NBER WORKING PAPER SERIES

PROMINENT JOB ADVERTISEMENTS, GROUP LEARNING AND WAGE DISPERSION

Julio J. Rotemberg

Working Paper 18638
http://www.nber.org/papers/w18638

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2012

Prominent Job Advertisements, Group Learning and Wage Dispersion
Julio J. Rotemberg
NBER Working Paper No. 18638
December 2012
JEL No. D83,J31,J64

ABSTRACT

A model is presented in which people base their labor search strategy on the average wage and the average unemployment duration of people who belong to their peer group. It is shown that, if the distribution of wage offers is not stationary so lower wage offers tend to arrive before higher wage ones, such learning can induce a great deal of wage inequality. An equilibrium model is developed in which firms can choose either to advertise their job openings prominently or not. Prominent ads are assumed to have more influence on more inexperienced job searchers who are less able to identify a multiplicity of viable jobs. Equilibria can then feature groups that learn naively from the experience of their members and accept low wage offers from prominent ads while other groups do not find these offers acceptable. A new test statistic is proposed that measures whether, as predicted by the model, the gains from increasing one's reservation wage are larger than either those that people expect or those predicted by models in which job offers are stationary.

Julio J. Rotemberg
Graduate School of Business
Harvard University, Morgan Hall
Soldiers Field
Boston, MA 02163
and NBER
jrotemberg@hbs.edu

The process of job search entails deepening one's knowledge about potential jobs by, for example, interviewing on site. This research takes time. If being employed limits one's capacity to do this research, unemployed workers have to rely on their subjective beliefs about likely future job prospects when they decide whether to accept a particular offer. But, where do these beliefs come from? A convenient assumption in the literature is that unemployed workers have an accurate mental model of the stochastic process followed by offers, so that the subjective probability they assign to receiving any particular offer at any particular point in time is identical to the true probability of receiving such an offer. The precise mechanism that leads to this correspondence of objective and subjective probabilities is not usually spelled out. Moreover, in an environment in which workers learn about the distribution of wage offers while engaging in costly search, they are likely to accept a job well before their beliefs converge to the truth.[1]

Because workers search for jobs only a few times in their entire lifetimes, their ability to learn about the distribution of wage offers from their own experience is extremely limited. They may thus base their decisions mostly on the information provided by people they know. To focus on this issue, I neglect individual learning altogether and focus on three bits of information that seem likely to be available from peers. First, information about the amount of time that unemployed individuals have spent looking for a job is relatively visible so that it may travel widely within a group. Wage information is both less visible and more sensitive, but individuals may still have a good sense of the average wage earned by the people who are both close and similar to them.[2] Lastly, little should stop people within a group from sharing their beliefs about what the minimally acceptable wage is, so that people should have good information about their group's reservation wage.

This leads me to focus on outcomes that are stable in the sense that people are happy to go along with the group's reservation wage when they believe both that the arrival rate of acceptable jobs is one over the group's average unemployment duration and that the average

---

[1]See Rothschild (1974) for an early and classic proof of a closely related proposition.

[2]For a protocol that allows a group to learn only the mean value of a sample when each member knows one observation, see Abbe *et al.* (2012).

1

wage they will receive by using this reservation wage is the group's mean wage. One rationalization for this is that people believe that they live in a stationary environment where all possible wage offers have constant arrival rates. When this stationarity assumption, which is ubiquitous in the search literature, happens to be true, the stable learning outcome coincides with the outcome in which workers have accurate beliefs (i.e., "rational expectations") about the wage distribution.

When the distribution of wage offers is not stationary, however, group learning of the form I consider need not converge to the outcome with accurate beliefs. I focus on the case where the offers that a worker receives early in her unemployment spell tend to be worse than those she receives later. This nonstationarity is not detectable if, for example, people obtain wage information from a different source than they obtain unemployment spell information. Moreover, it seems plausible to suppose that wage information comes from employed friends, who may mostly have forgotten the length of their past unemployment, while information about unemployment spells originates mostly from more distant individuals.[3] I thus consider situations where a group of workers continues to treat the facts to which they have access as coming from a stationary environment. The key finding of the paper is that the presence of two distinct groups, one that learns in this way while the is more sophisticated, can lead to wages that are substantially more dispersed than those that result from accurate beliefs. The reason for this is that the group that learns naively accepts the low wages that are available early in their unemployment spell. Newly unemployed members of the group thus expect wages to be low and accept these jobs as well.

More sophisticated workers wait for higher wages and learn from their group that it is rational to do so. A crucial assumption, then, is that groups do not learn from each others experience. While I do not pursue the root causes for this in detail, it is easy to imagine settings where the group that earns little miss-attributes the high earnings of the other group by blaming the difference in earnings, for example, on favoritism.

---

[3]Professional surveyors faces the same recall issues so that the best and most extensive sources of wage data for researchers do not coincide with the best and most extensive sources of data on unemployment spells.

A question this raises is whether there are any good reasons to expect early offers to be relatively less generous than later ones. I provide an answer to this question that is based on the inexperience of job searchers as they begin their spell of unemployment. I suppose that this inexperience limits the sophistication with which they process job advertisements. They thus spend time investigating ads that are highly visible or striking but which have a good chance of being inappropriate for them. With experience, searchers become less likely to be distracted by these "prominent" ads, and more likely to focus on ones that are suitable for them. I also imagine that workers are investigating several possible jobs at the same time, so that they may have access to more than one job offer in any given period.[4] Because inexperienced searchers are more likely to focus on dead ends, this is less likely for them than for experienced searchers.

Relative to firms posting more ordinary jobs, those posting prominent offers may thus reach fewer appropriate workers but the appropriate workers that they do reach are less likely to have competing offers. I demonstrate that this reduced competition leads firms posting prominent jobs to offer a distribution of wages that is stochastically dominated by the distribution of wages offered by less prominent job ads. This occurs even if all workers have rational expectations, though the effect can be magnified if a subset of workers draws incorrect (but plausible) inferences from the experience of their peer group.

It is worth asking at this point whether the idea that more prominently displayed jobs have lower wages has any empirical validity. The inter-industry wage observations of Katz and Summers (1989) are somewhat consistent with this. They show that retail employees, and particularly employees of eating and drinking establishments, have low wages even within narrowly defined occupations. In particular, Katz and Summers (1989) report that the lowest wages for janitors are earned in the "Eating and Drinking" industry, with "Other Retail Trade" not being far behind. These jobs are often advertised on the premises, and these advertisements have large audiences. Janitors that work in Banking or Insurance, industries where help wanted advertising on the premises is more unusual, earn more.

---

[4]This assumption has been used earlier by Lang (1991).

The question of whether imperfect information about the identity of firms that pay particular wages can lead to a realistic degree of wage dispersion has received a great deal of attention. Mortensen (2003) summarizes an extensive literature saying that it does. In an important recent paper that the current one follows closely, Hornstein, Krusell, and Violante (2011) (HKV) note that implausible parameters are needed in these models to obtain realistic levels of wage dispersion. They focus on the implications of search models for the ratio of the mean wage to the minimum wage that is paid/accepted in the market and show that, for empirically plausible parameters, the baseline search model implies that this ratio should equal 1.036. When they turn to Census observations, and after using a variety of controls including occupation and geographic area, the observed ratio is between 2.5 and 3. Using Census as well as other data, they estimate ratios of the mean wage to the 10th percentile (which is less affected by measurement error than the minimum) that hover around 1.6 - 1.7. I follow them both by computing implications of my model for these ratios and by adopting most of their model parameters. The paper can thus be read as an attempt at using social learning and the differential prominence of different job offers as explanations of the puzzle uncovered by HKV.

The paper implies that, for some people, there are outsize rewards from continuing to search. These rewards are higher both than the rewards that these individuals expect and than those that would be predicted if wage offers were in fact stationary. These two predictions turn out to be testable if one has access to the wage and unemployment duration outcomes of people with different reservation wages. In effect, one has to compare the "returns to delay" (which depend on the size of the wage gains of people with higher reservation wages to the extra amount of time they spend searching) to the ratio of the mean to the minimum wage.

For young white male workers, the returns to delay implied by the estimates in Holzer (1986) contradict the stationarity of job offers. This test is derived by ignoring on-the-job search, which HKV regard as having the potential to resolve the puzzle of high mean to minimum wages. The test statistic I propose, however, would actually yield stronger rejections

of stationarity if people were able to raise their wages by engaging in on-the-job search. To evaluate rational expectations one needs to know not only people's reservation wage but also the wages they expect to receive. If one supposes that Lancaster and Chesher's (1983) survey provides an accurate measure of the mean ratio of expected wages to reservation wages, the evidence in Holzer (1986) also contradicts rational expectations for both white and black young male workers.

The paper proceeds as follows. Section 1 studies group learning when the distribution of wage offers is exogenous. Its first subsection considers the standard case where this distribution is stationary while the second one lets early offers be less attractive than later ones. The stability of an outcome in which two groups have different average wages because they have different beliefs is demonstrated for this latter case.

In the rest of the paper, starting with Section 2, the distribution of wages is endogenous and is the result of letting firms post wages in a discrete time model whose structure borrows heavily from Burdett and Judd (1983). To allow comparison with the existing theoretical literature, the first subsection of Section 2 makes all offers equally visible to all unemployed individuals. Subsequent subsections, by contrast, distinguish between prominent and other offers and let firms decide which type of offer they wish to make. Section 2.2 studies the case where everyone has accurate beliefs while Section 2.3 focuses on the case where everyone belongs to a group that learns naively from the experience of other members. Section 2.4 demonstrates the existence of stable outcomes where one group is naive while another is sophisticated, and shows that the resulting equilibria can display a great deal of inequality. Section 3 derives the statistic that allows one to test the stationarity of offers and the rationality of expectations. Section 4 considers an extension to study whether the most productive firm necessarily pays the highest wage, while Section 5 offers some concluding remarks.

Related Literatures

This paper combines assumptions that are made in three distinct theoretical literatures. First, there is fairly large literature on individualistic learning in search models, of which

Nishimura and Ozaki (2004) provide a recent example. This literature neglects the possibility of learning from peers and I am unaware of examples in which the equilibrium wages posted by firms are incorporated into such models. A different departure from accurate beliefs in search is provided by Rotemberg (2002), who also offers a theory of equilibrium wage determination that can lead to inequality. However, Rotemberg's (2002) model does not explicitly involve group learning.

Second, there is a literature that, like this paper, starts from the assumption that people have a network of peers and investigates the effect of these connections on individual wages. The more extensive branch of this literature supposes that all individuals have the same knowledge about the wage distribution while peers are uniquely capable of referring each other to jobs. Two prominent examples are Montgomery (1991) and Calvó-Armengol and Jackson (2004). Given these assumptions, it is not surprising that these papers find that peers can only help improve an individual's wages, whereas here that is not the case. A case where groups provide guidance but not referrals, and which is thus closer to what is studied here, has been studied by Ahmed (2008). He considers a situation with directed search where some groups convince their members that they will be discriminated against and shows that there is an equilibrium where members of this group only apply to jobs with low wages. One important difference between Ahmed (2008) and the current papers is that he does not let groups learn because he considers only an extreme case in which different groups apply to wholly distinct jobs.

Third, there is literature of search in product markets that considers the case where some firms are systematically sampled before others. Both in Armstrong, Vickers and Zhou (2009) and Haan and Moraga-Gonzalez (2011), the firm that is sampled first offers a particularly good deal (a low price in the product search context). This is the opposite of what occurs here. One important difference between these models and the current one is that they suppose that products are imperfect substitutes in the eye of the consumer. Given perfect recall, this means that the firm that is sampled first in Armstrong, Vickers and Zhou (2009) faces a more elastic demand than other firms. Similarly, the result in Haan and Moraga-

6

Gonzalez (2011) is due to a selection effect in which customers who are dissatisfied by the product of the prominent firm reveal themselves to have an inelastic demand for the product of the second (and only other firm).

1 Group Learning with Exogenous Wages

Consider an unemployed worker in a setting where time is discrete, the worker is risk neutral and discounts next period's revenue using the factor ρ. Each period that this worker is unemployed, she investigates job postings. This time consuming research is needed for the worker to learn the characteristics of the job, including its wage w. The non-wage characteristics of jobs make them unsuitable for a randomly selected subset of workers. For simplicity, I treat all those that are suitable as having the same non-pecuniary benefits. In a given period, the worker may discover more than one suitable job, and would then prefer the one with the highest wage. At the end of the period, the unemployed individual must decide whether to accept any of the suitable jobs she has found. If she does, she earns the job's posted wage w in every period that she works. It is common knowledge that people who are employed at a particular point in time can cease to have their job at the beginning of the next period with probability σ. At that point, they become unemployed again.

People belong to groups, and groups have a common perception of the present value of earnings that an unemployed individual can expect to receive. I denote this perceived present value by $\tilde{U}(\tilde{w}, \bar{w}, \hat{\lambda})$. The value of $\tilde{U}$ depends on the three bits of information which, as discussed in the introduction, members of groups have access to. These are $\tilde{w}$, the reservation wage of her peers, $\bar{w}$, the average wage earned by them, and $\hat{\lambda}$, an estimate of the probability that a viable job paying at least $\tilde{w}$ arrives in any given period. Often, $\hat{\lambda}$ is simply the inverse of S, the average length of the group's unemployment spells, since this is the natural estimate of $\hat{\lambda}$ under stationary beliefs.[5] Indeed, a key characteristic of $\tilde{U}$ is that,

---

[5]Observations of long spells are more often incomplete than observations of short spells, and this may affect people's ability to infer the true average length of unemployment spells S (as is true for econometricians). An extension might thus consider a more complicated estimation exercise for individuals with more limited data about their group.

under the assumption that job offers are stationary, it is the best estimate of the present value of earnings for an unemployed individual that uses the group's reservation wage of $\tilde{w}$. If they believed this stationarity assumption, members of the group would agree with the group's $\tilde{U}$.

An unemployed worker is also able to form an estimate of $\tilde{V}(w)$, the expected present value of revenues she expects the moment she starts working at a job that pays w. If the job does not terminate after one period, the wage w is earned again and the present value of earnings is $\tilde{V}(w)$ once more. If it does terminate, the group would expects the individual to earn a present value of $\tilde{U}$ at this point. Accepting this expectation as her own, $\tilde{V}(w)$ obeys

$$\tilde{V}(w) = w + \rho(\sigma\tilde{U} + (1-\sigma)\tilde{V}(w) = \frac{w + \rho\sigma\tilde{U}}{1 - \rho(1-\sigma)}, \tag{1}$$

where the second equality follows from the first.

The key assumption regarding individuals that learn from their groups is that they accept all viable jobs for which

$$\tilde{V}(w) \geq \tilde{U}(\tilde{w}, \bar{w}, \hat{\lambda}) \qquad \text{or} \qquad w \geq w^* \equiv (1-\rho)\tilde{U}(\tilde{w}, \bar{w}, \hat{\lambda}), \tag{2}$$

and turn down the jobs for which these inequalities are strictly reversed. There are two equally valid interpretations for members' use of $(1-\rho)\tilde{U}$ as their reservation wage w^*. The first is that people agree completely with the calculations of the group, so they too regard the distribution of job offers as stationary. The second is that, even when they are uncertain whether this is true, they feel group pressure to accept jobs for which (2) is true and pressure to turn down jobs for which it is false.

Now turn to the group's calculation of $\tilde{U}$ itself. A worker that is unemployed in a particular period receives a flow of utility b in that period. The group expects that, with probability $\hat{\lambda}$, she will land a viable job that pays $\bar{w}$ on average as long as the worker uses the group's reservation wage $\tilde{w}$. Since $\tilde{V}(w)$ is linear in w, the expected present value of a viable job under this strategy is $\tilde{V}(\bar{w})$. Thus, under stationary beliefs, the group's $\tilde{U}$ is

$$\tilde{U}(\tilde{w}, \bar{w}, \hat{\lambda}) = b + \rho(\hat{\lambda}\tilde{V}(\bar{w}) + (1 - \hat{\lambda})\tilde{U}) = \frac{Db + \rho\hat{\lambda}\bar{w}}{(1 - \rho)(D + \rho\hat{\lambda})} \qquad \text{where } D \equiv 1 - \rho(1 - \sigma), \ (3)$$

and the second equality is obtained by using (1) to substitute for $\tilde{V}(\bar{w})$ and solving the resulting equation for $\tilde{U}$. It is important to emphasize that this solution for $\tilde{U}$ is being computed under the supposition that the worker follow her group's strategy of using the reservation wage $\tilde{w}$. The analysis does not require group members to have views on the present value of unemployment for any other reservation wage.[6]

If, however, the actual reservation wage w^* in (2) differs from $\tilde{w}$, the group reservation wage cannot be expected to remain equal to $\tilde{w}$. This leads to a natural definition of stability for outcomes that result from group learning. In particular, w^* is a stable value of $\tilde{w}$ if

$$\tilde{V}(w^*) = \tilde{U}(w^*, \bar{w}, \hat{\lambda}) \qquad \text{or} \qquad w^* = (1 - \rho)\tilde{U}(w^*, \bar{w}, \hat{\lambda}), \qquad (4)$$

In other words, accepting jobs with $w \geq w^*$ is stable under group learning if the mean wage and hazard rate of employment that is generated when the group uses the reservation wage w^* justify the use of w^* as the reservation wage.[7] Using (3), w^* satisfies

$$w^* = \frac{Db + \rho\hat{\lambda}\bar{w}}{D + \rho\hat{\lambda}}. \qquad (5)$$

1.1 Stationary Wage Offers

Groups in this model implicitly assume that the probability of finding a viable offer is constant through time and that the distribution of the most favorable wage available at any given point in time is time invariant as well. It is perhaps not surprising that, if both of these premises are valid, the stable group learning outcome is identical to the unique outcome that

---

[6]Calculating the present value of being unemployed for alternative reservation wages requires more detailed wage data than just the mean wage for the group. Because wage data is so sensitive, estimates of the group's mean wage may be much more reliable than more fine-grained estimates of the group's wage distribution.

[7]This stable outcome is "self-confirming" in the Fudenberg and Levine (1993) sense of not being inconsistent with what agents observe.

would obtain if people had a correct assessment of the true underlying stochastic process for job offers. This is shown in the next proposition.

Proposition 1. *Suppose that the probability that an unemployed worker will find a viable offer is λ_u per period and that the highest wage among the viable offers that the unemployed worker gets in any given period has a cumulative density function $G(w)$. If workers know λ_u and G, they accept viable offers if and only if they pay at least w^* in (5) where $\hat{\lambda} = \lambda_u(1 - G(w^*))$ while $\bar{w}$ is the mean of $G(w)$ conditional on w being greater than or equal to w^* so that it equals*

$$\bar{w} = \frac{\int_{w^*} w \, dG(w)}{1 - G(w^*)}. \tag{6}$$

Just like the rational expectations optimum to which it is equal, the stable group learning outcome is unique in this case.

Therefore, group learning leads to conventional outcomes when wage offers are stationary as they are in much of the theoretical labor search literature starting with McCall (1970). As shown by HKV, the resulting dispersion of wages is quite small when plausible parameters are used and inequality is measured by the ratio of mean to minimum wages. To make this point transparent, they suppose that b is equal to $\gamma \bar{w}$ for all workers, and I follow them in this assumption. Equation (5) then implies that

$$\frac{\bar{w}}{w^*} = \frac{D + \rho \hat{\lambda}}{D\gamma + \rho \hat{\lambda}} \tag{7}$$

On the basis of economy-wide average separation rates, job finding rates and real interest rates HKV set σ, $\hat{\lambda}$ and ρ to .02, .39 and .9959 respectively. Lastly, they set γ equal to .4 on the basis that smaller numbers both create problems for the model as a model of economic fluctuations and are fairly implausible given the rules governing unemployment insurance. These numbers imply that the ratio of mean to average wages $\bar{w}/w^*$ equals 1.03.

While keeping the same parameters for comparability, the next subsection shows that group learning can lead to considerably more dispersed wages if the arrival rate of high-wage jobs is delayed relative to the arrival rate of low-paying jobs. Under stationary arrivals,

10

the probability of obtaining an offer above w equals $\lambda_u(1 - G(w))$, which is declining in w. Thus, individuals with higher reservation wages must wait longer on average to obtain jobs. It might thus be thought that this assumption already implies that "higher wages come later." On the other hand, the stationary arrivals assumption predicts that, for a common reservation wage, the average wage of people who find jobs early is the same as the average wage of people who find jobs later. I drop this aspect of the stationary arrivals assumption from now on.

1.2 High Wage Offers Start Arriving in Period 2

To depart as little as possible from standard models of labor search, unemployed workers continue to have a probability λ of finding a viable job in every period. The distribution of offers is non-stationary, however. All viable job offers found by workers in their first period of unemployment carry a wage of w_1 while, in subsequent periods, viable jobs pay a wage $w_2 > w_1$ with probability η and pay w_1 with the remaining probability. The total population is normalized to equal one and all individuals have the same tastes and productivity levels.

If unemployed workers knew the stochastic process followed by job offers, they would realize that their decision environment is stationary in the sense that the probability distribution of *future* job offers is the same at every node at which workers decide whether to accept an offer or not. The distribution of wages in the first period is the only source of non-stationarity, and since all decisions are made after this, it has no effect on rational decision-making.

Except in the uninteresting case in which $b > w_2$, unemployed workers all accept offers of w_2 since there are no higher wages worth waiting for. The critical issue is whether any worker also accepts an offer of w_1. This decision depends on whether, as in (2), w_1 is greater or smaller than $(1 - \rho)\tilde{U}$, where $\tilde{U}$ represents the worker's expectation of the value of remaining unemployed and accepting only jobs that pay w_2. This decision thus depends on $\hat{\lambda}$. For individuals with accurate beliefs, this finding rate equals the hazard of finding jobs that pay w_2 from period 2 onward, which is $\lambda\eta$.

Now consider members of a group that has used the reservation wage w_2 in the past. The average length of unemployment spells for these individuals is $S = 1 + 1/\lambda\eta$. If the group were naive, and treated the environment as stationary, its value of $\hat{\lambda}$ would equal $1/S$, which is smaller that $\lambda\eta$. If the group were more sophisticated and noticed that no one found a job in period 1, it would alert its members that the constant arrival rate of jobs that pay w_2 starts in period 2. The group would then infer that the arrival rate of jobs is $1/(S-1)$, which would lead its members to conclude, correctly, that it equals $\lambda\eta$.

Many of the results below hinge on some groups being naive, so that they use the group's average wage and average length of unemployment spells to estimate the parameters of a stationary model. As discussed in the introduction, this naiveté may be reasonable in situations where people obtain their estimate of the mean wage from people they know closely and these, consistent with the parameters used below, have mostly been employed for some time. If the passage of time erases people's memory of the length of their unemployment spells, the close friends of unemployed people are not helpful in determining the joint distribution of wages and the length of these spells.

On the other hand, information about how long people who belong to one's own group have been unemployed is both more visible and less sensitive so that it may travel relatively widely within a group. For the most part, I suppose that the way this information travels is in condensed form, so that people know the average length of unemployment. If, however, there is a group that never accepts a job in the first period, this may become known to the group as well. As a result, the group may become sophisticated in the sense of concluding that viable offers start arriving in the second period. This non-stationarity becomes more difficult to discern if people start accepting offers in the first period. I start the analysis of nonstationary offers by studying the case where both w_1 and w_2 would be acceptable if people had accurate information.

Proposition 2. *Individuals with accurate beliefs have a reservation wage of w_1 if and only*

if

$$w_1 \geq \frac{Db + \rho\lambda\eta w_2}{D + \rho\lambda\eta}. \tag{8}$$

When this condition is met, and whether individuals are naive and treat the environment as stationary or sophisticated, there is no stable group learning outcome in which a group turns down offers of w_1.

This shows that the level of sophistication of group learners is not important when (8) holds. The reason is that, even if a group that learns naively were to turn down offers of w_1, the appeal of such offers to its unemployed members would not be reduced. In a sense, naive group learners obtain an accurate assessment of what happens if their members only accept high wages. Matters are different when (8) is violated. In particular

Proposition 3. *Suppose* (8) *is violated. As long as $\lambda < 1$, values of w_1 can be found for any $w_2 > b$ such that*

$$\frac{Db + \rho\lambda\eta(1-\lambda)w_2}{D + \rho\lambda\eta(1-\lambda)} \leq w_1 \leq \frac{Db + \rho\lambda\eta w_2}{D + \rho\lambda\eta}. \tag{9}$$

At these values there is a naive stable group learning outcome with a reservation wage w_1 even though individuals with accurate beliefs have a reservation wage of w_2. There is also a sophisticated group learning outcome with a reservation wage of w_2.

Proposition 3 shows that there exist values of w_1 that workers with rational expectations would turn down but that naive group learners would accept. This is possible because, once naive group learners accept low wages in the first period, their average wage falls. This leads naive group learners to overestimate the fraction of offers that has low wages *from the second period onward* so that they find offers of w_1 more attractive.

Even if a group of workers that learns naively accepts a wage w_1 that is significantly lower than w_2, (7) still keeps wage inequality among members of the group low. As long as there is also a group of workers that is more sophisticated, however, overall inequality can be large if the two groups have different average wages. To see this, suppose that (8) is violated and that a group of size N^L accepts offers with a wage of w_1 while a group of size

$N^H = (1 - N^L)$ does not. I refer to the members of the former as being of "type" L, while the members of the latter are of type H. Note that types differ only in the group whose experience they use to draw inferences about the labor market. The overall average wage in this economy is then

$$\bar{w} = w_2 + N^L[1 - (1 - \lambda)\eta](w_1 - w_2). \tag{10}$$

Thus, the ratio of the average wage $\bar{w}$ to the minimum wage w_1 is increasing in w_2/w_1. As in HKV, the size of w_2/w_1 is limited by the need to ensure that offers of w_1 are accepted, though the limit is quite different here.

Since individuals of type L accept lower wages, we can expect them to receive lower unemployment insurance payments on average, and this depresses their b relative to that of individuals of type H. To simplify the analysis, I follow HKV and do not let the level of an individual's b depend on his personal employment history. Instead, like HKV, I suppose that b equals a proportion γ of average wages. Because there are two distinct groups, however, I let the b of each type equal γ times the average wage of her group. Aside from its simplicity, the main aim of this assumption is to preserve comparability with HKV. With this assumption, the maximum value of w_2/w_1 that is consistent with the violation of (8) is

$$1 + \frac{D(1 - \gamma)}{(D\gamma + \rho\lambda)(1 - \lambda)\eta}.$$

Using this to substitute for w_2/w_1 in (10), the maximum value of $\bar{w}/w_1$ is

$$1 + \frac{D(1 - \gamma)}{(D\gamma + \rho\lambda)(1 - \lambda)\eta}\Big[1 - N^L[1 - (1 - \lambda)\eta]\Big].$$

As in HKV, an increase in γ lowers this maximum ratio because it implies that workers obtain a higher fraction of the mean wage while unemployed. Thus, the minimum wage they accept is closer to this mean. What is more novel is how this maximum depends on λ. The derivative of the expression above with respect to λ is

$$\frac{D(1 - \gamma)}{(D\gamma + \rho\lambda)(1 - \lambda)\eta}\left[-\frac{\rho}{D\gamma + \rho\lambda} - \eta N^L + \frac{1}{1 - \lambda}\right]$$

The first term inside the square brackets is negative and captures the effect emphasized by HKV, namely that an increase in the job finding rate leads workers to demand a higher wage

14

since they have less to lose from unemployment. The second term, which is also negative, is a composition effect. A higher λ leads more workers of type L to accept offers that pay w_1 in the first period, so that a smaller fraction of workers earns w_2 and the mean shifts towards w_1. The last term is positive, and becomes dominant for large λ. It captures the idea that a high λ leads more workers of type L to match with a job that pays w_1 in the first period and thereby reduces their estimate of the likelihood that they will obtain w_2 by waiting. It thus leads them to be more willing to accept a lower w_1. In the limit in which λ equals one, w_2 is irrelevant to their decision so that it can be arbitrarily high relative to w_1.

The expression above is positive only if λ is sufficiently high. This raises the question of whether such high values of λ are inconsistent with the properties of the aggregate hazard rate of exiting unemployment. In the case where all workers have accurate beliefs and w_1 is the lowest wage that workers accept, this hazard rate equals λ itself. A very high value of λ would then be incompatible with the fact that about half the people who become unemployed do not find a job during their first month of unemployment.

When groups differ in their beliefs, the computation of aggregate hazard rates is more complex. This is true in part because the willingness of workers of type L to accept lower wages implies that they leave unemployment more quickly. This implies that people of type L have a higher employment rate, which in turn means that a disproportionate number of the newly unemployed are of this type. To show this formally, let Ψ be the fraction of people who become unemployed in any given period who are of type L. Let M^i represent the number of people of type i who are employed in steady state, U^i the corresponding number of unemployed people, and U^{i1} the number of people of type i who are unemployed in the current period but were employed in the last one. For both types U^{i1} is equal to σM^i. Thus Ψ, which is defined by $U^{L1}/(U^{L1} + U^{H1})$, equals $M^L/(M^L + M^H)$.

For people of type L, the outflow from unemployment equals λU^L, and this must equal the inflow σM^L. Given that $U^L + M^L = N^L$, the equality of λU^L and σM^L implies that

$$M^L = \frac{\lambda}{\lambda + \sigma} N^L \qquad\qquad U^L = \frac{\sigma}{\lambda + \sigma} N^L. \qquad (11)$$

Individuals of type H start receiving acceptable offers only in the second period of unemployment so that the outflow from unemployment equals $\lambda\eta(U^H - U^{H1})$, where U^{H1} equals the inflow into unemployment σE^H. Since $U^H + M^H = (1 - N^L)$, we have

$$M^H = \frac{\lambda\eta(1 - N^L)}{\lambda\eta + \sigma(1 + \lambda\eta)} \qquad\qquad U^H = \frac{\sigma(1 + \lambda\eta)}{\lambda\eta + \sigma(1 + \lambda\eta)}. \tag{12}$$

After some rearranging, Ψ thus equals

$$\Psi = \frac{\lambda(\sigma + \lambda\eta(1 + \sigma))N^L}{\lambda(\sigma + \lambda\eta(1 + \sigma)) - \sigma\lambda(1 - \eta(1 - \lambda))(1 - N^L)},$$

which has been written so that it is clear that it is larger than N^L.

The number of individuals of type L that are newly unemployed in any given period equals $\Psi(U^{L1} + U^{H1})$. Because these people have a probability λ of accepting a job in each period that they are unemployed, a fraction $(1 - \lambda)^{\tau-1}$ of them also experience a τ'th period of unemployment, while a fraction $\lambda(1 - \lambda)^{\tau-1}$ finds a job during this τ'th period.

Similarly, $(1 - \Psi)(U^{L1} + U^{H1})$ individuals of type H become newly unemployed in a given period. None of them accepts a job in their first period of unemployment and, afterwards, those who are unemployed have a probability $\lambda\eta$ of accepting a job. Thus, for $\tau \geq 2$, a fraction $(1 - \lambda\eta)^{\tau-2}$ experiences a τ'th period of unemployment and a fraction $\lambda\eta(1 - \lambda\eta)^{\tau-2}$ accept jobs in that period.

To an outside observer who does not distinguish between these groups, the "hazard" of finding a job in the τ'th period of being unemployed is the fraction of the $U^{L1} + U^{H1}$ who become unemployed in any given period that find jobs in the τ'th period of their unemployment divided by the fraction that was still unemployed in their τ'th period. It thus equals $\lambda\Psi$ in the first period. For $\tau \geq 2$ it equals

$$\frac{\lambda(1 - \lambda)^{\tau-1}\Psi + \lambda\eta(1 - \lambda\eta)^{\tau-2}(1 - \Psi)}{(1 - \lambda)^{\tau-1}\Psi + (1 - \lambda\eta)^{\tau-2}(1 - \Psi)}. \tag{13}$$

Since $(1 - \lambda)^{\tau-1}$ becomes negligible faster than $(1 - \lambda\eta)^{\tau-2}$, this converges to $\lambda\eta$ as τ gets large though it differs from this exact value unless λ is equal to one. There is thus only one case in which the hazard is constant, and equal to η. This is when λ equals one while $\eta = \Psi$.

Insofar a constant hazard is a good approximation to what is observed, a high value of λ appears justified.[8] Admittedly, the case where $\eta = \Psi$ is somewhat arbitrary since there is no force in the model that forces the fraction of workers who accept low wages to correspond to the fraction of jobs that offer w_2 from period 2 onwards.

2 Equilibrium in a Discrete Time Wage Posting Model

The inequality obtained in Section 1.2 hinges on wage offers being lower in the first period than they are subsequently. The plausibility of this requirement may not be immediately obvious, particularly because firms play a role in setting wages and it may be doubted that they would not compete by raising the wages that individuals see when they first look for work. To respond to this concern, this section derives such an outcome in an equilibrium where firms can pay any wage that they wish. The source of the rise in wage offers over time is that, when they first become unemployed, workers are more susceptible to visible, and likely inappropriate, opportunities. Firms whose jobs are prominent then have an incentive to offer low wages. For comparability with the literature, this discrete time model is first analyzed under the assumption that all offers are equally visible at all times and that workers have accurate beliefs.

2.1 Stationarity with a Single Type of Wage Offer

Firms have productive opportunities in which the marginal product of labor is a constant R and can post job offers at a cost of c. The total number of job postings equals v. Once a worker accepts a posting, he becomes employed. As before, employees face a probability σ of becoming unemployed in the next period. The value to a firm of having an employee who accepted a posting that promised a wage w is

$$\Pi(w) = R - w + \rho(1-\sigma)\Pi(w) = \frac{R-w}{D}. \tag{14}$$

---

[8] See Machin and Manning (1999) for a discussion suggesting that the evidence against a constant hazard rate, once one controls for individual heterogeneity, is weak.

Following the idea of Burdett and Judd (1983), unemployed workers have a probability $\lambda(1-\delta)$ of identifying a single viable job and a probability $\lambda\delta$ of identifying two. One way of thinking about this is that most job offers are inappropriate for most workers so that workers must pore over ads and decide which to explore at more length. In each period, they have time to do this for only a limited number of jobs so that a period can end without finding any attractive jobs. Because the search process is stretched over time, one can think of the worker as following various exploratory steps in parallel for several jobs so that the period can end with the worker having identified more than one opportunity.

This particular interpretation of the model supposes that firms are passive and willing to take all comers. However, as discussed in Lang (1991), the idea that workers are somewhat likely to have several job offers at once does not require this. What it does require is the parallel pursuit of several leads and the assumption that firms cannot make exploding offers that disappear unless the worker accepts them instantly. In practice, workers can usually claim that there are aspects of the job that they are uncertain about even after they receive a formal offer so that they are given some time to decide. Since applications are being pursued simultaneously, another job may become available while the worker is still thinking about the first. The idea that workers have some probability of having access to competing offers thus seems applicable beyond the model considered here. The current model with passive job posting firms captures this idea in a simple manner but deserves to be extended to the case where each firm also gets to decide how many offers to extend to its pool of applicants.

Now consider a particular unemployed worker-offer pair and suppose for the moment that the offer has a wage that exceeds the worker's reservation wage. The probability that this particular offer is the single one that this particular individual is willing and able to accept is $\lambda(1-\delta)/v$. At the same time, the probability that this offer is one of two offers that this individual deems viable is $2\lambda\delta/v$. So, for a posting firm, the probability of facing a competing offer conditional on having its own offer deemed viable by a worker is $2\delta/(1+\delta)$, while the probability of not facing a competing offer conditional on its own offer being acceptable is $(1-\delta)/(1+\delta)$.

Given U unemployed workers, the expected number of workers that explore a particular offer and find it viable is Q/v, where Q is the total expected number of viable contacts made by all offers and equals $\lambda(1 + \delta)U$. If a worker finds two viable offers that pay the same wage, he chooses one at random, whereas he chooses the one that pays more if their wages are different. An equilibrium requires the expected present value of profits of a firm that makes an offer be independent of w for any w that is actually posted. Burdett and Judd's (1983) argument then implies that the distribution of offered wages has neither mass points nor holes. If it had a mass point at a wage w below R, a firm would be able to raise profits by raising its wage slightly above w because this would cause a jump in the probability of having its offer accepted. Moreover, paying R leads to negative profits since $c > 0$. Similarly, if the distribution had a hole, a firm could increase its profits by lowering its wage slightly from the upper bound of the hole, and this would not lower its probability of having its offer accepted.

Let $F(w)$ represent the cdf of wages across job postings, so that a fraction $F(w)$ of them offer less than w. Given the analysis so far, and taking $F(w)$ as given, the expected present value of the profits that accrue from offering a job that pays w is

$$\frac{\lambda U}{Dv}\Big(1 - \delta + 2\delta F(w)\Big)(R - w) - c.$$

A free entry equilibrium requires that cv adjust so that this is zero for any w in the support of F and negative for wages outside this support. It follows that, in such an equilibrium,

$$F(w) = \frac{\theta}{2\delta(R - w)} - \frac{1 - \delta}{2\delta} \qquad \text{where} \qquad \theta \equiv \frac{cvD}{\lambda U}. \tag{15}$$

The minimum wage w^* and maximum wage w^m implied by this satisfy

$$w^* = R - \frac{\theta}{1 - \delta} \qquad\qquad w^m = R - \frac{\theta}{1 + \delta}. \tag{16}$$

These equations still depend on the endogenous value of cv (or θ). For θ to be consistent with equilibrium, it is necessary that unemployed workers find this minimum wage acceptable given the mean wage implied by $F(w)$. In other words, cv (or θ) must adjust so that (7) is

19

satisfied as well. To calculate this equilibrium θ, one must first compute $\bar{w}$, the mean wage earned by workers. For unemployed individuals who find two viable jobs in a given period, the probability that they both offer wages below w is $F(w)^2$. Given that all firms must offer more than the reservation wage in equilibrium, the cdf of the best offer received by workers in a period, $G(w)$, is given by

$$G(w) = (1 - \delta)F(w) + \delta F(w)^2.$$

Using (15), this becomes

$$G(w) = \frac{\theta^2}{4\delta(R - w)^2} - \frac{(1 - \delta)^2}{4\delta} = \frac{(1 - \delta)^2}{4\delta}\left(\left(\frac{R - w^*}{R - w}\right)^2 - 1\right), \tag{17}$$

where the second equality uses the first equation in (16). The average wage received by workers is thus

$$\bar{w} = \int_{w^*}^{w^m} w \, dG(w) = \int_{R - \frac{\theta}{1-\delta}}^{R - \frac{\theta}{1+\delta}} \frac{\theta^2 w \, dw}{2\delta(R - w)^3} = R - \theta, \tag{18}$$

where the substitution of w by $R - w$ in the integral simplifies the computation that leads to the last equality. Combining (18) and the first equation in (16) yields

$$\bar{w} = \delta R + (1 - \delta)w^*,$$

which makes it clear that the effect of competition in the form of a higher δ is to raise offers towards R. This linear equation in $\bar{w}$ and w^* can be combined with (16), which is also linear in these two variables. This yields the equilibrium minimum wage

$$w^* = \frac{\delta(D + \rho\lambda)R}{\delta(\gamma D + \rho\lambda) + D(1 - \gamma)}. \tag{19}$$

Differentiation of this equation leads to the conclusion that the minimum wage is increasing in the level of competition δ and in the arrival rate λ. The latter results from the effect we saw earlier, namely that a higher arrival rate of jobs makes unemployed individuals less willing to accept a wage that is below the average wage. For equilibrium wages to rise with

λ, of course, it is necessary that the cost of posting jobs decline in λ. This is indeed what happens, as is apparent once one uses the first equation of (16) to obtain θ:

$$\theta = (1-\delta)(R - w^*) = \frac{(1-\delta)^2(1-\gamma)DR}{\delta(D\gamma + \rho\lambda) + D(1-\gamma)}.$$

Since θ is proportional to cv, this shows that an increase in λ lowers cv. Note that the model is silent as to how changes in cv are distributed between changes in c and v. It is thus consistent with many technologies for posting jobs, including the one in which the cost per posting c is constant and the one in which the total number of postings v is fixed while their cost c is determined in a competitive market. The nature of this technology does matter in Section 4.

2.2 Prominent Job Advertisements with Accurate Beliefs

Unemployed workers continue to have accurate beliefs in this subsection, but a nonstationarity is introduced in that some job advertisements are more prominent than others. Only prominent ads are seen in the first period. As unemployed workers gain experience they get better at tuning out prominent ads and have a better chance of seeing ordinary ones. This is captured by supposing that, in each period starting in period 2 on, individuals have a probability η of seeing only ordinary ads and a probability $(1-\eta)$ of seeing only prominent ones. For people who have already been unemployed for one period, one can thus imagine that there is an indicator variable κ in each period that equals 1 with probability $(1-\eta)$ and 2 with probability η. In the first period of unemployment, κ is sure to equal 1. Using a parallel notation, the number of prominent postings and their cost is denoted by v_1 and c_1 respectively while the corresponding magnitudes for ordinary ads are denoted by v_2 and c_2.

Prominent ads are assumed to be distracting in another sense as well, namely that individuals who pay attention to them (i.e. who are not able to filter these out) are less likely to have access to two viable offers. This assumption is crucial, and leads the distribution of wages offered by prominent jobs to be dominated by the distribution of wages offered by other jobs. The assumption itself can be interpreted in two related ways. The first is that

people who do not tune out prominent ads are less efficient at processing ads in general so that they are less likely to find appropriate jobs. The second is that the prominent ads themselves, precisely because they are prominent, tend to snare people who are not well-matched for the job into wasting time investigating them.

In either case, one can capture this aspect of prominent ads by supposing that $\delta_2 > \delta_1$, the probability that an individual finds two viable jobs in a period is $\lambda \delta_\kappa$, the probability she finds only one is $\lambda(1 - \delta_\kappa)$, and κ is the individual's indicator variable in that period. While this leads the temporal structure of wages to resemble somewhat that of Section 1.2, this subsection demonstrates that the existence of prominent job offers does not increase wage dispersion when beliefs are accurate. Indeed, the ratio of mean to minimum wages is actually lowered by differences between δ_1 and δ_2, where the case where they are the same is equivalent to the one studied in Section 2.1.

Before carrying out any detailed computations regarding wages, it is worth giving a simple argument demonstrating that accurate beliefs lead workers to have a constant reservation wage w^{*a} even though some ads are prominent. The reason is that the future stochastic process for offers is invariant from the first decision point (at the end of period 1) onward, and thus the same is true for wage expectations in the case where these are accurate. Since no firm wants to make an offer below all workers' reservation wages, the constancy of the reservation wage implies that any unemployed individual who identifies a viable offer becomes employed. Thus, λ is once again the hazard rate of leaving unemployment. Total employment M thus equals $\lambda N/(\lambda + \sigma)$ and the number of people that are unemployed for the first time in any given period, U^1, equals $\sigma \lambda N(\lambda + \sigma)$.

These magnitudes allow one to compute the total number of viable contacts with unemployed individuals made by all postings of type i, which are denoted by Q_i. An individual whose indicator variable equals κ can expect $\lambda(1 + \delta_\kappa)$ contacts with viable offers because

he expects $\lambda(1 - \delta_\kappa)$ contacts with a single offer and $\lambda\delta_\kappa$ contacts with two. Therefore,

$$Q_1 = (1 + \delta_1)\lambda\Big(U^1 + (1 - \eta)(U - U^1)\Big) = \frac{\lambda\sigma(1 + \delta_1)(1 - \eta(1 - \lambda))N}{\lambda + \sigma} \tag{20}$$

$$Q_2 = (1 + \delta_2)\lambda\eta(U - U^1) = \frac{\lambda\sigma(1 + \delta_2)\eta(1 - \lambda)N}{\lambda + \sigma}. \tag{21}$$

A firm that posts an offer of type i with a wage of w greater than or equal to workers' reservation wage thus earns expected profits equal to

$$\frac{Q_i}{v_i}\frac{R - w}{D(1 + \delta_i)}\Big(1 - \delta_i + 2\delta_i F_i(w)\Big) - c_i, \tag{22}$$

where $F_i(w)$ is the cdf for the wages paid by offers of type i. It follows that the lowest wage offered by each type of ad equals the workers' common reservation wage. If the lowest wage offered by ads of type i were higher, a firm of this type could raise its profits by having a slightly lower wage. It would hire workers just as often and make additional profits when it did.

In an equilibrium with zero profits, (22) is zero for all wages that are actually offered so that we have

$$F_i(w) = \frac{\theta_i}{2\delta_i(R - w)} - \frac{1 - \delta_i}{2\delta_i} \qquad \text{where} \qquad \theta_i \equiv \frac{c_i v_i D(1 + \delta_i)}{Q_i}. \tag{23}$$

Since the lowest wage offered by both types of ads is w^{*a}, it must be the case that $w^{*a} = F_1^{-1}(0) = F_2^{-1}(0)$ so that

$$\frac{\theta_1}{1 - \delta_1} = \frac{\theta_2}{1 - \delta_2} = R - w^{*a}. \tag{24}$$

Using this in (23),

$$F_i(w) = \frac{1 - \delta_i}{2\delta_i}\left[\frac{R - w^{*a}}{R - w} - 1\right]. \tag{25}$$

Since $\delta_1 < \delta_2$, $F_1(w) \leq F_2(w)$ with equality when $w = w^{*a}$ and strict inequality otherwise. Thus the distribution of wages offered in ordinary ads dominates the distribution of wages in prominent ones. This is a direct effect of the lower level of competition in the latter ones. This result helps rationalize the idea that unemployed workers are better off learning to ignore prominent ads.

The cdf of the wages earned by those workers who find viable jobs when their indicator variable is κ is

$$G_\kappa(w) = (1 - \delta_\kappa)F_\kappa(w) + \delta_\kappa F_\kappa(w)^2 = \frac{\theta_\kappa^2}{4\delta_\kappa(R-w)^2} - \frac{(1-\delta_\kappa)^2}{4\delta_\kappa}, \tag{26}$$

where the second equality is obtained using the logic that leads to (17). The average wage earned by such workers is thus

$$\bar{w}_\kappa = \int_{w^{*a}} w \, dG_\kappa(w) = R - \theta_\kappa, \tag{27}$$

where the second equality follows from the argument that leads to (18).

With accurate beliefs, the average wage that unemployed workers can expect to earn from period 2 onward is

$$\tilde{w} = (1 - \eta)\bar{w}_1 + \eta\bar{w}_2. \tag{28}$$

Because all workers who take jobs in the first period draw their wage from the cdf $G_1(w)$, the overall average wage earned by workers $\bar{w}$ is

$$\bar{w} = \lambda\bar{w}_1 + (1 - \lambda)\tilde{w} = \bar{w}_1 + \eta(1 - \lambda)(\bar{w}_2 - \bar{w}_1) \tag{29}$$

This equation shows that the fraction of workers who obtain their jobs through a prominent ad is $1 - (1 - \lambda)\eta$ while the rest obtain them through ordinary ads.

The value to a worker of accepting a job that pays w remains $\tilde{V}(w)$, which is given by (1). With accurate beliefs, the value $\tilde{U}$ under the strategy of using w^{*a} as the reservation wage is

$$\tilde{U} = \gamma\bar{w} + \rho(\lambda\tilde{V}(\tilde{w}) + (1 - \lambda)\tilde{U}) = \frac{D\gamma\bar{w} + \rho\lambda\tilde{w}}{(1 - \rho)(D + \rho\lambda)}.$$

Since w^{*a} must ensure that $\tilde{V}(w^{*a})$ equals $\tilde{U}$, it must equal $(1 - \rho)\tilde{U}$ so that

$$w^{*a} = \frac{D\gamma\bar{w} + \rho\lambda\tilde{w}}{D + \rho\lambda}. \tag{30}$$

This implies

Proposition 4. *With accurate beliefs regarding the distribution of offers,*

$$w^{*a} = \frac{D\gamma[\delta_1 + \eta(1-\lambda)(\delta_2 - \delta_1)] + \rho\lambda[\delta_1 + \eta(\delta_2 - \delta_1)]}{D(1-\gamma) + D\gamma[\delta_1 + \eta(1-\lambda)(\delta_2 - \delta_1)] + \rho\lambda[\delta_1 + \eta(\delta_2 - \delta_1)]} R. \tag{31}$$

Moreover

$$\frac{\bar{w}}{w^*} \leq \frac{D + \rho\lambda}{D\gamma + \rho\lambda}, \tag{32}$$

with equality if $\eta = 0$ or $\delta_1 = \delta_2$ and strict inequality otherwise.

The Proposition shows that inequality as measured by the mean/min ratio is lower when there are prominent ads. This follows almost directly from the fact that the wages of ordinary ads dominate those from prominent ads. As a result, people who take jobs in the first period of unemployment earn relatively little, and this brings the average wage closer to w^{*a}.

As HKV have shown, one benefit of focusing on the ratio of average to minimum wages is that, under a broad set of conditions, this ratio depends only on worker behavior. Nonetheless, empirical attempts to measure this may be more sensitive to measurement error than measures such as the ratio of the mean to the 10th percentile. This ratio can be computed from the overall distribution of wages, which of course depends on the equilibrium behavior of firms as well. Since a fraction $1 - (1-\lambda)\eta$ of workers obtain their jobs from prominent ads, this overall distribution is

$$G(w) = [1 - (1-\lambda)\eta]G_1(w) + (1-\lambda)\eta G_2(w) \tag{33}$$

2.3 Naive Group Learning by a Single Group

Notice first that the rational expectations equilibrium computed in Section 2.2 is not consistent with naive group learning. The reason is that the average wage earned by all workers is below the average wage that workers who have a reservation wage of w^{*a} can expect to earn by turning down their wage offer in period 1. Unemployed individuals whose expectation is that a reservation wage of w^{*a} will lead to an average wage of $\bar{w}$ will accept jobs that pay less than w^{*a}. Indeed, it would seem fairly difficult to learn the true mean wage $\tilde{w}$ that one can earn in the future by using the reservation wage w^{*a}. Knowledge of the average wage

earned by people who accept their jobs at different times conditional on their reservation wage, seems beyond even what is known in the scholarly literature. Attempts to collect such data using interviews is likely to be subject to considerable measurement error. On the other hand, it is not at all clear how a good estimate of $\tilde{w}$ can be computed without such data if, as in the equilibrium of Section 2.2, everyone is just as likely to accept jobs in the first as in subsequent periods.

This subsection studies the opposite extreme, namely a situation where everyone is a naive group learner. As in Section 1, all people act as they would in a stationary world where the average length of unemployment spells equals the inverse of the hazard rate of finding a job while the economy-wide average wage always equals the average wage that one can expect to earn by using the common reservation wage. Since all the unemployed individuals regard the economy as stationary, their reservation wage is constant once again. Moreover, the logic of (5) ensures that wage inequality within this group remains small. Instead, the main conclusion from the subsection is that w^{*b}, the common reservation wage when everyone is a naive group learner, is smaller than w^{*a}. For certain parameters, the difference between the two can be substantial.

Except for one key difference, almost all the steps used to analyze the case of accurate beliefs in Section 2.2 can be applied to this case as well. For example, the lowest wage offered by both types of ads must equal the reservation wage in equilibrium. The reason is, again, that offering a lower wage leads to the waste of c_i whereas having a lowest wage that exceeds w^{*b} implies that firms can raise their profits by undercutting this lowest wage. These two facts imply that the hazard of leaving unemployment remains λ for all workers so that the total number of viable contacts made by offers of type i equal the Q_i values displayed in (20) and (21).

Other parallels between the cases can be seen in the demonstration of Proposition 5, which characterizes the equilibrium in this case

Proposition 5. *If all unemployed workers infer that the use of the reservation wage w^{*b} yields a job with probability λ in each period whose average wage is the economy-wide average*

26

wage of $\bar{w}$, the reservation wage satisfies

$$w^{*b} = \frac{(D\gamma + \rho\lambda)\bar{w}}{D + \rho\lambda},$$ (34)

so that the unique equilibrium reservation wage is

$$w^{*b} = \frac{(D\gamma + \rho\lambda)[\delta_1 + \eta(1-\lambda)(\delta_2 - \delta_1)]}{D(1-\gamma) + (D\gamma + \rho\lambda)[\delta_1 + \eta(1-\lambda)(\delta_2 - \delta_1)]}R.$$ (35)

The economy-wide average wage is given by (29) while

$$\bar{w}_i = \delta_i R + (1 - \delta_i)w^{*b}.$$ (36)

Lastly,

$$\frac{c_i v_i D(1 + \delta_i)}{Q_i} = R - \bar{w}_i.$$ (37)

Comparison of w^{*b} in (35) and w^{*a} in (31) shows that w^{*a} is obtained from the formula for w^{*b} after multiplying the coefficient of $\rho\lambda(\delta_2 - \delta_1)$ in the numerator and denominator by $1/(1-\lambda)$, which exceeds one. This brings w^{*a} closer to R so that it is larger than w^{*b}. This is to be expected because the unemployed workers who take jobs with probability λ in the first period depress the average wage. With naive group learning, this lowers the wage expected from continued search and thus lowers reservation wages. Note that this effect is more pronounced the larger is the difference between δ_2 and δ_1. The reason for this is that large differences between these parameters imply that firms with prominent ads have a larger incentive to set low wages.

With large values of λ and $(\delta_2 - \delta_1)$, these effects can be dramatic. Suppose that, as in HKV, σ, ρ, and γ are set to .02, .9959, and .4 respectively, while R is normalized to equal 1. If λ, δ_1, δ_2 and η are set equal to .99, .01, .8, and .6 respectively, w^{*a} equals .96, while w^{*b} equals .48.

Equation (34) ensures that the mean wage is almost the same as the minimum wage in the case of economy-wide group learning. Indeed, for the parameters above, the ratio of the average to the minimum wage is 1.015. This is even smaller than the value given in HKV because λ is assumed to be larger so that workers have even less cause to accept a wage

27

below the mean wage. With accurate beliefs, these parameters yield a ratio of the mean to the minimum wage of 1.0006, which is even smaller.

Because the minimum wage is hard to measure accurately, HKV also report measures of the mean to the 10th percentile. To compute this here, one has to start from the fact that the distribution of wages offered by firms continues to be given by (25) as long as one replaces w^{*a} with w^{*b}. With the resulting values of $F_i(w)$, the formula in (26) gives the distribution of wages received by workers who accept an offer from a firm of type i. Finally, the formula in (33) gives the overall distribution of wages. Therefore the wage such that a fraction x of workers earn less than this wage is given by

$$x = \left([1 - (1 - \lambda\eta)] \frac{(1 - \delta_1)^2}{4\delta_1} + (1 - \lambda)\eta \frac{(1 - \delta_2)^2}{4\delta_2} \right) \left(\left(\frac{R - w^{*b}}{R - w} \right)^2 - 1 \right).$$

Using this formula, the ratio of the mean to the 10th percentile wage is 1.012, so that dispersion is very small once again. The above formula for the wage such that a fraction x earns less is also valid when everyone has accurate beliefs as long as w^{*b} is replaced with w^{*a}. The ratio of the mean to the 10th percentile wage is then even smaller and equals 1.0005.

While the model with economy-wide social learning does not generate any additional wage inequality beyond that in HKV, the difference between w^{*a} and w^{*b} suggests it should be possible to induce inequality between groups if two groups learn differently. The next subsection shows that such outcomes can indeed be stable.

2.4 Two Groups with Stable Heterogeneous Beliefs

In this subsection I consider the stability of outcomes in which a group of size N^L has a reservation wage of w^{*L} and the rest of the population $(N - N^L)$ has a reservation wage of $w^{*H} > w^{*L}$. As in Section 1.2, I treat the members of the former group as being of type L while those of the latter are of type H. For the two groups to believe that they are optimizing when they accept jobs, it must be the case that people of type L incorrectly expect future wage offers to be lower than they are. This, in turn, is possible only if the average wage of people of type L is low, which requires that many of them accept low wages in period 1.

Letting U^{i1} denote the number of people of type i that become unemployed in a particular period, this requires that the fraction of U^{L1} that find jobs in the first period exceeds the corresponding fraction of U^{H1}. Letting U^{it} be the number of individuals of type i that have been unemployed for t periods, U^{Ht}/U^{H1} must therefore exceed U^{Lt}/U^{L1} for all t greater than one. Since the total number of unemployed individuals of type i, U^i equals the sum over t of U^{it}, we must have

$$\frac{U^L - U^{L1}}{U^{L1}} < \frac{U^H - U^{H1}}{U^{H1}}. \tag{38}$$

This inequality turns out to help sustain equilibria in which prominent ads offer wages below w^{*L} while ordinary ads offer wages above. The reason is that ordinary ads end up being seen mostly by people of type H that have high reservation wages, so that it is not as profitable to offer low wages with such ads. By the same token, prominent ads are seen disproportionately by people of type L, and the benefit of posting high wages in such ads is correspondingly lower.

The section proceeds as follows. First, formulas are presented for the number of contacts Q that offers of type i make with workers who find the wage w acceptable. Then, Proposition 6 gives conditions under which a particular type of ad only offers wages above w^{*H} and shows that this is more likely for ordinary ads. After this, Proposition 7 gives conditions under which a particular type of ad only has wages below w^{*H}, which turns out to be more likely for prominent ads. Proposition 8 is the analytic heart of the subsection because it provides conditions on the parameters such that the two types of ads have non-overlapping wage ranges with only people of type L accepting jobs from prominent ads.

Using the analysis that leads to (20), the total contacts of prominent ads with unemployed workers are given by

$$Q_1^H = \lambda(1 + \delta_1)\Big(U^{L1} + (1 - \eta)(U^L - U^{L1}) + [U^{H1} + (1 - \eta)(U^H - U^{H1})]\Big). \tag{39}$$

Similarly, the total contacts between ordinary ads and unemployed workers are

$$Q_2^H = \lambda\eta(1 + \delta_2)\Big(U^L - U^{L1} + [U^H - U^{H1}]\Big). \tag{40}$$

An individual offer of type i, of course, has only $1/v_i$ as many contacts as the total contacts of all offers of type i. All these contacts find the job acceptable if its wage is greater than or equal to w^{*H}, and this is captured by the H superscript in the definitions of Q above. If a posting pays a w less than w^{*L}, it is rejected by everyone while, if $w^{*L} \leq w < w^{*H}$, the unemployed workers that appear in square brackets in the expressions above do not find the job acceptable while the rest do. It is thus helpful to define Q_i^L by the expressions for Q_i^H when the terms in square brackets are set to zero. Then, Q_i^L/v_i denotes the number of unemployed individuals that find an offer of type i acceptable if its wage satisfies $w^{*L} \leq w < w^{*H}$. Using these definitions, one can show that,

Proposition 6. *If*

$$\frac{Q_i^L(R - w^{*L})}{Q_i^H(R - w^{*H})} < 1 \tag{41}$$

no job postings of type i offers a wage below w^{*H}. In this case, free entry implies that the distribution of wages paid by postings of type i satisfies (23) with Q_i replaced by Q_i^H. As a result, θ_i equals $(1 - \delta_i)(R - w^{*H})$. If the inequality (41) is reversed, a positive fraction of postings of type i offer less than w^{*H}.

If some ordinary offers pay a wage below w^{*H}, then some prominent ones do as well. The converse is not true.

The proposition shows that, if wages below w^{*H} are offered at all, they must also be offered by prominent ads. At the same time,

Proposition 7. *If*

$$\frac{Q_i^L(R - w^{*L})(1 - \delta_i)}{Q_i^H(R - w^{*H})(1 + \delta_i)} > 1 \tag{42}$$

no wage greater than or equal to w^{*H} is offered. In this case, the distribution across postings of wages offered is given by (23) with Q_i replaced by Q_i^L. The variable θ_i is then equal to $(1 - \delta_i)(R - w^{*L})$.

Suppose (41) and (42) are reversed. Then, for any values of c_i and v_i, the condition that firms make the same profits at any wage they post implies that no wage between $w^{mL} < w^{*H}$

and w^{*H} is posted though both wages above w^{*H} and wages between w^{*L} and w^{mL} are. The wage w^{mL} is defined by

$$\frac{Q_i^L(R - w^{mL})}{Q_i^H(R - w^{*H})} = 1. \tag{43}$$

For sufficiently low δ_1, (42) holds for prominent firms whenever (41) is violated for such firms. The Proposition then implies that their wages are all below w^{*H}. I thus focus on the existence of free entry equilibria in which people of type H work only for ordinary firms while prominent firms hire only people of type L. This requires that (41) hold for $i = 2$ while (42) holds so that (41) is reversed for $i = 1$. The next proposition gives conditions for such equilibria.

Proposition 8. *If*

$$\frac{w^{*H}}{R} = \frac{\delta_2(D\gamma + \rho\lambda\eta)}{D(1 - \gamma) + \delta_2(D\gamma + \rho\lambda\eta)}, \tag{44}$$

and

$$\frac{w^{*L}}{R} = \frac{(D\gamma + \rho\lambda)\left[(1 - (1 - \lambda)\eta)\delta_1 + \frac{(1-\lambda)\eta\delta_2(D+\rho\lambda\eta)}{D(1-\gamma)+\delta_2(D\gamma+\rho\lambda\eta)}\right]}{D(1 - \gamma) + (D\gamma + \rho\lambda)((1 - (1 - \lambda)\eta)\delta_1 + (1 - \lambda)\eta)}, \tag{45}$$

while, for these values of w^{*L} and w^{*H},

$$\left(1 + \frac{(\lambda + \sigma)(1 - N^L)}{(\sigma + \lambda\eta(1 + \sigma))(1 - \lambda)N^L}\right)\frac{R - w^{*H}}{R - w^{*L}} > 1, \tag{46}$$

$$\left(1 + \frac{(\lambda + \sigma)(1 - N^L)}{(\sigma + \lambda\eta(1 + \sigma))N^L}\right)\frac{R - w^{*H}}{R - w^{*L}} < 1, \tag{47}$$

there exists a free entry equilibrium in which firms of type 1 only offer wages strictly below w^{*H} while firms of type 2 only offers wages above. At the same time, it is stable for the N^L workers who are naive group learners to set w^{*L} as their reservation wage. Whether the workers of group H have accurate beliefs or are sophisticated group learners, it is stable for them to have a reservation wage of w^{*H}.

This proposition gives necessary conditions for an equilibrium where it to be stable for the two types to have different reservation wages while firms behave optimally. It does not say whether parameters exist such that (46) and (47) hold when the reservation wages are given by (45) and (44) respectively. However, inspection of these inequalities does suggest

that they are likely to hold if both λ and δ_2 are high while δ_1 is low. The reason is that a high λ implies that (46) holds regardless of the values of w^{*i}. A high value of λ also ensures that w^{*L} in (45) is low when δ_1 is low. Lastly, because $D(1-\gamma)$ is small relative to $(D\gamma + \rho\lambda\eta)$, a high value of δ_2 leads w^{*H} to be fairly high as well. The result is that $(R - w^{*H})/(R - w^{*L})$ is low so that (47) holds as well.

For N^L equal .4, the parameters used as illustration in Section 2.3 do indeed satisfy all the conditions of Proposition 8. Recall that, as in HKV, σ, ρ, and γ were equal to .02, .9959, and .4 respectively, while R was normalized to equal 1. The parameters λ, δ_1, δ_2 and η were set equal to .99, .01, .8, and .6. Here, this leads w^{*L} to equal .52 while w^{*H} equals .97. The former is somewhat larger than the common reservation wage when everyone is a naive group learner while the second is slightly larger than the common reservation wage when everyone has accurate beliefs. The fact that w^{*L} exceeds the wage when everyone is a naive group learner is easy to understand. It comes about because more sophisticated workers are willing to wait until competition for workers becomes more intense and thus receive higher wages. Because some naive group learners also receive these higher wages, their average wages are increased, and this raises w^{*L}.

The differences between the wage earned by more sophisticated workers in this subsection and the wage earned when everyone has accurate beliefs are more subtle. The former is given by (44) while the latter is given by (31). Comparing the two formulas, it is seen that coefficient of $\rho\lambda$ is larger for w^{*a}. This effect, which tends to raise w^{*a} relative to w^{*H} comes about because workers with accurate beliefs receive acceptable offers more frequently if all workers have accurate beliefs. Otherwise, they often encounter offers that are meant for more naive workers. The second difference between the formulas is that the coefficient of $D\gamma$ is larger in (44), which tends to raise w^{*H} relative to w^{*a}. The difference here is that the high value of λ implies that, when all workers have accurate beliefs, many of them accept jobs in the first period. This depresses their average wage relative to the case where they all wait and lowers the minimum wage they require. For the parameters in this example, this effects slightly dominates.

The overall average wage in this economy $\bar{w}$ is given by

$$\bar{w} = \Big([1 - (1 - \lambda)\eta]\bar{w}_1 + (1 - \lambda)\eta\bar{w}_2\Big)N^L + \bar{w}_2(1 - N^L). \tag{48}$$

For the parameters described above, this equals 1.55 times the minimum wage w^{*L}.

Since the distribution of wages offered by postings of type i is given by (23), the distribution of wages received by workers who obtain them from postings of type i is given by (26). Because the fraction of workers that obtains jobs from offers of type 1 equals $N^L[1 - (1 - \lambda)\eta]$, and all these workers earn less than w^{*H}, the wage w such that a fraction x of workers earns less than this wage is given by

$$x = N^L[1 - (1 - \lambda)\eta]\frac{(1 - \delta_1)^2}{4\delta_1}\left(\left(\frac{R - w^{*L}}{R - w}\right)^2 - 1\right), \tag{49}$$

when x is below $N^L[1 - (1 - \lambda)\eta]$. For the parameters above, the 10th percentile is below $N^L[1 - (1 - \lambda)\eta]$ so this formula can be used to compute the ratio of the mean wage to the 10th percentile. For the parameters above this ratio equals 1.54, which is not far from the empirical estimates of HKV.

3 Tests of Stationarity and Accurate Beliefs

The main source of wage dispersion according to the theory presented here is that some people accept jobs too soon because they learn from their peers that this is a good idea. If these people raised their reservation wage, they would be surprised at how quickly they would receive offers that met this higher reservation wage. A simple test of this idea involves comparing the wage gains from having a slightly higher reservation wage to corresponding increases in the time it takes to find a job. If the former are high relative to the latter, the standard assumptions that wage offers originate from a stationary distribution and that workers have accurate beliefs about this distribution can be rejected.

Questionnaire evidence makes it clear that individuals differ in the reservation wage. Leaving aside differences in their opportunity sets, this can be for the result of differences in tastes, in the level of unemployment insurance coverage, or, as stressed here, in beliefs.

Suppose then that individuals have a constant probability λ of receiving offers that are drawn from the constant distribution $G(w)$. Whether time is continuous, as is standard in this literature, or discrete, the expected duration of unemployment S for an individual with a reservation wage w^* is

$$S = \frac{1}{\lambda(1 - G(w^*))}.$$

Meanwhile, the average wage received upon employment by such an individual is given by $\bar{w}$ in (6). Differentiating these equations, the effect of changes in w^* on the percentage changes in S and $\bar{w}$ is

$$\frac{dS}{S} = \frac{g(w^*)}{1 - G(w^*)}dw^* \qquad \frac{d\bar{w}}{\bar{w}} = \frac{g(w^*)}{1 - G(w^*)}\left(1 - \frac{w^*}{\bar{w}}\right)dw^*. \tag{50}$$

Taking the ratio of the two elasticities in (50) and calling it r

$$r \equiv \frac{d\bar{w}/\bar{w}}{dS/S} = 1 - \frac{w^*}{\bar{w}} \tag{51}$$

This remarkably simple formula is the basis of the tests I propose.[9] The ratio r is akin to a "return" earned by people for spending more time looking for a job. Those that set higher reservation wages turn down more jobs so their unemployment duration is longer, and r measures the benefit in terms of obtaining higher wages from doing so.

One reason to be interested in r is that it is amenable to measurement and, indeed, one can obtain estimates of it from Holzer (1986)[10]. Holzer (1986) uses retrospective data on unemployed men aged 16 through 21 from the NLS New Youth Cohort. He uses the 1980 survey to obtain the wage individuals received upon employment as well as their duration of unemployment and relates these to these individual's stated reservation wage in 1979. As controls, he uses 1979 data on the individual's past occupation, industry and union status as well as schooling, experience, household income, region, marital status, "Knowledge of the World of Work," and the existence of a library card in the home. In some of his specifications, past individual wages are used as controls instead.

---

[9]Derivatives of the form of (50) can be found in Holzer (1986), but the ratio in (51) appears to be new to the literature.

[10]Krueger and Mueller (2011) provide indirect evidence that higher reservation wages lead to higher wages and longer unemployment durations, but their published estimates do not permit a calculation of r.

Holzer (1986) presents regressions of the log wage ultimately earned and of the log duration of unemployment on the individual's log reservation wage as well as on controls. The ratio r is thus equal to the ratio of the primary coefficients obtained in these two regressions. When using weighted least squares (based on sample weights), the resulting ratios are either 2.33 or 1.85 for whites depending on controls.[11] For blacks, they are either .32 or .18.

The stationarity of wage offers combined with the existence of a well-defined reservation wage implies that these values of r should be the same as those on the right hand side of (51), so they should be close to $(1 - w^*/\bar{w})$. HKV provide estimates of $\bar{w}/w^*$ based on the average and minimum wages for narrow occupations in particular locations, and obtain values between 2 and 3. The resulting values of $(1 - w^*/\bar{w})$ are not out of line with the r's for blacks discussed above but they are quite inconsistent with the r's for whites. Indeed, these exceed 1, and the right hand side of (51) cannot do so for any $w^*/\bar{w}$ that is estimated by the ratio of the minimum to the average of a series of wage observations.

It is worth noting that the inconsistency of the "return to waiting" r and the observed ratio of average to minimum wages is not a simple reiteration of HKV's finding that the ratio of average to minimum wages is too large for the standard search model. In particular, HKV emphasize that high values of the ratio of average to minimum wages might be consistent with search models in which workers continue to search and find jobs after they start working. A key property of these models is that employees accrue wage gains that are not associated with any unemployment, and this turns out to intensify the puzzle discussed here. The reason is that on-the-job search implies that the measured average wage exceeds the mean wage that people receive in their first job after a spell of unemployment. In equation (51), $\bar{w}$ represents the latter. For a given value of measured minimum and average wages, the theory-consistent value of r is thus lower when there is on-the-job search than when there isn't. On-the-job-search thus makes it more difficult to explain high returns from raising

---

[11]The OLS results seem less applicable than the WLS results as estimates of the typical effect in the population he considers. Nonetheless, it is worth noting that the OLS coefficients of reservation wages in the equation explaining the unemployment duration of whites are negative, which is a sign of misspecification. These duration regressions do not include the past wage as a control.

one's reservation wage.

If one is armed not only with people's reservation wages but also with their own estimates of how much they expect to earn when they get a job, one can use (51) to test whether their expectations are accurate (conditional on offers being stationary and people using their reservation wage consistently). Lancaster and Chesher (1983) have data of this sort. Their appendix presents a joint frequency distribution for their respondents expected and reservation wages. Unfortunately, the responses concerning both wages are amalgamated into 11 discrete values for wages, with the highest of these being over five times larger than the smallest. The large distance between these wage "buckets" is presumably responsible for the fact that the reported frequencies in Lancaster and Chesher (1983) includes a great many observations in which the reservation wage and the expected wage coincide. As a result, computation of $w^*/\bar{w}$ from these data is somewhat perilous. Still, it seems relevant that the average of the ratio of the expected wage to the reservation wage in their reported distribution equals 1.15. When this is interpreted as being equal to $\bar{w}/w^*$, it implies that r should equal .13, which remains smaller than any of the estimates based on Holzer (1986). While the Lancaster and Chesher (1983) and the Holzer (1986) samples are not the same, this suggests that the returns to increasing the time spent searching could be larger than what is implied by their own subjective beliefs.

One can calculate a variant of r for the model developed in Section 2.4 by taking the ratio of the log difference in average wages for the two groups divided by the log difference in their durations. To calculate this, note that the average wage earned by workers who obtained their job through a prominent offer is $\bar{w}_1 = w^{*L} + \delta_1(R - w^{*L})$ while the average wage of those that obtained their job through an ordinary ad is $\bar{w}_2 = w^{*H} + \delta_2(R - w^{*H})$. The average wage earned by workers of type H is simply $\bar{w}_2$ while that of people of type L equals $\bar{w}_1 + (1 - \lambda)\eta(\bar{w}_2 - \bar{w}_1)$. The expected unemployment duration for workers of type L is $1/\lambda$ whereas that for employees of type 2 is $(1 + 1/\lambda\eta)$.

For the parameters used in Section 2.4, the value of r based on the ratio of log differences of the average wages for the two types and the log difference of their average durations is .51,

which is actually lower than the average of the four Holzer (1986) estimates reported above. More important than the model's capacity to yield relatively high values of r is its capacity to yield values of r that exceed the right hand side of (51). As discussed below equation (48), these parameters imply that $\bar{w}/w^*$ equals 1.55. The right hand side of (51) thus equals .35, which is indeed smaller than r. This shows that instances in which measurements of the left hand side yield smaller values than measurements of the right hand side do indeed correspond to rejections of the null hypothesis that the models derived from stationarity and accurate beliefs are valid against an alternative hypotheses that is consistent with the current model.

4 Productivity and Wages

The reason more productive firms pay higher wages in the variant of the Burdett and Mortenson (1998) model discussed in Mortensen (2003 p. 21) is that these firms are particularly keen to have a high probability of recruiting workers (or to recruit a large number of them). In their model, the only way to raise the probability of recruitment is to raise wages, and the result follows. Here, however, firms also have to decide whether to purchase ordinary or prominent ads. Insofar the recruitment probability is maximized using the latter type, the most productive firms may not have the highest wages.

To see this, consider two firms with marginal products of labor equal to R' and R'' respectively where $R'' > R'$. Let Q' and Q'' denote, respectively, their Q values. Supposing that the conditions of Proposition 8 are met for all possible values of the marginal product of labor including R' and R'', Q' and Q'' must equal either Q_1^L or Q_2^H. Thus, the identity of the firms' Q also determines their v's, their c's, their F's and their δ's, where I continue to use primes and double primes to denote the values for the two firms. Keeping this in mind, and supposing that the two firm's wages are w' and w'', let A' and A'' be given by

$$A' = \frac{1 - \delta' + 2\delta' F'(w')}{Dv'(1 + \delta')} \qquad A'' = \frac{1 - \delta'' + 2\delta'' F''(w'')}{Dv''(1 + \delta'')}.$$

The firm with R' must at least weakly prefer the combination of Q' and w' to the com-

bination of Q'' and w'', with the reverse being true for the firm with R''. Therefore

$$Q''A''(R'' - w'') - c'' \geq Q'A'(R'' - w') - c' \qquad Q'A'(R' - w') - c' \geq Q''A''(R' - w'') - c''.$$

Subtracting one of these inequalities from the other, we have

$$Q''A''(R'' - R') \geq Q'A'(R'' - R').$$

Since R'' exceeds R', $Q''A''$ must be at least as large as $Q'A'$. If $Q' = Q''$ so the firms use the same types of ads, $A'' \geq A'$. Since A is rising in the wage for a given type of ad, the most productive firm must have the highest wage paid by any firm that advertises in the same way.

To see that it need not have the highest wage overall, imagine that essentially all firms have a marginal product of labor of R' and that a single infinitesimal firm has the marginal product of labor R''. The analysis of Section 2.4 then remains intact, except that the firm with $R = R''$ pays either the highest wage of firms of type 2, w_2^m or the highest wage of firms of type 1, w_1^m. In either case, $F_i(w_i^m) = 1$ so that A' and A'' at these wages equal $1/Dv'$ and $1/Dv''$ respectively. The firm with $R = R''$ thus fails to have the highest wage overall if it prefers a prominent ad, which occurs if

$$\frac{Q_2^H}{Dv_2} < \frac{Q_1^L}{Dv_1} \tag{52}$$

whereas it does have the highest wage if the inequality is reversed.

For firms with $R = R'$, offering a wage of w_1^m with a prominent advertisement and offering a wage of w_2^m with an ordinary one should yield the same profits of zero. Therefore

$$\frac{Q_2^H(R - w_2^m)}{Dc_2 v_2} = \frac{Q_1^L(R - w_1^m)}{Dc_1 v_1}.$$

If the conditions of Proposition 8 are met, w_2^m is higher than w_1^m. This implies that, if the cost of the two postings c_1 and c_2 are the same, (52) is reversed and the most productive firm offers the highest wage. For a sufficiently high c_1/c_2 however, (52) holds and the most productive firm offers only w_1^m which is less even that w^{*H}.

The purchase of the most expensive ads can be an attractive solution for the most productive firms because, in equilibrium, ads are expensive if they have a high likelihood of attracting employees. This can occur, for example, if prominent job advertisements are both in fixed supply and scarce relative to the number of susceptible unemployed individuals. They can then be a very effective recruitment device with a high shadow cost. Productive firms may then choose these ads and feel no pressure to pay more than the highest wage offered by other prominent advertisements. If, instead, prominent ads are cheap, the low wages that they advertise attract a large number of firms that try their chance at hiring workers in this way, so that more productive firms prefer to use ordinary ads and pay more.

5   Conclusions

This paper has presented a very stylized model that generates wage inequality as a result of people's reliance on peers for information about expected wages and the unemployment durations. The key idea of the model is that people whose peers have low wages and short unemployment spells come to expect that all jobs have relatively low wages so they accept low-wage jobs relatively quickly. People with peers that have higher wages are, instead, more choosy and wait for better jobs.

While the model is consistent with a great deal of inequality, its current incarnation does not yield continuous and concave distributions of wages of the sort that are empirically observed and displayed in Mortensen (2003, p. 48-51). This raises the question of whether extensions that allow for a more diverse set of groups or for a richer range of productivity differences across firms can fit the wage distribution better.

The model implies that some people (those of type L) would benefit from staying unemployed longer. This need not be be true for everyone, however. Indeed, being too reliant on the wages earned by one's peer group as indicators of one's future offers might well lead people who belong to high-wage groups to remain unemployed too long, and thereby lose marketability. An extension incorporating this idea might fit with the observation in Lollivier and Rioux (2010) that some people experience declines in the offers they receive as

they stay unemployed, suggesting that they should have accepted earlier offers.

A different extension that seems worth pursuing is to allow workers to quit endogenously, as in Mortensen (2003). This ought to lead workers of type L to quit often since both their wages and their unemployment durations are low. High wage firms would then have much lower turnover than low-wage ones, as is documented for example in Abowd, Kramarz and Roux (2006).

This paper stands in a complementary relation with the theoretical literature on job referrals by peers. That literature emphasizes the role of peers in communicating information about particular job openings, which this paper ignores, but neglects the role of peers in assessing whether a particular job is worth taking, which this paper emphasizes. While modeling both of these roles of peers simultaneously may be challenging, it seems worthwhile because several empirical papers on job referrals suggest that jobs obtained through peer contacts sometimes carry low wages.[12] This seems consistent with the idea suggested in this paper that peer groups can lead their members to accept relatively unattractive jobs by causing them to believe that better opportunities are more scarce than they actually are. Firms that pay low wages should be particularly eager to enroll their employees in the recruitment of others, and this might reinforce the phenomenon discussed here.

---

[12]For examples, see Loury (2006), which contains several references from the sociology literature, and Bayer, Ross and Topa (2008).

References

Abowd, John M., Francis Kramarz, and Sebastien Roux, "Wages, Mobility and Firm Performance: Advantages and Insights from Using Matched Worker-Firm Data," *Economic Journal,* 116, June 2006, 245–85.

Ahmed, Ali M., "If You Believe that Discrimination Exists, it Will," *The Manchester School,* 76, December 2008, 613–28.

Armstrong, Mark, John Vickers and Jidong Zhou, "Prominence and Consumer Search," *RAND Journal of Economics,* 40, July 2009, 209–33.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa, "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy,* 116, December 2008, 1150–96.

Burdett, Kenneth and Kenneth L. Judd, "Equilibrium Price Dispersion," *Econometrica,* 51, July 1983, 955–69.

Burdett, Kenneth and Dale T. Mortensen, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review,* 39, May 1998, 257–73.

Calvó-Armengol, Antoni and Matthew O. Jackson, "The Effects of Social Networks on Employment and Inequality," *American Economic Review,* 94, June 2004, 426–54.

Fudenberg, Drew and David K. Levine, "Self-Confirming Equilibrium," *Econometrica,* 61, May 1993, 523–45.

Haan, Marco A. and Jose L. Moraga-Gonzalez, "Advertising for Attention in a Consumer Search Model," *Economic Journal,* 121, May 2011, 552–579.

Holzer, Harry J. "Reservation Wages and their Labor Market Effects for Black and White Male Youth," *Journal of Human Resources,* 21, April 1986, 157–77.

Hornstein, Andreas, Per Krusell, and Giovanni L. Violante, "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *American Economic Review,* 101, December 2011, 2873–98.

Katz, Lawrence F. and Lawrence H. Summers, "Industry Rents: Evidence and Implications," *Brookings Papers on Economic Activity,* 1989, 209–75.

Krueger, Alan B. and Andreas Mueller, "Job Search, Emotional Well-Being, and Job Finding in a Period of Mass Unemployment: Evidence from High-Frequency Longitudinal Data," *Brookings Papers on Economic Activity*, April 2011, 1–57.

Lancaster, Tony and Andrew Chesher, "An Econometric Analysis of Reservation Wages," *Econometrica,* 51, November 1983, 1661–76.

Lang, Kevin, "Persistent Wage Dispersion and Involuntary Unemployment," *Quarterly Journal of Economics,* 106, February 1991, 181–202.

Lollivier, Stefan and Laurence Rioux, "An Empirical Examination of the Sources of Changes over Time in the Job Finding Rate Using Reservation Wages and Rejected Wage Offers," *International Economic Review,* 51, November 2010, 1039–69.

Loury, Linda Datcher, "Some Contacts Are More Equal than Others: Informal Networks, Job Tenure, and Wages,"*Journal of Labor Economics,* 24, April 2006, 299–318.

Machin, Stephen and Alan Manning, "The Causes and Consequences of Longterm Unemployment in Europe," in Orley Ashenfelter and David Card (eds.), *Handbook of Labor Economics*, 3, New York and Oxford: Elsevier Science, North-Holland, 1999, 3085–3139.

McCall, John J., "Economics of Information and Job Search," *Quarterly Journal of Economics,* 84, February 1970, 113–26.

Montgomery, James D., "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis," *American Economic Review,* 81, December 1991, 1408–18.

Mortensen, Dale T., *Wage Dispersion: Why Are Similar Workers Paid Differently?,* Cambridge, MA: MIT Press, 2003.

Nishimura, Kiyohiko G. and Hiroyuki Ozaki, "Search and Knightian Uncertainty," *Journal of Economic Theory,* 119, December 2004, 299–333.

Rotemberg, Julio J., "Perceptions of Equity and the Distribution of Income," *Journal of Labor Economics,* 20, April 2002, 249–88.

Rothschild, Michael, "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory,* 9, October 1974, 185–202.

Appendix: Proofs of Propositions

Proof of Proposition 1: Let $V(w)$ represent the present value of revenues for workers who are currently employed at a firm that pays w. The argument in the text still applies so that $[1 - \rho(1-\sigma)]V(w)$ equals $w + \rho\sigma U$ for workers who know λ_u and G, where U represents the value of being unemployed. For the reservation wage w^*, $V(w^*)$ must equal U, so they both equal $w^*/(1-\rho)$. As a result, $[1 - \rho(1-\sigma)](V(w) - U)$ equals $w - w^*$, which I use below. Given a reservation wage w^*, the value of U for a worker who knows λ_u and G is

$$U = b + \rho\Big[U + \lambda_u \int_{w^*} (V(w) - U)dG(w)\Big].$$

Therefore

$$w^* = b + \frac{\rho\lambda_u}{1 - \rho(1-\sigma)} \int_{w^*} (w - w^*)dG(w).$$

Once one carries out the substitutions spelled out in the Proposition, this is a restatement of (5). For values of w^* below b, the left hand side is below the right hand side, while the opposite is true for the largest value of w in the support of G (as long as this exceeds b). Moreover, the derivative of the left hand side with respect to w^* is positive while that of the right hand side is negative. The equation thus has a unique solution.

Proof of Proposition 2: Using (2), individuals accept w_1 rather than turning it down and setting their reservation wage to w_2 if and only if

$$w_1 \geq \frac{Db + \rho\hat{\lambda}w_2}{D + \rho\hat{\lambda}},$$

where the right hand side is based on the present value of the utility of being unemployed under the strategy of using a reservation wage of w_2. Thus, w_1 is the reservation wage for individuals with accurate beliefs if this inequality holds when $\hat{\lambda}$ is replaced with $\lambda\eta$. In this case, sophisticated social learners also accept offers of w_1. Reductions in $\hat{\lambda}$ lower the right hand side of the above inequality. This implies that naive social learners, whose $\hat{\lambda}$ is lower, accept offers of w_1 as well.

Proof of Proposition 3: The second inequality in (9) implies that (8) is violated so that individuals with accurate beliefs have a reservation wage of w_2. If a group is sophisticated, there is also a stable group learning outcome where the reservation wage is w_2. This follows from computing $\tilde{U}$ under a reservation wage of w_2 coupled with these beliefs and noting that the violation of (8) leads individuals with $\hat{\lambda}$ equal to $\lambda\eta$ to turn down offers of w_1. I now demonstrate that the first inequality implies there is a stable group learning outcome with a reservation wage of w_1. If a group uses this reservation wage, its average unemployment spell is $1/\lambda$ so that $\hat{\lambda}$ is equal to λ itself. Its average wage $\bar{w}_1$ is given by

$$\bar{w}_1 = \lambda w_1 + (1 - \lambda)(\eta w_2 + (1 - \eta)w_1),$$

where the first term captures that a fraction λ of newly unemployed workers find a job in the first period. This job pays w_1. The second term captures that the rest get an offer of w_2 with probability η and an offer of w_1 with the remaining probability. Using (3), the expected value of being unemployed for this group is

$$\tilde{U}_1 = \frac{Db + \rho\lambda[(1 - \eta(1 - \lambda))w_1 + \eta(1 - \lambda)w_2]}{(1 - \rho)(D + \rho\lambda)}.$$

Inequality (2) then implies that this group accepts a wage of w_1 if the first inequality in (9) holds. The last step is to show that values of the parameters can be found that satisfy both inequalities. Notice that both the left and the right hand side of (9) are convex combinations of b and w_2, with the weight on b being larger on the left hand side if $\lambda < 1$. Thus, as long as $w_2 > b$, there is a range of values that satisfies both inequalities.

Proof of Proposition 4: Equations (24) and (27) imply that

$$\bar{w}_i = \delta_i R + (1 - \delta_i) w^{*a}.$$

Using this in (28) and (29), and plugging the results in (30), (31) follows. Moreover, for $\eta = 0$ or $\delta_1 = \delta_2$, both $\bar{w}$ and $\tilde{w}$ equal $\bar{w}_1$ so that (32) holds as an equality. Otherwise $\bar{w}$ is smaller than $\tilde{w}$ so that it holds as a strict inequality.

Proof of Proposition 5: Since the Q_i are the same, the expected profits of a firm using an ad of type i and offering a wage w are given by (22) once again. Therefore (23) holds at a zero profit equilibrium, though the values of θ_i can be different. It follows that both (24) and (25) hold when w^{*a} is replaced by w^{*b}. Since (26) still defines the cdf of the wages earned by people who obtain their job using an ad of type i, (27) still gives the value of the average wages $\bar{w}_i$ as long as w^{*a} is replaced by w^{*b}. This implies (36).

Meanwhile, (29) continues to define the economy-wide wage $\bar{w}$ and $\tilde{V}(w)$ in (1) still gives the value to a worker of accepting a job. With group learning, the expected value of $\tilde{U}$ when using the reservation wage w^{*b} is

$$\tilde{U} = \gamma \bar{w} + \rho(\lambda \tilde{V}(\bar{w}) + (1 - \lambda)\tilde{U}) = \frac{(D\gamma + \rho\lambda)\bar{w}}{(1 - \rho)(D + \rho\lambda)}.$$

Given this expectation, the actual reservation wage of workers w^{*b} ensures that $\tilde{V}(w^{*b})$ equals $\tilde{U}$, so that it equals $(1 - \rho)\tilde{U}$. Equation (34) follows. Using (29) and (36) in this equation, the unique solution for w^{*b} is (35).

Average wages are then given by (29) and (36). Since (27) still holds and θ_i continues to be defined by (23), (37) follows.

Proof of Proposition 6: There is no mass point of offers at either w^{*H} or w^{*L} because firms would be better off offering slightly more. If a wage below w^{*H} is offered, w^{*L} must be offered as well. The reason is that, if the lowest wage were higher, higher profits would be earned by undercutting this lowest wage slightly. The profits from offering w^{*L} are

$$\frac{Q_i^L(R - w^{*L})}{(1 + \delta_i)Dv_i}(1 - \delta_i) - c_i. \tag{53}$$

Even in the case where $F(w^{*H}) = 0$, the expected profits from offering a wage of w^{*H} equal

$$\frac{Q_i^H(R - w^{*H})}{(1 + \delta_i)Dv_i}(1 - \delta_i) - c_i, \tag{54}$$

and they are higher still if $F(w^{*H}) > 0$. Thus offering a wage of w^{*H} strictly dominates offering any lower wage when (41) holds. With free entry, the lowest wage offered is w^{*H}, which implies that the distribution of wages is given by (23) with Q_i replaced by Q_i^H and that $\theta_i = (1 - \delta_i)(R - w^{*H})$.

If no one offers a wage below w^{*H}, w^{*H} is the lowest wage offered. This means that, if the inequality in (41) is reversed, expected profits are strictly higher by offering w^{*L} rather than w^{*H}.

Using the definitions of Q_i^j given in (39), (40) and the discussion below

$$\frac{Q_H^1}{Q_L^1} = 1 + \frac{U^{H1} + (1-\eta)(U^H - U^{H1})}{U^{L1} + (1-\eta)(U^L - U^{L1})}$$

$$\frac{Q_H^2}{Q_L^2} = 1 + \frac{U^H - U^{H1}}{U^L - U^{L1}}.$$

Since (38) implies that

$$\frac{U^{H1}}{U^{L1}} < \frac{U^H - U^{H1}}{U^L - U^{L1}},$$

it follows that $Q_H^1/Q_L^1 < Q_H^2/Q_L^2$. Therefore (41) holds for $i = H$ if it holds for $i = L$ but the converse need not be true.

Proof of Proposition 7: If offers above w^{*H} are made, so are offers of w^{*H}. The expected profits from making such offers are

$$\frac{Q_i^H(R - w^{*H})}{(1 + \delta_i)Dv_i}(1 - \delta_i + 2\delta_i F_i(w^{*H})) - c_i, \tag{55}$$

so they are bounded above by

$$\frac{Q_i^H(R - w^{*H})}{(1 + \delta_i)Dv_i}(1 + \delta_i) - c_i. \tag{56}$$

The inequality in (42) implies that (41) is reversed so that ads of type i include offers of w^{*L} and profits at this wage are given by the expression in (53). Therefore (42) implies that the highest possible profits from setting a wage greater than or equal to w^{*H} are below those of setting a wage of w^{*L}. Therefore, these higher wages are not offered.

Free entry then ensures that the expression in (53) equals zero so that (23) with Q_i replaced by Q_i^L gives the distribution of wages while $\theta_i = (1 - \delta_i)(R - w^{*L})$.

Now consider the case where both (41) and (42) are reversed There must then exist an $0 < F_i(w^{*H}) < 1$ such that the expression in (55) equals the expression in (53). Moreover, because (41) is reversed, there exists a value of $w^{*L} < w^{mL} < w^{*H}$ such that (43) is satisfied. Expected profits at the posted wage of w^{mL} are then the same as at w^{*H} as long as $F_i(w^{*H}) = F_i(w^{mL})$.

For fixed c_i and v_i, profits at all wages must equal those in (53). Thus, for $w^{*L} \le w \le w^{mL}$, the cdf of wages $F_i(w)$ is given by

$$(R - w)(1 - \delta_i + 2\delta_i F_i(w)) = (R - w^{*L})(1 - \delta_i),$$

while it is given by

$$Q_i^H(R - w)(1 - \delta_i + 2\delta_i F_i(w)) = Q_i^L(R - w^{*L})(1 - \delta_i),$$

for $w \ge w^{*H}$.

Proof of Proposition 8: The proof starts by supposing that, indeed, prominent job advertisements offer wages between w^{*L} and $w^{mL} < w^{*H}$ while ordinary job advertisements offer wages greater than or equal to w^{*H}. It first computes the distribution of wages at a free entry equilibrium of this sort. It then shows that members of H should set their reservation wage to w^{*H} whether they have accurate beliefs or are sophisticated group learners, while it is stable for naive group learners of type L to set it to w^{*L}. Lastly, it shows that (46) and (47) are sufficient to prevent both the deviations in which job advertisements of type 1 offer wages greater than or equal to w^{*H} and the deviations in which job advertisements of type 2 offer wages smaller than w^{*H}.

Suppose that prominent ads offer wages below w^{*H} while ordinary ads offer wages that are above. For zero profits, the cdf's $F_i(w)$ must lead (22) to be zero when Q_1 is equated with Q_1^L while Q_2 is equated with Q_2^H. These cdf's must therefore satisfy (23) with these Q's while, as in (24) the values of $c_i v_i$ ensure that

$$(R - w^{*L})(1 - \delta_1) = \theta_1 \qquad (R - w^{*H})(1 - \delta_2) = \theta_2.$$

It follows that the people who accept jobs from ads of type i have a cdf of wages $G_i(w)$ given by (26) so that their average wage $\bar{w}_i$ satisfies (27). Together with the equation above, this implies that

$$\bar{w}_1 = \delta_1 R + (1 - \delta_1)w^{*L} \qquad \bar{w}_2 = \delta_2 R + (1 - \delta_2)w^{*H} \qquad (57)$$

Now consider individuals of type H. Since wages greater than or equal to w^{*H} are only offered by ordinary ads, they arrive with probability $\lambda\eta$ starting in period 2. Thus, whether these workers have accurate beliefs or are sophisticated learners who realize that wages above w^{*H} start arriving in period 2, their subjective hazard of receiving such offers in the next period, $\hat{\lambda}$, equals the objective hazard $\lambda\eta$. Proposition 1 thus implies that, in either case, their reservation wage is w^{*H} if this satisfies (5) when $b = \gamma w^{*H}$. Using (57), this requires that

$$(D + \rho\lambda\eta)w^{*H} = (D\gamma + \rho\lambda\eta)(\delta_2 R + (1 - \delta_2)w^{*H})$$

which is equivalent to (44). When combined with (57) this implies that

$$\frac{\bar{w}_2}{R} = \frac{\delta_2(D\gamma + \rho\lambda\eta)}{D(1 - \gamma) + \delta_2(D\gamma + \rho\lambda\eta)}$$

Now turn to individuals of type L. Suppose that they accept all viable offers above w^{*L}, so that their estimated and their actual hazard of leaving unemployment is λ while their average wage $\bar{w}^L$ is given by the expression in (29). Given that the people of type L are naive group learners, w^{*L} is a stable reservation wage for them as long as it satisfies (5) with $b = \gamma\bar{w}^L$. Noting that $\bar{w}^L = \bar{w}_1 + (1 - \lambda)\eta(\bar{w}_2 - \bar{w}_1)$ and using (57) for $\bar{w}_1$, this requires that

$$(D + \rho\lambda\eta)w^{*L} = (D\gamma + \rho\lambda\eta)\Big((1 - \eta(1 - \lambda))(\delta_1 R + (1 - \delta_1)w^{*L}) + \eta(1 - \lambda)\bar{w}_2\Big).$$

Using the solution for $\bar{w}_2$ above, this is equivalent to (45).

With this worker behavior the total number of employees and unemployed individuals of types L and H are given by (11) and (12) respectively. Given that the number of people of type i who become unemployed in the current period U^{i1} is given by σM^i, we have

$$\frac{U^H - U^{H1}}{U^L - U^{L1}} = \frac{(\sigma + \lambda)(1 - N^L)}{(1 - \lambda)(\sigma + \lambda\eta(1 + \sigma))N^L} \qquad \frac{U^{H1} + (1 - \eta)(U^H - U^{H1})}{U^{L1} + (1 - \eta)(U^L - U^{L1})} = \frac{(\sigma + \lambda)(1 - N^L)}{(\sigma + \lambda\eta(1 + \sigma))N^L}$$

As a result, (46) ensures that (41) holds for $i = 2$ while (47) ensures that (42) holds for $i = 1$.