.807	1.827	1.031	1.030
$\mathbb{E}_\theta\left[maxM[w\|\theta]\right]$	1.744	1.740	3.442	4.504
$\mathbb{E}_\theta\left[\mathbb{V}[w\|\theta]\right]$	47.288	48.045	4.357	10.609
$\mathbb{E}_\theta\left[\mathbb{V}[\log w\|\theta]\right]$	0.137	0.136	0.012	0.017
$\mathbb{E}_\theta[\mathbb{E}_K[\mathbb{V}[w\|\theta, K]]]$			0.046	0.070
$\mathbb{E}_\theta[\mathbb{E}_K[\mathbb{V}[\log w\|\theta, K]]]$			0.001	0.001

Cross-section wage dispersion

	$\widehat{\theta}_1$	$\widehat{\theta}_2$	Model 0	Model 1
$\mathbb{V}_\theta\left[\mathbb{E}\left[w\|\theta\right]\right]$	19.242	24.771	52.533	40.108
$\mathbb{V}_\theta\left[\mathbb{E}\left[\log w\|\theta\right]\right]$	0.186	0.186	0.315	0.308
$\mathbb{CV}\left[\mathbb{E}[w\|\theta]\right]$	0.296	0.336	0.494	0.440
Contribution individual variance w	0.474	0.482	0.076	0.209
Contribution individual variance $\log w$	0.424	0.423	0.035	0.052

$\mathbb{E}_x$ denotes expectation with respect to x; $\mathbb{V}_x$ denotes expectation with respect to x, $\mathbb{CV}$ denotes coefficient of variation. Measures of estimated productivity $\widehat{\theta}_1$ and $\widehat{\theta}_2$are based on OLS estimation of augmented Mincer regressions. For detailed explanation of the computation of these estimations of productivity, see Section 4.

repeated data of matches with clear identification of workers and firms involved. On the other hand, such a demanding empirical and computational task lacks comparability with previous studies on wage dispersion. Moreover, implications about the relation between unemployment and residual wage dispersion would be hardly evaluated using available employer-employee matched data sets since a large fraction of unemployed workers have very short or none employment records. Although *a priori* limited by the characteristics of the data, I try to isolate residual wage dispersion that does not relate to firm or worker productivity. The approach in this paper is to construct some measures of productivity $\widehat{\theta}$ by fitted values of log wages obtained from augmented Mincer regressions.

The dataset I use contains merged information of the CPS Outgoing Rotation Panel (CPS-ORG) and the CPS March Supplement (CPS-MAR). I focus on hourly wages for the sake of comparability with the rest of the literature.[16] Since two wages are observed for a large sample of individuals, I proceed to estimate an augmented Mincer regression using the following strategy. First, using the previous year's data from the CPS-MAR I run a Mincer regression such as

$$\log w_{it}^{\text{MAR}} = A_1 X_{it} + A_2 Z_{it}^{\text{MAR}} + \mu_{it}$$

where X_{it} is a vector of worker characteristics that are directly related to individuals' productivities, and Z_{it} is a vector of other firm or worker characteristics that are not related with the individuals' productivity. The residual obtained from the OLS estimation of the former equation, $\widehat{\mu}_{it}$, represents an unobserved productivity component of the worker plus measurement error. I then proceed to the second step in which I run the following augmented Mincer regression.

$$\log w_{it}^{ORG} = B_0 \widehat{\mu}_{it} + B_1 X_{it} + B_2 Z_{it}^{\text{ORG}} + \epsilon_{it} \tag{29}$$

I obtain a measure of residual wage dispersion by computing the OLS residuals $\widehat{\epsilon}_{it}$ of (29). I also compute a measure of log productivity as

$$\widehat{\theta} = \widehat{B}_0 \widehat{\mu}_{it} + \widehat{B}_1 X_{it} \tag{30}$$

and a measure of composition adjusted log wages –which is free from firm/demographic characteristics– as

$$\begin{aligned} \log w_{it}^{c} &= \widehat{B}_0 \widehat{\mu}_{it} + \widehat{B}_1 X_{it} + B_2 \overline{Z}^{\text{ORG}} + \widehat{\epsilon}_{it} \\ &= \overline{\log w^{ORG}} + \widehat{\theta} - \widehat{B}_0 \overline{\widehat{\mu}}_{it} + \widehat{B}_1 \overline{X}_{it} + \widehat{\epsilon}_{it} \end{aligned} \tag{31}$$

[16]In Villena-Roldan (2008) I use weekly earnings so the numbers obtained in this paper are not directly comparable with the former.

where $\overline{Z}^{\text{ORG}}$ represents the average firm and demographic characteristics that are unrelated to productivity and the overlines represent sample averages as usual. The intention is to build a measure of wages that is compatible with the theoretical environment of the model. For this reason, I cleaned the data from firm specific components to obtain wages paid by "ex ante identical firms" to workers that only differ in terms of productivity.

It is not readily clear what should be the best definition for X and Z variables in the previous framework. Since it is debatable I pursue two approaches. I denote $\widehat{\theta}_1$ as an estimated log productivity which only considers a set of educational indicators[17] and the residual of the first stage, $\widehat{\theta}$. A second measure, $\widehat{\theta}_2$, considers the effect of experience by adding an age quartic polynomial to the vector X. On the other hand, the vector Z includes age polynomials (for the first case $\widehat{\theta}_1$), sex, race, state dummies, year dummies, unemployment rate, 23 industry dummies, 6 occupation dummies[18], married dummy, union dummy and weeks unemployed last year (to capture possible stigma effects). For the first stage on CPS-MAR, 5 employer size dummies are also available.

I use this framework to empirically assess the implications derived from the model. Since the model predicts a negative correlation between the residual wage dispersion and the unemployment rate, I regress the squared residual of the second stage $\widehat{\epsilon}_{it}^2$ on different measures of the unemployment rate. The evidence in Table 3 shows that high-unemployment labor markets have significantly lower residual wage dispersion regardless of the definition of the labor market. I consider different measures of unemployment: from broad measures defined by just year and state to much narrower ones (including sex, two educational categories and four age groups). Since the original Mincer equation does control for unemployment rate, I conclude that the variance of the residual depends on the unemployment level and it is not an artifact due to an erroneous econometric specification. I also consider two alternative definitions of unemployment: the traditional one, named U_t, used by the CPS and an alternative definition, which includes the group of "passive searchers" into the unemployment,[19] denoted by U_p. The change of definition does not alter the results. In the column B of Table 3, I control for year and state dummies to rule out that omitted time trends or geographical idiosyncratic patterns are generating the results. This is specially important in light of the large literature on the increasing inequality trend of wages in the US (Lemieux 2006). In column C, I finally introduce

[17]Here I follow Lemieux (2006) and I define the same seven educational categories he does in order to deal with the discontinuous measurement practices of educational attainment in the CPS. As this author I use 0-4, 5-8, 9, 10, 11, 12, 13-15, 16 and 17+.

[18]The exact codes of the industry and occupation dummies are available on request.

[19]Passive searchers declare to want to work although they have not done any search effort in the last four weeks before the interview. For an analysis of the behavior of this group see Jones and Riddell (1999)

occupational regressors to account for possible changes of the occupational composition of the labor force that may be confounded with changes in unemployment rates. The negative relation between the two variables is surprisingly robust to the different definitions of unemployment rates, labor markets and controls considered.[20]

Table 3: Residual wage dispersion vs Unemployment rate

	A		B		C	
Main regressor	Coef	t-stat	Coef	t-stat	Coef	t-stat
U_t, (year, state)	-0.173	5.176	0.030	0.656	0.031	0.673
U_p, (year, state)	-0.146	5.843	0.039	1.067	0.041	1.100
U_t, (year, state, skill)	-0.183	8.593	-0.123	5.176	-0.142	5.726
U_p, (year, state, skill)	-0.136	8.717	-0.085	4.813	-0.101	5.457
U_t, (year, state, skill)	-0.127	6.932	-0.076	3.826	-0.096	4.654
U_p, (year, state, skill, sex)	-0.103	7.356	-0.057	3.732	-0.055	3.514
U_t, (year, state, skill, age)	-0.184	9.870	-0.141	6.968	-0.157	7.514
U_p, (year, state, skill, age)	-0.146	10.481	-0.109	7.149	-0.124	7.817
U_t, (year, state, skill, age, sex)	-0.130	8.219	-0.094	5.632	-0.109	6.366
U_p, (year, state, skill, age, sex)	-0.108	8.917	-0.077	5.922	-0.077	5.795

Notes: Regression coefficients of the Unemployment rate on squared residuals. Definitions of unemployment rates in the text. Specification A does not have other controls. Specification B controls for state and year indicator variables; Specification C controls are state, year and occupation variables (Management, Professional, Production worker, Service worker, Sales, Clerical). Specification C also includes interaction dummies for years after 2002 to consider the effect a the change in occupation classification.

The model also predicts a positive relation between individual wage dispersion and labor market productivity dispersion. To assess this hypothesis I run regressions of squared Mincerian residuals and productivity dispersion measures for various definitions of labor markets. Table 5 displays these estimates for different labor markets defined by year, region, [21] occupation and sex. Column A shows the regression coefficients obtained from a regression between the squared residuals and the market productivity dispersion without additional controls. Column B includes year and state dummies, while Column C incorporates additional occupation dummies as controls (with interaction for years after 2003 due

[20]Only the two upper rows in Columns B and C show positive but statistically insignificant results. For this specifications I use unemployment rates defined by year and state. Hence, a great deal of the variation of this definition of unemployment rate is explained by these controls.

[21]In order to keep a reasonable number of observations per cell, I classified the US states in Northeast, Midwest, South and West.

Table 4: Residual wage dispersion vs Productivity dispersion

	A		B		C	
Main regressor	Coef	t-stat	Coef	t-stat	Coef	t-stat
$SD(\widehat{\theta}_1)$ (year, state)	0.584	18.86	0.300	7.039	0.293	6.861
$SD(\widehat{\theta}_1)$ (year, state, sex)	0.512	20.01	0.326	10.534	0.293	9.415
$SD(\widehat{\theta}_1)$ (year, state, sex, occup)	0.252	24.99	0.217	20.940	0.284	20.240
$SD(\widehat{\theta}_2)$ (year, state)	0.596	18.92	0.275	6.338	0.268	6.182
$SD(\widehat{\theta}_2)$ (year, state, sex)	0.461	19.17	0.270	9.591	0.214	7.480
$SD(\widehat{\theta}_2)$ (year, state, sex, occup)	0.278	25.05	0.240	21.052	0.254	18.072

Notes: Regression coefficients of the std. deviation of log productivities on squared residuals. Productivity dispersions are computed in cells defined by variables in the first column. Specification A does not have other controls. Specification B controls for state and year indicator variables; Specification C controls are state, year and occupation variables (Management, Professional, Production worker, Service worker, Sales, Clerical). Specification C also includes interaction dummies for years after 2002 to consider the effect a the change in occupation classification.

to a classification change in CPS data). The results show that there is a highly significant and positive relation between individual wage dispersion and labor market productivity wage dispersion. To my knowledge, this paper proposes the first theoretical explanation that links the first-order importance worker heterogeneity (i.e. productivity) to residual wage dispersion. The result is remarkably robust to the introduction of controls and to the definition of the labor market. In particular, it is important that the effect remains significant after controlling for annual dummy variables in light of the series of papers reporting secular trends of increasing wage inequality in the US during the last decades. Likewise, controlling for occupational and geographical effects confirms that the result is not generated by compositional changes of the labor force. The fact that the result remains unaltered under different definitions of labor market reinforces the idea that the phenomenon is a quite general finding in CPS data.

I finally examine the predictions regarding the relationship between individual wage dispersion and individual productivity. The model predicts that level measures of dispersion increase in productivity while relative measures decrease in productivity. To a great extent, I find support for this prediction in the data. However, for very high productivity workers the evidence shows an increasing relationship for relative measures of individual wage dispersion. I follow two strategies to assess this hypothesis. First, I create cells defined by biannual periods, region and 20 percentiles (ventiles) of both estimated log productivity distributions. For each of these cells, I compute the following individual residual wage dispersion measures: standard deviation of log residual wages, Mean-min ratio and Max-min ratio. Mimicking Hornstein et al. (2007), I use percentiles 1^{st}, 5^{th}, 10^{th} for min wages and 90^{th}, 95^{th} and 99^{th} for max wages. I also use the composition-corrected wage measure computed in (31) to compute the standard deviation of levels, expected wage and the coefficient of variation in levels. The results of this exercise are displayed in Table 5. The column labeled "All cells" shows a positive relation between the Mean-min ratio and the estimated productivity. In the even columns, the same analysis is conducted without the top 10% of productivities of the sample. The results change dramatically, suggesting that the left tail of the distribution of productivities behaves very different from the bottom 90%.

Table 5: Individual wage dispersion vs Productivity

	All cells, θ_1		Restricted, θ_1		All cells, θ_2		Restricted, θ_2	
Dep. Var	Coef	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat
SD($\log w^r$)	-0.003	5.1	-0.123	281.3	0.006	13.7	-0.088	224.6
SD(w^c)	5.061	226.8	2.319	90.4	5.162	249.7	2.886	119.8
$\mathbb{E}(w^c)$	14.606	1586.7	12.236	1514.6	14.951	1467.6	12.255	1412.1
CV(w^c)	-0.158	123.5	-0.278	179.2	-0.133	116.3	-0.214	150.8
Mean-min1	0.876	159.2	-0.213	40.1	0.874	170.6	-0.100	19.5
Mean-min5	0.465	241.9	-0.023	15.0	0.467	259.5	0.026	17.8
Mean-min10	0.238	199.0	-0.084	98.3	0.224	209.0	-0.041	49.1
Max-mean99	-0.194	193.9	-0.422	441.7	-0.139	177.4	-0.301	379.7
Max-mean95	-0.471	326.7	-0.788	551.9	-0.377	326.1	-0.599	490.6
Max-mean90	-1.278	334.3	-1.811	409.2	-1.100	303.6	-1.441	326.6

Notes: "Individual" residual wage dispersion measures are computed by dividing the data in cells according to biannual periods, region and 20 quantiles of estimated productivity. The table reports coefficients obtained from a regression of the dispersion measure at each cell on the average productivity measure of the cell. "All cells" columns use information of all available cells with more than 100 individuals; "Restricted cells" estimates additionally drop from the cells the top 10% of most productive workers.

An alternative, more graphical approach to obtain an empirical measure of the relation between productivity and individual wage dispersion is to measure the individual wage dispersion "locally" that is, for individuals with similar productivity level. To implement this idea I develop a rolling nearest-neighbor estimation of the local dispersion measures. The algorithm is very simple. I sort all the workers in the economy according to a measure of productivity and compute "local" measures of residual wage dispersion using a moving sample of size $\sqrt{N}$. In Figures 3 and 4, the pictures obtained from applying this method on the data show that the residual wage dispersion measures decline in productivity for a large number of individuals. The pattern is also found for the Max-mean (upper left), Mean-min (upper right) and standard deviation of log residual wages and coefficient of variation of corrected wages (lower left). Nevertheless, for the very far left tail of the distribution, the relation turns out to be positive. As a histogram of the distribution of estimated log productivities shows, the positively sloped portion of the curve only occurs for a small group with very high productivity. This observation is fully consistent with the results obtained in Table 5. Once the top 10% of the observations are dropped, there is a clear negative and significant relation between the residual wage dispersion measures: standard deviation log residual wages, Mean-min and Max-mean ratios.

Therefore, the model does a good job in predicting the shape of the profile for most workers except for the ones with very high productivities. On the other hand the magnitudes of residual wage dispersion are considerably larger than the ones generated by reasonable calibrations of the model. The lower right panel of Figures 3 and 4 depict the local standard deviation of levels of composition-corrected wages and average level of composition-corrected wages. As it is theoretically shown in the model, both series are increasing in productivity.

All in all, considering that the residuals and the estimated productivities are orthogonal by construction, it is somewhat surprising that such a clear cross sectional pattern emerges from this analysis. The class of search models analyzed by Hornstein, Krusell, and Violante (2007) fails to achieve the upper bound of residual wage dispersion found in CPS data. However, the same framework in the context of a recruiting selection model generates a cross sectional pattern that mimics the cross sectional patterns of residual wage dispersion for about 90% of the labor force.

5 Conclusions

In this paper I outline an alternative theoretical approach of wage dispersion, different from those relying on the acceptance decision of workers. Instead, the focus is put on the probability that the worker receives a job offer, which seems to be a more significant driving force of unemployment spells according to several structural search model estimations surveyed by Devine and Kiefer (1991) and Eckstein and van den Berg (2007). Although this framework shares the same basic structure of the class of standard sequential search models analyzed by Hornstein, Krusell, and Violante (2007) conditional on individual productivity, it theoretically constrains the scope of the residual wage dispersion the model may be able to explain. Hence, the match-specific contribution to the total variance of wages is the reasonable target for search models to match provided the impact of individual heterogeneity is controlled for. In particular, the recruiting selection approach predicts a negative relationship between unemployment duration and wages, as found in CPS data. The model also builds a bridge between human capital theory and residual wage dispersion.

The calibrated model can obtain reasonable magnitudes of residual wage dispersion, especially in wage levels. This can be as large as 20% according to a reasonable calibration using a second-price auction wage-setting mechanism. Although the amount of residual wage dispersion obtained from the model is not a success, it shows that the gap between the search model and the data is considerably smaller than previous research suggests.

The novelty of the approach advanced in this paper suggests to look into new evidence. The predicted cross-sectional relation between log and level residual wage dispersion and productivity has strong support in the data except for the very high wage workers as shown in Section 4. Similarly, the theoretical link between residual wage dispersion and productivity dispersion has also support in the data.

References

Abbring, J. H., G. J. van den Berg, and J. C. van Ours (2001, Oct). Business cycles and compositional variation in U.S. unemployment. *Journal of Business & Economic Statistics 19*(4), 436–448.

Addison, J. T. and P. Portugal (1989). Job displacement, relative wage change and duration of unemployment. *Journal of Labor Economics 7*, 281–301.

Albrecht, J. W. and B. Axell (1984). An equilibrium model of search unemployment. *The Journal of Political Economy 92*(5), 824–840.

Belzil, C. (1995, aug). Unemployment duration stigma and re-employment earnings. *The Canadian Journal of Economics 28*(3), 568–585.

Bils, M. J. (1985). Real wages over the business cycle: Evidence from panel data. *The Journal of Political Economy 93*(4), 666–689.

Bontemps, C., J.-M. Robin, and G. J. van den Berg (1999, Nov). An empirical equilibrium job search model with search on the job and heterogeneous workers and firms. *International Economic Review 40*(4), 1039–1074.

Burdett, K. and D. Mortensen (1998, May). Wage differentials, employer size and unemployment. *International Economic Review 39*(2), 257–273.

Christensen, B. (2002, January). Reservation wages, offered wages and unemployment duration - new empirical evidence. Kiel Institute of World Economics, Working

Paper No 1095.

David, H. A. (1970). *Order Statistics*. Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, Inc.

Devine, T. J. and N. M. Kiefer (1991). *Empirical Labor Economics: The Search Approach*. Oxford University Press.

Eckstein, Z. and G. J. van den Berg (2007). Empirical labor search: A survey. *Journal of Econometrics 136*, 531–564.

Gaumont, D., M. Schindler, and R. Wright (2006, 5). Alternative theories of wage dispersion. *European Economic Review 50*(4), 831–848.

Haefke, C., M. Sonntag, and T. van Rens (2007, July). Wage rigidity and job creation. Mimeo, Universitat Pompeu Fabra.

Hornstein, A., P. Krusell, and G. L. Violante (2007, November). Frictional wage dispersion in search models: A quantitative assessment. Working Paper 13674, National Bureau of Economic Research.

Jones, S. R. G. and W. C. Riddell (1999, Jan). The measurement of unemployment: An empirical approach. *Econometrica 67*(1), 147–162.

Lemieux, T. (2006, June). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review 96*(3), 461–498.

Machin, S. and A. Manning (1999). *The causes and consequences of longterm unemployment in Europe*, Volume 3 of *Handbook of Labor Economics*, Chapter 47. Elsevier Science.

Mortensen, D. T. (2005). *Wage Dispersion: Why are similar workers paid differently?* The MIT Press.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review 95*(1), 25–49.

Solon, G., R. Barsky, and J. A. Parker (1994). Measuring the cyclicality of real wages: How important is composition bias. *The Quarterly Journal of Economics 109*(1), 1–25.

Villena-Roldan, B. (2008). Aggregate implications of employer search and recruiting selection. Mimeo, University of Rochester.

Woodcock, S. (2008). Wage differentials in the presence of unobserved worker, firm, and match heterogeneity. *Labour Economics 15*(4), 772.

Appendix A: Proofs

Proof of Proposition 2

In general, for an non-central moment of order M we have that

$$\frac{\partial}{\partial\theta}\sum_{k=1}^{\infty}\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}\int_{\underline{\theta}}^{\theta}(v-\underline{w}(v))^M(k-1)\frac{F(v)^{k-2}}{F(\theta)^{k-1}}f(v)dv$$
$$=\frac{f(\theta)}{F(\theta)}\sum_{k=1}^{\infty}\frac{e^{-\lambda}\lambda^{k-1}}{(k-2)!}\left[(\theta-\underline{w}(\theta))^M-\int_{\underline{\theta}}^{\theta}(v-\underline{w}(v))^M\frac{F(v)^{k-2}}{F(\theta)^{k-1}}f(v)dv\right]$$

Using Mean Value Theorem for integrals and simplifying, the latter expression becomes

$$\frac{\lambda f(\theta)}{F(\theta)}\left[(\theta-\underline{w}(\theta))^M-(\theta_k^m-\underline{w}(\theta_k^m))^M\right]$$

With $M=1,2,...$ and $\theta_k^m\in(\underline{\theta},\theta)\quad\forall k=1,2,....$

An useful intermediate result is established in the following lemma.

Lemma 4 *For any moment order $M\geq 2$, it holds that $\theta_k^m-\underline{w}(\theta_k^m)>\theta_k^1-\underline{w}(\theta_k^1)$ for all $k=1,2,3,....$*

Proof. By Jensen's inequality for any convex function and a probability measure $G(v)$ we know that

$$\int_{\underline{\theta}}^{\theta}(v-\underline{w}(v))^M dG(v)>\left(\int_{\underline{\theta}}^{\theta}(v-\underline{w}(v))dG(v)\right)^M$$

Replacing on both sides of the inequality by using the Intermediate Value Theorem for integrals yields

$$(\theta_k^m-\underline{w}(\theta_k^m))^M>(\theta_k^1-\underline{w}(\theta_k^1))^M\Rightarrow\theta_k^m-\underline{w}(\theta_k^m)>\theta_k^1-\underline{w}(\theta_k^1)$$

which proves the lemma. ■

Back to the main proof, deriving the variance in (17) with respect to θ yields

$$\frac{\partial\mathbb{V}[w(\theta)]}{\partial\theta}=\left(\frac{\alpha^2f(\theta)}{F(\theta)}\sum_{k=2}^{\infty}\frac{e^{-\lambda}\lambda^{k-1}}{(k-2)!}\left[(\theta-\underline{w}(\theta))^2-\int_{\underline{\theta}}^{\theta}(v-\underline{w}(v))^2\frac{F(v)^{k-2}}{F(\theta)^{k-1}}f(v)dv\right]\right.$$
$$\left.-2D(\theta)\left[(\theta-\underline{w}(\theta))-\int_{\underline{\theta}}^{\theta}(v-\underline{w}(v))\frac{F(v)^{k-2}}{F(\theta)^{k-1}}f(v)dv\right]\right)$$

Then, by applying the Mean Value Theorem for integrals we obtain

$$\frac{\partial\mathbb{V}[w(\theta)]}{\partial\theta}=\frac{\alpha^2f(\theta)\lambda}{F(\theta)}\left[(\theta-\underline{w}(\theta))^2-(\theta_k^2-\underline{w}(\theta_k^2))^2-2(\theta_k^1-\underline{w}(\theta_k^1))((\theta-\underline{w}(\theta))-(\theta_k^1-\underline{w}(\theta_k^1)))\right]$$

Denoting $x = \theta - \underline{w}(\theta)$, $x_k^1 = \theta_k^1 - \underline{w}(\theta_k^1)$ and $x_k^2 = \theta_k^2 - \underline{w}(\theta_k^2)$ by the Lemma 4 it follows that $x > x_k^2 > x_k^1$ for all k. Because of the latter result, the sign of the latter expression is proved to be positive for all k.

$$\begin{aligned} x^2 - (x_k^2)^2 - 2x_k^1(x - x_k^1) &= (x - x_k^2)(x + x_k^2) - 2x_k^1(x - x_k^1) \\ &> (x - x_k^2)(x + x_k^2) - 2x_k^1(x - x_k^2) = (x - x_k^2)((x - x_k^1) + (x_k^2 - x_k^1)) > 0 \end{aligned}$$

Since the lower bound of the latter expression is always positive for all k, the latter result proves the proposition.

Proof of Proposition 3

In order to prove that the variance of wage decreases in the mean number of applicant $\partial \mathbb{V}[w(\theta, \theta_2)|\theta]/\partial \lambda < 0$, I first establish an intermediate result

Lemma 5 *For an application pool of size greater or equal than 2 ($K \geq 2$), the variance of the wage is strictly decreasing in K, i.e. $\mathbb{V}[w(\theta, \theta_2)|\theta, K = k-1] > \mathbb{V}[w(\theta, \theta_2)|\theta, K = k]$.*

Proof. For this proof, I denote $\mu_{n,h}^{(m)}$ as the m-th uncentered moment of the distribution of wages of the h-th highest draw among n draws. Consider the expected difference of the two moments of the distribution of unemployed workers conditional on being k and $k-1$ applicants respectively

$$\mu_{k,k}^{(m)} - \mu_{k-1,k-1}^{(m)} = \int v^m (kF(v)^{k-1} - (k-2)F(v))f(v)dv$$

After a bit of algebra, the last expression equals

$$\frac{1}{k-1} \int v^m ((k(k-1)F(v)^{k-2}(1-F(v)) + (k-1)F(v)^{k-2})f(v)dv = \frac{1}{k-1}(\mu_{k,k-1}^{(m)} + \mu_{k-1,k-1}^{(m)})$$

The same result is obtained from a more general identity in David (1970). Due to this result, we have that $\mu_{k-1,k}^2 + (k-1)\mu_{k,k}^2 = k\mu_{k-1,k-1}^2$ and $\mu_{k-1,k} + (k-1)\mu_{k,k} = k\mu_{k-1,k-1}$. Therefore we have the following

$$\begin{aligned} k\mu_{k-1,k-1}^{(2)} &> (k-1)\mu_{k,k}^{(2)} + (\mu_{k-1,k})^2 \quad \text{Due to Jensen's inequality} \\ &= (k-1)\mu_{k,k}^{(2)} + (k\mu_{k-1,k-1} - (k-1)\mu_{k,k})^2 \end{aligned}$$

Then, rearranging it follows that

$$\begin{aligned} k(\mu_{k-1,k-1}^{(2)} - \mu_{k,k}^{(2)}) &> k^2(\mu_{k-1,k-1})^2 - 2k(k-1)\mu_{k-1,k-1}\mu_{k,k} + (k-1)^2(\mu_{k,k})^2 \\ &= k(\mu_{k-1,k-1})^2 - 2(k-1)\mu_{k-1,k-1}\mu_{k,k} + 2(k-2)(\mu_{k,k})^2 \\ &> k(\mu_{k-1,k-1})^2 - 2(\mu_{k,k})^2 \quad \text{Because } \mu_{k,k} > \mu_{k-1,k-1} \end{aligned}$$

This result implies

$$\mu^{(2)}_{k-1,k-1} - \mu^{(2)}_{k,k} > (\mu_{k-1,k-1})^2 - (\mu_{k,k})^2 \quad \text{For } k \geq 2$$

Thus, $\mu^{(2)}_{k-1,k-1} - (\mu_{k-1,k-1})^2 = \mathbb{V}[\theta|K = k-1] > \mathbb{V}[\theta|K = k] = \mu^{(2)}_{k,k} - (\mu_{k,k})^2$ which is the desired result. ■

Now, back to the main proof. Since $\mathbb{V}[w(\theta, \theta_2)|\theta, K = k] = \alpha^2\mathbb{V}[\theta_2 - \underline{w}(\theta_2)|\theta, K = k]$, we can apply the Lemma 5 to moments of a function of θ such as $\theta - \underline{w}(\theta)$. Applying this result I can rewrite (17) as

$$\begin{aligned}
&\mathbb{V}[\theta_2 - \underline{w}(\theta_2)|\theta] \\
&= \alpha^2 \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} \int_{\underline{\theta}}^{\theta} [(v - \underline{w}(v)) - \mathbb{E}[\theta_2 - \underline{w}(\theta_2)|\theta, K = k]]^2 \frac{(k-1)F(v)^{k-2}f(v)}{F(\theta)^{k-1}} dv \\
&= \alpha^2 \sum_{k=2}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} \mathbb{V}[\theta_2 - \underline{w}(\theta_2|\theta, K = k]
\end{aligned}$$

Deriving the previous expression with respect to λ we obtain

$$\begin{aligned}
&\frac{\partial \mathbb{V}[\theta_2 - \underline{w}(\theta_2|\theta]}{\partial \lambda} \\
&= -\alpha^2 \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} \mathbb{V}[\theta_2 - \underline{w}(\theta_2|\theta, K = k] + \alpha^2 \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^{k-2}}{(k-2)!} \mathbb{V}[\theta_2 - \underline{w}(\theta_2|\theta, K = k] \\
&= \alpha^2 \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} [\mathbb{V}[\theta_2 - \underline{w}(\theta_2|\theta, K = k+1] - \mathbb{V}[\theta_2 - \underline{w}(\theta_2|\theta, K = k]]
\end{aligned}$$

From Lemma 5 we know that all the differences of conditional variances are negative, except the first one because $\mathbb{V}[\theta_2 - \underline{w}(\theta_2|\theta, K = 1] = 0$. Therefore, the derivative is clearly negative, as desired.

Computational Algorithm

In this subsection, I describe the algorithm I use to solve the model. It is quite similar to the one used in Villena-Roldan (2008), so only the most relevant differences are stressed.

Algorithm 1: Solve Volterra integral equation for reservation wages (14)

It assumes that there is a targeted unemployment rate $\mathcal{U}$, average number of applicants per vacancy λ, separation rate η and probability of full screening χ. With a Coincidence Ranking Equilibrium, the problems of endogenous distribution of the unemployed $F(\theta)$ and the reservation schedule determination are treated separately. Using a grid of points, follow Algorithm 1 in Villena-Roldan (2008) to solve the Volterra integral equation in (26). Using the vector of values representing this function proceed as follows

Step 1 Start iteration j. Enter a guess $\underline{w}^j(\theta)$.

Step 2 Solve the right-hand side of the linear Volterra equation for reservation wages using (14) using a (Gaussian) quadrature rule and linear interpolation of $\underline{w}^j(\theta)$. To compute the order statistic densities use your stored results for $F(\theta)$. In this way, obtain an updated guess $\underline{w}^j(\theta)$

Step 3 If $\|\underline{w}^{j+1}(\theta) - \underline{w}^j(\theta))\| < \epsilon$, stop. Otherwise, start a new iteration at Step 1 using the updated distribution.

Algorithm 2: Targeting mean and variance of log weekly earnings via Newton-Raphson

This algorithm finds the values of the log normal distribution of productivities to match targeted moments of the distribution of log weekly earnings, such as the mean and variance

Step 0 Have an initial guess for the mean and variance of log productivities $\mu_\theta^{(0)}$ and $\sigma_\theta^{2(0)}$. Set a $j = 0$, a tolerance $\epsilon > 0$ and a marginal change to compute numerical derivatives $\varrho > 0$. Set a target sample mean and variance of log wages $\overline{\log w}$ and $S^2(\log w)$.

Step 1 Using the guessed parameters, solve the distribution of unemployed workers $F(\theta)$ and wage reservation schedule $\underline{w}(\theta)$ according to Algorithm 1.

Step 2 Using some quadrature rule, compute the mean and variance of the generated log wage distribution of employed workers using the following moments (for $p = 1, 2$)

$$\mathbb{E}[\log(w(\theta))^p] =$$

$$\int \left(\sum_{k=2}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} \int \log\left(\alpha(\theta - v + \underline{w}(v)) + (1-\alpha)\underline{w}(\theta)\right)^p (k-1)\frac{F(v)^{k-2}}{F(\theta)^{k-1}} f(v)dv \right) f^E(\theta)d\theta$$

$$+ e^{-\lambda} \int \log(\alpha\theta + (1-\alpha)\underline{w}(\theta))^p f^E(\theta)d\theta$$

Then, $\mathbb{V}[\log(w(\theta))] = \mathbb{E}[\log(w(\theta))^2] - \mathbb{E}[\log(w(\theta))]^2$

Step 3 Evaluate these moments again using small changes. First proceed[22] with changes for $\mu_\theta + \varrho$ and then for $\sigma_\theta^2 + \varrho$.

[22]Although it is generally recommended to compute two-side numerical derivatives, one-side numerical ones perform very well and require one-half the function evaluations.

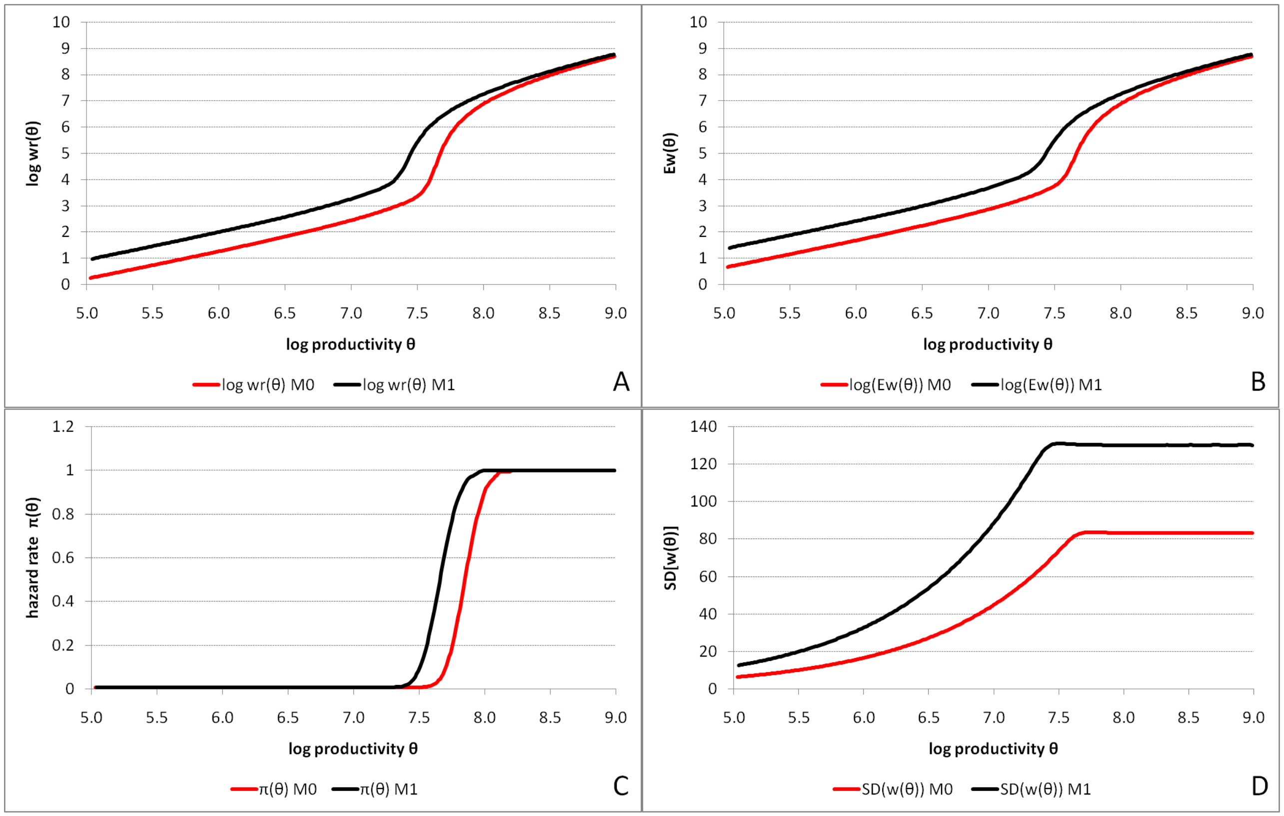
Step 4 Compute the Jacobian matrix D with numerical derivatives and update the guess for μ_θ and σ_θ^2 using $\begin{bmatrix} \mu_\theta^{(j+1)} \\ \sigma_\theta^{2(j+1)} \end{bmatrix} = \begin{bmatrix} \mu_\theta^{(j)} \\ \sigma_\theta^{2(j)} \end{bmatrix} - D^{-1} \begin{bmatrix} \mathbb{E}[\log(w)]^{(j)} - \overline{\log w} \\ \mathbb{V}[\log(w)]^{(j)} - S^2(\log w) \end{bmatrix}$. With the updated guess, go back to Step 1.

Step 5 Having computed $F(\theta)$ and $\underline{w}(\theta)$ to match the targeted moments, compute a consistent vacancy-posting cost κ to satisfy the free entry condition.

$$\kappa = \chi \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} \int \frac{\theta - \mathbb{E}[w|\theta]}{1-\beta(1-\eta)} kF(\theta)^{k-1} f(\theta) d\theta + (1-\chi) \int \frac{\theta - \mathbb{E}[w|\theta]}{1-\beta(1-\eta)} f(\theta) d\theta$$

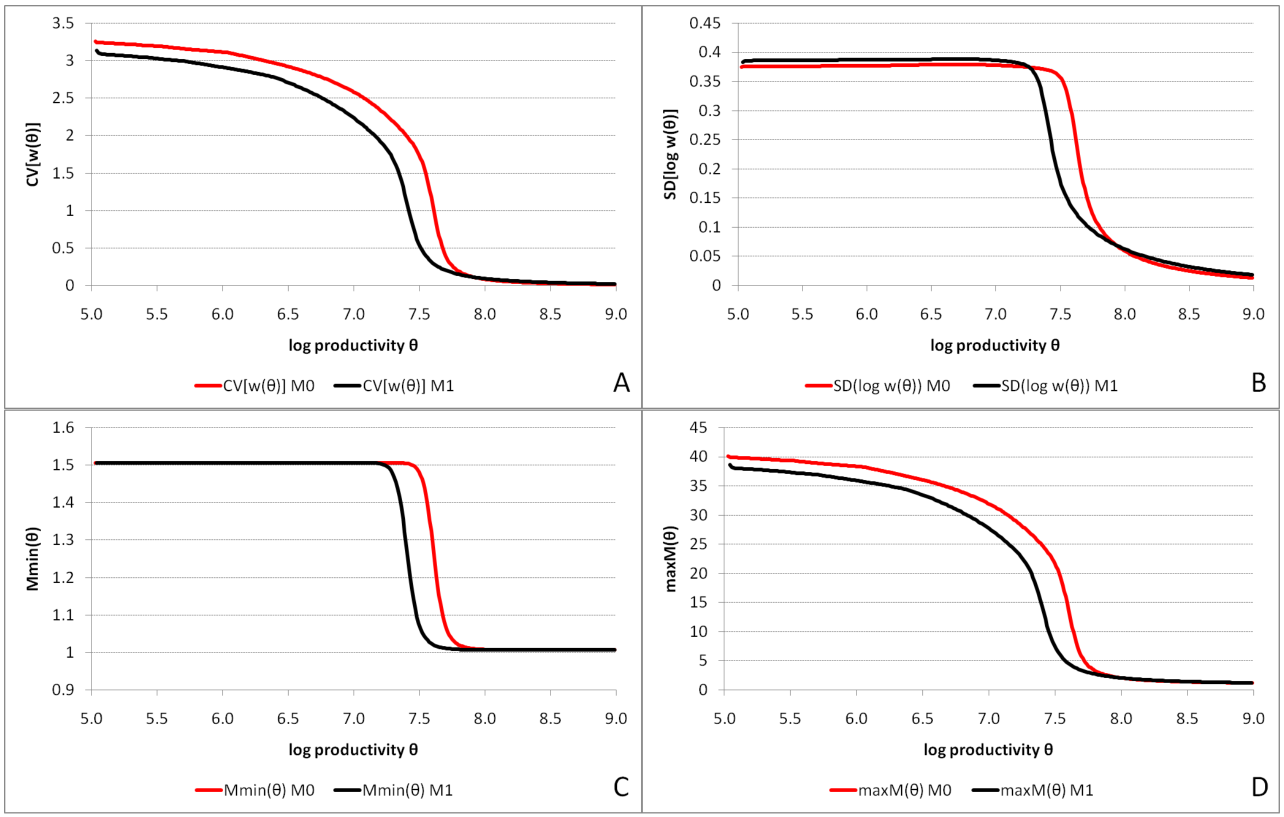
Appendix B: Figures

Figure 1: Results Models Wage Dispersion



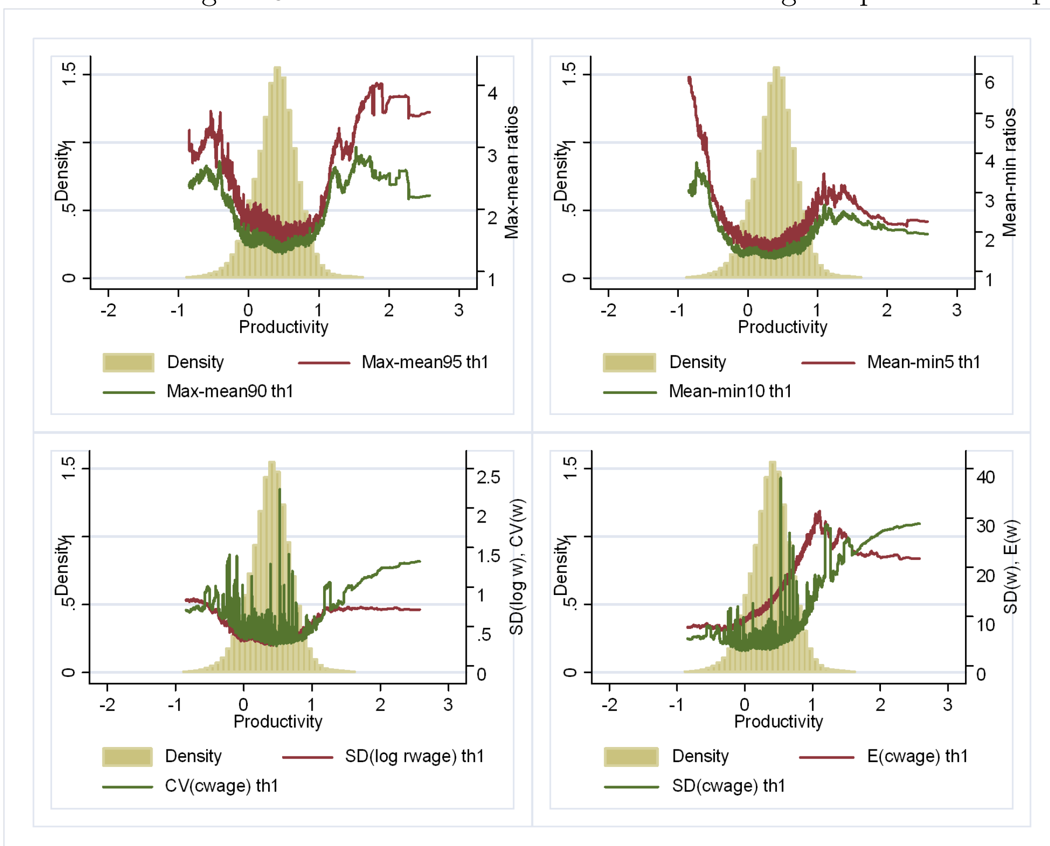
Panel A depicts the log of the equilibrium reservation wage , $\log \underline{w}(\theta)$ schedule for Models 0 and 1. Panel B shows the log of expected wage for both models $\log \mathbb{E}[w|\theta]$. Panel C shows the equilibrium hazard rate of unemployment $\pi(\theta)$. Panel D displays the individual wage dispersion of wage level, $SD[w|\theta] = \sqrt{\mathbb{V}[w|\theta]}$.

Figure 2: Relative Measures of Individual Wage Dispersion



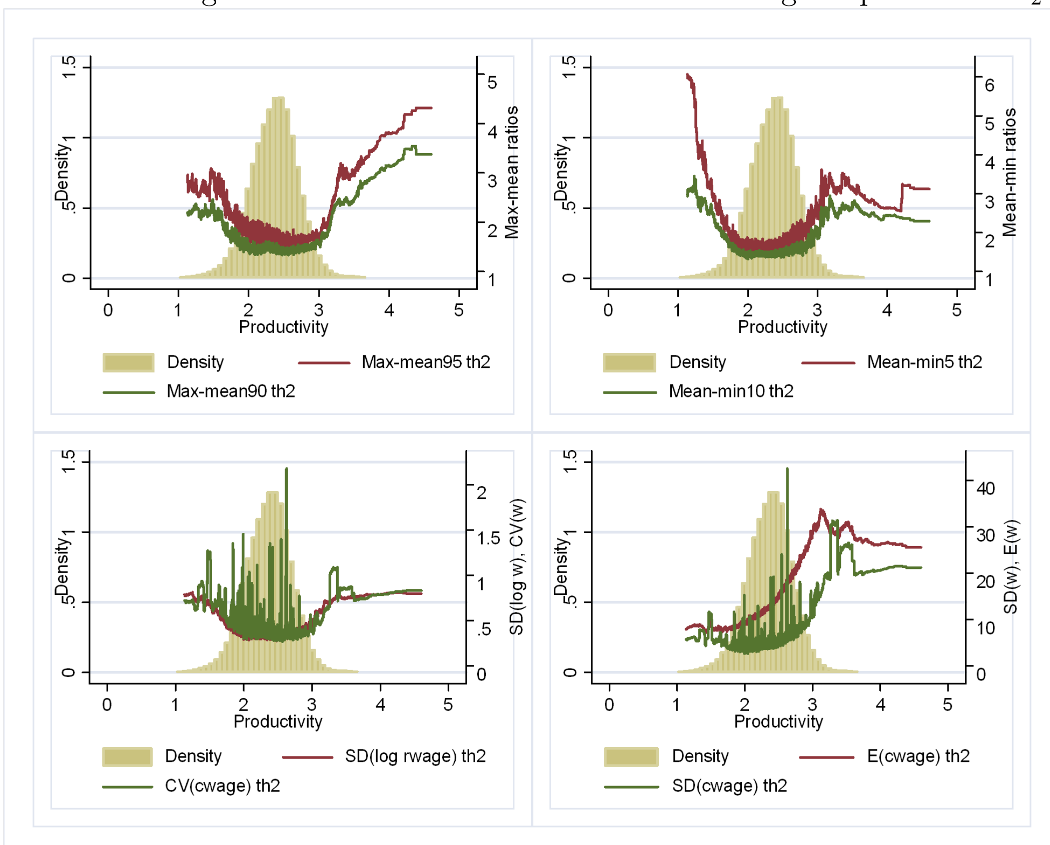
Panel A depicts the coefficient of variation of level of wages conditional on θ for Models 0 and 1; Panel B shows the standard deviation of log wages conditional on θ. Panel C shows the Mean-min ratio conditional on θ. Panel D displays the max-Mean ratio conditional on θ.

Figure 3: Measures of individual residual wage dispersion for $\widehat{\theta}_1$



All "local" statistics are computed using a rolling subsample of the nearest $\sqrt{N}$ observations to all productivities in the sample.

Figure 4: Measures of individual residual wage dispersion for $\widehat{\theta}_2$



All "local" statistics are computed using a rolling subsample of the nearest $\sqrt{N}$ observations to all productivities in the sample.