Group Learning, Wage Dispersion, and Nonstationary Offers

Julio J. Rotemberg*

This version: October 15, 2014

Abstract

It is shown that differences in beliefs can be an important source of inequality even if everyone is equally productive and people are reasonably sophisticated in the way that they learn about their economic environment. As is standard in the search literature, people believe that the wage offers they obtain while searching for a job are drawn from a stationary distribution. They then base their job acceptance decision on the average wage and the average unemployment duration of people who belong to their peer group. If, in fact, the distribution of wage offers is not stationary and 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 capable of identifying multiple viable jobs. Equilibria can then feature groups that accept low wage offers from prominent ads while other groups do not find these offers acceptable. A new statistic is proposed whose application to existing studies suggests that the nonstationarity considered here is present in data. JEL: D83, J31, J64

*Harvard Business School, Soldiers Field, Boston, MA 02163, jrotemberg@hbs.edu. I wish to thank Francis Kramarz, Pawel Krolikowski, Benjamin Schoefer and audiences at Ca’ Foscari University, the Banque de France, the University of Michigan, and the NBER Summer Institute for comments. All errors are my own. I also thank the Harvard Business School Division of Research for support.
Piketty (1995) and Di Tella and Dubra (2008) show that beliefs can affect productivity, which can in turn affect inequality. But can differences in beliefs be an important source of inequality even if people are equally productive and, in addition, are reasonably sophisticated in the way that they learn about their economic environment? This paper shows that beliefs can play this role in the context of a labor market search model that stays close to Hornstein, Krusell and Violante (2011) (HKV) and is consistent with at least some of the salient facts concerning the dynamics of unemployment. Two key assumptions yield this result. First, people use observations of their peers’ average labor market outcomes to assess the benefits of continuing to search by remaining unemployed. Second, these observations are insufficient to get a full picture of their prospects so that some people have incorrect beliefs about the gains to continued search.\(^1\)

One issue that turns out to play a key role is the stationarity of the wage distribution from which workers draw wage offers as they search. The standard assumption in the extensive literature starting with McCall (1970, p. 115) is that the wages accompanying “[j]ob offers” are “independent random selections” from a constant distribution. When people base their job acceptance decision on observations of peers, labor market outcomes depend on both the extent to which this assumption is true and the extent to which workers believe it to be true. If the wage offers that people receive are drawn from an invariant distribution and people believe that this is so, the learning process I propose leads all groups to end up with identical reservation wages and identical wage distributions. Moreover, this common reservation wage is the one that would be chosen by people with rational expectations.

Groups go astray, however, if their members interpret group-level averages under the supposition that wage offers are drawn from a constant distribution when this is not the case. This paper focuses on a particular form of nonstationarity in which people who are at the beginning of their unemployment spell typically receive less attractive offers than people who are more experienced job searchers. This form of nonstationarity does not appear to

---

\(^1\)While there is no direct evidence on the accuracy of these beliefs, there is evidence that some students underestimate the gains they would accrue by staying in school (Jensen 2010).
have been considered in the labor market context before, though it has been studied in product markets. This nonstationarity arises naturally if some job opportunities are more visible (or prominent) than others.

As a prototypical example, small retail establishments often post their personnel needs on windows that face the street, ensuring that they are easily seen by passerby. And, consistent with the idea that prominence allows firms to pay low wages, 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. Janitors that work in Banking or Insurance, industries where help wanted advertising on the premises is more unusual, earn more. As workers increasingly search for jobs online, prominence may come to depend less on geography and more on willingness to pay for visibility.

If workers knew that early offers with low wages would be followed by more attractive offers later, they would wait for the latter. Unfortunately, this nonstationarity is difficult to detect. While I argue below that there is some evidence for it in the existing empirical literature, the scholarly evidence on this is sparse. Leaving this aside temporarily, individuals considering a job offer will naturally ask their peers about the costs and benefits of staying unemployed. For this purpose, their peers’ mean wage and mean duration of unemployment seem particularly informative.

What seems much harder is for peer groups to provide individuals with wage averages that are conditional on the length of an unemployment spell. One reason for this difficulty is that the sources of information about wages and unemployment durations are likely to be different. Discussing wages is sensitive, so wage information is mostly obtained from close friends. The information these individuals have about how long it has taken them to find a job will tend to be stale and hard to recall. On the other hand, information about unemployment durations is both less sensitive and more observable. Therefore, good infor-

---

3This is already standard practice in the case of ads for goods and services displayed by search engines.
4The length of unemployment depends also on the intensity with which people search. While I abstract from the variability of this intensity below, the fact that it is largely unobservable provides another rationale for the absence of reliable wage information that is conditional on how much people have already searched.
information about it is available from a wider group and one will tend to ignore one’s close friends poorly remembered experiences. This disconnect in sources of information is mirrored by the practices of statistical agencies. Governments obtain most of their information regarding wages from samples of currently employed workers, while they gather information about the typical duration of unemployment from workers who are either currently searching or just finished searching for jobs. Armed only with these distinct pieces of information, one is not compelled to depart from the simple McCall (1970) hypothesis that the distribution of wage offers is independent of how long one has been searching. And, indeed, this hypothesis remains ubiquitous in the economics literature.

If every worker had access to the same economy-wide information, wage inequality would not be greater than in HKV. Inequality can be more substantial if different groups reach different conclusions about the distribution of wages. This turns out to be possible even if everyone views offers as stationary. To understand this possibility, imagine that one group accepts early low wage offers while the other has a higher reservation wage. Having access only to their own group’s statistics, members of the former group expect future offers to have low wages as well, and are therefore more inclined to accept such offers than members of the latter group.

But should one believe that a significant number of workers receives lower-paying offers when they are just starting their unemployment spell? This paper proposes a statistic that measures departures from the null hypothesis that job offers are stationary. This measure is the ratio of the percent increase in the average wage people obtain when they raise their reservation wage to the percent increase in the length of their unemployment spell. Holzer (1986) and DellaVigna and Paserman (2005) report that, in many cases, the wage gains from higher reservation wages are substantial while the resulting lengthening of unemployment spells is small. This is inconsistent with stationarity, which requires that the ratio above be equal to the percent difference between the average and the reservation wage divided by one hundred.

This paper also provides a theoretical reason for offers to be nonstationary, namely that
inexperienced job seekers have a lower likelihood of having access to two attractive competing offers. Following Lang (1991), competing offers are possible if workers investigate several job possibilities simultaneously. This paper adds the idea that inexperienced job searchers are more likely to waste their time on dead ends, and are accordingly more likely to have a single offer. Moreover, this single offer is more likely to originate from a firm whose job posting is relatively easy to find (i.e. prominent). This implies that firms whose job offers are prominent are less likely to face direct competition for workers, and this leads them to offer lower wages in equilibrium.

The paper proceeds as follows. Section 1 introduces group learning with an exogenous offer distribution. It focuses on a particularly simple illustrative case in which the wages offered to people in their first period of unemployment are lower than those offered subsequently. Even if all workers expect offers to be stationary, this can result in substantial inequality as one group accepts low wage offers and another does not. In spite - or perhaps because of - this heterogeneity, the average hazard of leaving unemployment can be constant as in Machin and Manning (1999).

Starting with Section 2, the distribution of wages is endogenous. In particular, it 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. Subsection 2.2 studies the case where everyone has accurate beliefs while Subsection 2.3 focuses on the case where everyone belongs to a group that learns from economy-wide statistics. Subsection 2.4 then shows that much more inequality is possible if people only have access to statistics concerning the average outcomes of their own distinct group. Section 3 discusses a statistic that measure departures from stationarity. Section 4 offers some concluding remarks.

**Related Literature**

This paper fits into the literature in which people have an imperfect understanding of the
stochastic process of job offers (Rothschild 1974, Nishimura and Ozaki 2004). This imperfect understanding degrades the decisions people make, which provides a rationale for learning from one’s peers, as in Bala and Goyal (2010) and here. In their setting, Bala and Goyal (2010) show that the equilibrium actions of individuals who are connected to each other in a group all have the same payoff. This corresponds to choosing an identical reservation wage in my setting. A key difference between Bala and Goyal (2010) and the current paper is that they suppose that each individual observes the full history and payoffs of her neighbors, and this makes it possible to converge to the optimum. By contrast, here agents only observe some summary statistics concerning the actions and payoffs of their group.

For simplicity, I suppose that the information-sharing groups that people belong to have no overlap, and this assumption is clearly unrealistic. Indeed, Watts (1999) notes that most people in social networks are connected to all others in the sense that one can find a set of pairwise links that lead from one member of the network to any other. This paper is based on a different feature of social networks that is also stressed by Watts (1999): social networks are clustered in the sense that a large fraction of the people who are linked to an individual have direct links with one another. By contrast, the links across people in different clusters are sparser. Onnela et al. (2007) show that this is not just a property of the number of connections that people have with each other but also of the time they spend exchanging information. They show that communication across groups is relatively rare and that information is “trapped in communities” by demonstrating that the largest connected component of their network disintegrates if one eliminates links that involve relatively short amounts of time spent in conversation. In the context of the current paper, this can be interpreted to mean that the transmission of statistics that help people decide whether to accept jobs require regular and time-consuming exchanges.

This paper focuses on the transmission of information about the overall environment as opposed to about individual opportunities in this environment. The latter is the focus of the literature on peer referrals (Montgomery 1991, Calvó-Armengol and Jackson 2004, Dustmann, Glitz, and Schönberg 2011). These papers assume that workers all have correct
beliefs regarding the distribution of wage offers. As a result, peer referrals enrich a worker’s opportunity set and can only increase wages. Empirically, on the other hand, several studies have shown that jobs obtained through peer referrals carry lower wages.\textsuperscript{5}

This paper is more closely related to two papers that focus on the effects of beliefs about the labor market as a whole on the search behavior of individuals. Rotemberg (2002) considers beliefs about the “fairness” of wage determination and shows that the nature of these beliefs can affect inequality. Ahmed (2008) considers beliefs about discrimination and shows that wage inequality across groups can persist if some groups adhere to the belief that they will never be selected if they apply for jobs with high wages. These two papers neglect learning so they leave open the question studied here, namely how much inequality can persist if a reasonable amount of learning is allowed.

1 Group Learning with Exogenous Wages

All individuals have access to the same job offers, time is discrete, and the total population is normalized to one. Workers are risk neutral and use the factor $\rho$ to discount future earnings. Each period that a worker is unemployed, she investigates job postings. By carrying out time consuming research, the worker learns the characteristics of jobs, including their wages $w$. The non-wage characteristics of postings make them unsuitable for a randomly selected subset of workers. For simplicity, all suitable (or “viable”) job postings have the same non-pecuniary benefits. At the end of the period, the unemployed individual must decide whether to accept any of the viable jobs she has found. If she does, she earns the job’s posted wage $w$ in every period that she remains on this job. Anyone who has a job at a point in time has a probability $\sigma$ of not having it in the next period and returning to unemployment. In any

\textsuperscript{5}For examples, see Loury (2006), which contains several references from the sociology literature, Bayer, Ross and Topa (2008), and Bentolila, Michelacci and Suarez (2010). In an indirect study of referrals that relies on the extent to which firms hire minorities, Dustmann, Glitz, and Schönberg (2011), find that minority workers working for firms where their own minority constitutes a larger share of employment (which they interpret as implying that members of this minority have been referred) have lower wages. They put more emphasis on the results with the opposite sign that they obtain after including firm fixed effects but it is not clear that this provides an improved measure of the effect of referrals.
period the worker is unemployed, she earns \( b \).

This paper studies steady states. In most of the analysis, people also believe their environment is stationary in the sense that the probability of getting a job offer with a particular wage is constant over the course of their own unemployment spell. As a result, their optimal strategy consists of accepting offers above a reservation wage. People belong to groups whose reservation wage is the same because they have the same information. Given a group’s reservation wage \( w^* \), the actual stochastic process of offers determines the group’s mean wage \( \bar{w}(w^*) \) and its mean duration of unemployment \( S(w^*) \).\(^6\) These are the two key observations made by members of the group.\(^7\)

I focus on outcomes that are “stable” or, to use Fudenberg and Levine’s (1993) terminology, “self confirming” in the sense that the data that people see lead people to take actions that reproduce the data they see. Here this means that the observed values of \( \bar{w} \) and \( S \) must make it optimal for individuals to choose a reservation wage \( w^* \) that leads to these values of \( \bar{w} \) and \( S \). If people believe the environment is stationary, the conditions required for an outcome to be stable turn out to be similar to conditions for optimality that have been derived under rational expectations.

Let \( U \) represent the expected value of being unemployed at a stable outcome according to members of a group. Then, employed members of the group believe that the expected present value of having a job that pays \( w \) is

\[
V(w) = w + \rho \left( \sigma U + (1 - \sigma) V(w) \right) = \frac{w + \rho \sigma U}{1 - \rho (1 - \sigma)},
\]

where the second equality follows from the first. On the basis of their belief in stationarity and their observation of the group’s outcomes, people believe that an unemployed person’s probability of finding a job that pays \( \bar{w} \) in any given period is \( \hat{\lambda} = 1/S \). Given that \( V \) is

\(^6\)I ignore the problem of estimating \( S \) that result from the fact that, in any given small sample, longer spells of unemployment are underrepresented relative to shorter ones.

\(^7\)If, as in McCall (1970), the actual environment is stationary in that offers drawn from \( G(w) \) arrive with Poisson probability \( \lambda \), \( \bar{w} = E_G(w | w \geq w^*) \) and \( \hat{\lambda} = \lambda(1 - G(w^*)) \) but neither of these relationships needs to hold if the actual process of receiving offers is nonstationary.
linear in $w$, the value of $U$ that is consistent with the group’s beliefs is thus

$$ U = b + \rho \left( \hat{\lambda} V(\bar{w}) + (1 - \hat{\lambda})U \right) = \frac{Db + \rho \hat{\lambda} \bar{w}}{(1 - \rho)(D + \rho \hat{\lambda})} \quad \text{where} \quad D \equiv 1 - \rho(1 - \sigma). \quad (2) $$

Here, the second equality is obtained by using (1) to substitute for $V(\bar{w})$ and solving the resulting equation for $U$.

If workers believe this value of $U$, they find it optimal to accept any offer whose wage $w$ makes $V(w) \geq U$. Since $V$ is increasing in $w$, this leads them to accept jobs that pay more than a reservation wage $w^r$. When wages can take on a continuum of values, $w^r$ satisfies $V(w^r) = U$. Moreover, (2) implies that $U$ depends only on $\bar{w}$ and $\hat{\lambda}$ and, as discussed above, both these variables depend only on the group’s reservation wage $w^*$. Thus, an outcome is stable under group learning if $w^r$ equals $w^*$. When wages take continuous values

$$ V(w^*) = U(\bar{w}(w^*), \hat{\lambda}(w^*)). $$

Using equation (1) to substitute for $V(w^*)$, the reservation wage is then

$$ w^* = (1 - \rho)U(w^*, \bar{w}, \hat{\lambda}) = \frac{Db + \rho \hat{\lambda} \bar{w}}{D + \rho \hat{\lambda}}, \quad (3) $$

where the second equality follows from (2).

### 1.1 Stationary Wage Offers

While (3) was derived without imposing rational expectations, (1) is valid under this assumption as well, as long as $U$ is the correct present discounted value of earnings for someone that is unemployed. The equation determining $U$, (2), would be valid under rational expectations as well if offers paying more than $w^*$ had a constant arrival probability $\hat{\lambda}$ and always had the same mean wage $\bar{w}$. It is thus not surprising that,

**Proposition 1.** If the environment is stationary so that there is a constant probability $\lambda$ that an unemployed worker finds at least one viable offer in any given period while the highest wage paid by the viable offers she finds has a constant cumulative density function $G(w)$, the
rational expectations optimum coincides with the unique stable group learning outcome. The resulting reservation wage satisfies (3) with \( \hat{\lambda} = \lambda(1 - G(w^*)) \) and

\[
\bar{w} = \frac{\int w dG(w)}{1 - G(w^*)}.
\]

As shown by HKV, the ratio of mean to minimum wages is quite small at this rational expectations optimum when plausible parameters are used. To make this point transparent, they suppose that \( b \) is equal to \( \gamma \hat{w} \) for all workers. Equation (3) then implies that

\[
\frac{\bar{w}}{w^*} = \frac{D + \rho \hat{\lambda}}{D\gamma + \rho \hat{\lambda}}.
\]

On the basis of economy-wide average separation rates, job finding rates and real interest rates HKV set \( \sigma, \hat{\lambda} \) and \( \rho \) to .02, .39 and .9959 respectively. Lastly, they set \( \gamma \) equal to .4 on the basis that smaller numbers are both implausible given the rules governing unemployment insurance and make it harder to explain economic fluctuations. These numbers imply that the ratio of mean to minimum 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.

1.2 High Wage Offers Start Arriving in Period 2

This subsection considers an extremely stark example of nonstationarity to provide intuition for the more general results that follow. Unemployed workers have a constant probability \( \lambda \) of finding a viable job in every period. All viable job offers found by workers in their first period of unemployment carry a wage of \( w_1 \) while, in subsequent periods, the most attractive viable job pays a wage \( w_2 > w_1 \) with probability \( \eta \) and pays \( w_1 \) with the remaining probability. Except in the uninteresting case in which \( b > w_2 \), offers of \( w_2 \) are accepted since there are no higher wages worth waiting for. The critical issue is whether workers also accept offers of \( w_1 \). I start by giving conditions under which they would accept them if they had rational expectations.
Proposition 2. Individuals with accurate beliefs accept offers of $w_1$ if and only if

$$w_1 \geq W_1 = \frac{Db + \rho \lambda \eta w_2}{D + \rho \lambda \eta}. \quad (6)$$

When (6) holds, there is no stable group learning outcome in which a group that believes the environment is stationary turns down offers of $w_1$.

The reason that group learners accept offers of $w_1$ whenever workers with accurate beliefs do so is the following. If a group were to confine itself to accepting only $w_2$ offers, its average wage would be $w_2$ and its average duration of unemployment would be $(1 + 1/\lambda \eta)$. Workers who believe the environment is stationary and use these statistics would infer that $w_2$ offers arrive with probability $\lambda \eta/(1 + \lambda \eta)$ per period, which is lower than the true probability $\lambda \eta$. Thus, compared to workers with accurate beliefs, newly unemployed members of such a group would be more keen to accept a first period offer of $w_1$.

This difference between group learners and people with accurate information matters when (6) is violated so that latter do not accept jobs in the first period. Workers who learn from their group now accept $w_1$ offers if $w_1$ exceeds $U$, where this expectation depends on the reservation wage used by the group. There are two reservation wages for which $U$ can be computed. First, if the group’s reservation wage is $w_2$, the belief that such offers arrive with probability $\lambda \eta/(1 + \lambda \eta)$ implies that the value of $U$ in (2) is

$$U_2 = \frac{Db(1 + \lambda \eta) + \rho \lambda \eta w_2}{(1 - \rho)[D(1 + \lambda \eta) + \rho \lambda \eta]}.$$

Second, if the group’s reservation wage is $w_1$, the group’s unemployment duration is $1/\lambda$ and its average wage is

$$\lambda w_1 + (1 - \lambda)[\eta w_2 + (1 - \eta)w_1] = w_1 + \eta(1 - \lambda)(w_2 - w_1).$$

In the first equality, the first term captures that a fraction $\lambda$ of newly unemployed workers find a job in the first period and the second reflects that, among those who find their job later, a fraction $\eta$ receives $w_2$ while the rest receives $w_1$. Using (2), the expected value of being unemployed for this group is then

$$U_1 = \frac{Db + \rho \lambda [(1 - \eta(1 - \lambda))w_1 + \eta(1 - \lambda)w_2]}{(1 - \rho)(D + \rho \lambda)}.$$
As the following proposition demonstrates, it is possible to find values of \( w_1 \) such that \( V(w_1) \) exceeds \( U_1 \) and/or \( U_2 \) even when (6) is violated so that employed people with accurate beliefs all earn \( w_2 \).

**Proposition 3.** As long as \( \lambda < 1, \) and \( \eta < 1 \) and \( b < w_2 \)

\[
\frac{Db + \rho \lambda \eta (1 - \lambda) w_2}{D + \rho \lambda \eta (1 - \lambda)} < (1 - \rho) U_2 < W_1. \tag{7}
\]

For \( w_1 \) between \((1 - \rho) U_2 \) and \( W_1 \), the only outcome that is stable under group learning is for workers to accept offers of \( w_1 \). For \( w_1 \) between the left hand side of (7) and \((1 - \rho) U_2 \), there is both a stable group learning outcome in which offers of \( w_1 \) are accepted and one where they are rejected. Finally, for \( w_1 \) below the left hand side of (7), only offers of \( w_2 \) are accepted at a stable outcome.

Groups that believe offers are stationary accept low wages when workers with accurate beliefs do not for two reasons. First, the former underestimate the probability of receiving an offer of \( w_2 \) when their group’s reservation wage is \( w_2 \). This effect implies that, when \( w_1 \) is between \((1 - \rho) U_2 \) and \( W_1 \), there is no stable outcome in which members of such groups turn down offers of \( w_1 \). Second, if such groups do accept offers of \( w_1 \) (which are particularly common in the first period), their members expect that there is a larger chance that they will only be offered \( w_1 \) in the future. They thus feel they have less to gain by rejecting a current offer of \( w_1 \). This logic leads to multiple equilibria when \( w_1 \) is between \( U_1 \) (or, equivalently, the left hand side of (7)) and \((1 - \rho) U_2 \).

Inequality among members of a group that accepts offers of \( w_1 \) is limited by the requirement that \( w_1 \) must exceed the left hand side of (7). This is a manifestation of the more general inequality-limiting condition (5). It is possible, however, for \( w_2 \) to be significantly larger than \( w_1 \). In particular, (7) allows \( w_1 \) to be significantly lower than \( w_2 \) as long as \( \lambda \eta (1 - \lambda) \) is small. One way this can be accomplished is by having a high value of \( \lambda \), which leads many unemployed workers to encounter an offer of \( w_1 \) in the first period. If they accept it, the average wage ends up being near \( w_1 \) so that inequality within the group as measured by \( \bar{w}/w^* \) is low even if \( w_2 \) is high.
On the other hand, overall inequality can be large as long as there are two groups of workers with different reservation wages. To see this, suppose 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 their beliefs. Given Proposition 3, both types can believe that offers are stationary and base their decisions to accept jobs on the average wage and average unemployment duration of their group. The overall average wage in the economy is then

$$\bar{w} = n_L[w_1 + \eta(1 - \lambda)(w_2 - w_1)] + (1 - n_L)w_2 = w_1 + (1 - n_L + n_L\eta(1 - \lambda))(w_2 - w_1). \quad (8)$$

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 this now requires only that $w_1$ exceed the left hand side of (7). If beliefs were assumed to be accurate, $w_1$ would only be accepted if it were larger than the right hand sides of (6) and of (7). This obviously leads to a lower $w_2/w_1$.

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 $\gamma$ of average wages. Because there are two distinct groups, however, I let the $b$ of each type equal $\gamma$ 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 $w_1$ being larger than the left hand side of (7) is

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

Using this to substitute for $w_2/w_1$ in (8), the maximum value of $\bar{w}/w_1$ is

$$1 + \frac{D(1 - \gamma)}{\eta(1 - \lambda)(D\gamma + \rho\lambda)} \left[1 - n_L + n_L\eta(1 - \lambda)\right].$$

As in HKV, an increase in $\gamma$ lowers this maximum ratio because it makes unemployment more attractive so that the minimum wage workers accept is closer to this mean. What is
more novel is how this maximum depends on $\lambda$. The derivative of the expression above with respect to $\lambda$ is

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

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 fear unemployment less, so that they demand a higher wage. The second term, which is also negative, is a composition effect. A higher $\lambda$ leads more workers of type $L$ to accept offers that pay $w_1$ in the first period, so fewer earn $w_2$ and the mean wage shifts towards the minimum wage $w_1$. The last term is positive, and becomes dominant for large $\lambda$. It captures the idea that a high $\lambda$ 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 $\lambda$ 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 $\lambda$ is sufficiently high. This raises the question of whether such high values of $\lambda$ 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 $\lambda$ itself. A very high value of $\lambda$ would then be incompatible with the fact that about half the people who become unemployed do not accept a job during their first month of unemployment.

As I now show, this problem need not arise when groups differ in their beliefs, though the computation of aggregate hazard rates is more complex in this case. One complexity is that the willingness of workers of type $L$ to accept lower wages implies that they leave unemployment more quickly. People of type $L$ thus have a higher employment rate, which in turn means that a disproportionate number of the people who become unemployed in a particular period are of this type.

More formally, 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 $u_i'$ the number of people of type $i$ who are freshly unemployed in a given period. Then, use $\Psi$ to denote the fraction of
newly unemployed people who are of type $L$. Since $u_f^L$ is equal to $\sigma m^i$ for both types and $\Psi$ is defined by $u_f^L/(u_f^L + u_f^H)$, $\Psi$ equals $m^L/(m^L + m^H)$.

For people of type $L$, the outflow from unemployment equals $\lambda u^L$, and this must equal the inflow $\sigma m^L$. Given that $u^L + m^L = n^L$, the equality of $\lambda u^L$ and $\sigma m^L$ implies that

$$m^L = \frac{\lambda}{\lambda + \sigma} n^L, \quad u^L = \frac{\sigma}{\lambda + \sigma} n^L. \quad (9)$$

Individuals of type $H$ start receiving acceptable offers only in the second period of unemployment so that their outflow from unemployment equals $\lambda \eta (u^H - u_f^H)$, where $u^H_f$ equals the inflow into unemployment $\sigma E^H$. Since $u^H + m^H = (1 - n^L)$, we have

$$m^H = \frac{\lambda \eta (1 - n^L)}{\lambda \eta + \sigma (1 + \lambda \eta)} \quad \quad u^H = \frac{\sigma (1 + \lambda \eta)}{\lambda \eta + \sigma (1 + \lambda \eta)}. \quad (10)$$

After some rearranging, $\Psi$ 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 $\Psi (u_f^L + u_f^H)$. Because these people have a probability $\lambda$ of accepting a job in each period that they are unemployed, a fraction $(1 - \lambda)^{\tau - 1}$ of them also experience a $\tau$'th period of unemployment, while a fraction $\lambda (1 - \lambda)^{\tau - 1}$ finds a job during this $\tau$'th period.

Similarly, $(1 - \Psi) (u_f^L + u_f^H)$ 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 $\tau$'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 $\tau$'th period of being unemployed is the fraction of the individuals who become unemployed in any given period that find jobs in the $\tau$'th period of their unemployment divided by the fraction that spent this $\tau$'th period unemployed. It thus
equals $\lambda \Psi$ in the first period. For $\tau \geq 2$ it equals
\[
\frac{\lambda(1 - \lambda)^{\tau-1}\Psi + \lambda \eta (1 - \lambda \eta)^{\tau-2}(1 - \Psi)}{(1 - \lambda)^{\tau-1}\Psi + (1 - \lambda \eta)^{\tau-2}(1 - \Psi)}.
\] (11)

Since $(1 - \lambda)^{\tau-1}$ becomes negligible faster than $(1 - \lambda \eta)^{\tau-2}$, this converges to $\lambda \eta$ as $\tau$ gets large though it differs from this exact value unless $\lambda$ is equal to one. There is thus only one case in which the hazard is constant, and equal to $\eta$. This is when $\lambda$ equals one while $\eta = \Psi$. Insofar a constant hazard is a good approximation to what is observed, a high value of $\lambda$ appears justified.\(^8\) Admittedly, the case where $\eta = \Psi$ is somewhat arbitrary since there is no force in the model that requires the fraction of workers who accept low wages to correspond to the fraction of jobs that offer $w_2$ from period 2 onwards.

2 Endogenous Wages

The purpose of Section 1.2 is to show that, in an extremely stylized setting, the presence of first period offers with relatively low wages can lead to dispersions in beliefs and inequality. This raises the question of whether it is at all plausible for the offers that unemployed individuals receive early in their unemployment spell to have lower wages than the offers they receive subsequently. This section shows that this can be an endogenous consequence of supposing that workers are less sophisticated in the potential jobs that they explore when they start their unemployment spell. The idea is that firms post job ads and that unemployed individuals spend time finding out both how suitable these jobs are for themselves along nonwage dimensions and learning their wage.

The nonstationarity assumed in the previous section then emerges endogenously if workers are more susceptible to opportunities that are simultaneously visible and relatively likely to be inappropriate. Such workers are then unlikely to find two different jobs that are appropriate for them. The resulting lack of competition for workers should lead firms whose ads are relatively visible to offer relatively low wages. To show how inattention combines with

\(^8\)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.
social learning, I start with a discrete model that includes neither so that all jobs are equally visible at all times and individuals have accurate beliefs. I then introduce inattention while maintaining accurate beliefs and follow this with a setup that includes both inattention and social learning. Here too, I start with the simplest case, namely when all workers learn from the average outcomes of the entire population and then conclude by demonstrating an equilibrium in which two groups have different beliefs.

2.1 Stationary Wages in a Burdett-Judd Model

Firms can post job offers at a cost of $c$ and have productive opportunities in which the per-period marginal product of labor is a constant $R$. The steady state number of job postings equals $v$. Once a worker accepts a posting, he becomes employed. As before, employees face a probability $\sigma$ 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

$$\Pi(w) = R - w + \rho(1 - \sigma)\Pi(w) = \frac{R - w}{D}. \quad (12)$$

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 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 acceptable job offers. 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 also end with the worker having identified more than one opportunity.\footnote{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.} 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.

Now consider a firm posting a particular offer. The probability that a particular unemployed worker ends up viewing this as her only viable offer is $\lambda(1 - \delta) / v$ while the probability that she ends up considering this to be one of her two viable offers $2\lambda\delta / v$. Thus, the probability that an unemployed worker who views a particular offer as viable also has a competing viable offer is $2\delta/(1 + \delta)$. With the remaining probability, $(1 - \delta)/(1 + \delta)$, she has no other viable offer.

Let $Q$ represent the expectation of the total number of viable contacts made by all offers. Since unemployed workers have a probability $\lambda(1 - \delta)$ of finding one viable offer and a probability $\lambda\delta$ of finding two, $Q$ must equal $\lambda(1 + \delta)$ times the number of unemployed workers $u$. For a particular offer, then, the expected number of workers that find it viable in a period is $Q/v$.

Since any firm can post an ad with a wage $w$ at cost $c$, profits must be the same for any $w$ that is actually offered. 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, as 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 posting of job that pays $w$ leads to an expected present value of profits of

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

A free entry equilibrium requires that $cv$ adjust so that profits are zero for any $w$ in the
support of $F$ and negative for wages outside this support. Therefore,

$$F(w) = \frac{\theta}{2\delta(R-w)} - \frac{1-\delta}{2\delta} \quad \text{where} \quad \theta \equiv \frac{cvD}{\lambda U}. \tag{13}$$

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

$$w^* = R - \frac{\theta}{1-\delta} \quad w^x = R - \frac{\theta}{1+\delta}. \tag{14}$$

These equations still depend on the endogenous value of $cv$ (or $\theta$). For $\theta$ 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 $\theta$) must adjust so that (5) is satisfied as well. To calculate this equilibrium $\theta$, 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 (13), this becomes

$$G(w) = \frac{\theta^2}{4\delta(R-w)^2} - \frac{(1-\delta)^2}{4\delta} = \frac{(1-\delta)^2}{4\delta} \left( \left( \frac{R-w^*}{R-w} \right)^2 - 1 \right), \tag{15}$$

where the second equality uses the first equation in (14). The average wage received by workers is thus

$$\bar{w} = \int_{w^*}^{w^x} wdG(w) = \int_{R-\frac{\theta}{1+\delta}}^{R-\frac{\theta}{1+\delta}} \frac{\theta^2 w dw}{2\delta(R-w)^3} = R - \theta, \tag{16}$$

where the substitution of $w$ by $R-w$ in the integral simplifies the computation that leads to the last equality. Combining (16) and the first equation in (14) yields

$$\bar{w} = \delta R + (1-\delta)w^*, \tag{17}$$

which makes it clear that the effect of competition in the form of a higher $\delta$ is to raise offers towards $R$. This linear equation in $\bar{w}$ and $w^*$ can be combined with (3), which is also linear.
in these two variables in the case where the finding rate is given by a constant \( \lambda \). This yields the equilibrium minimum wage

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

(17)

Differentiation of this equation leads to the conclusion that the minimum wage is increasing in the level of competition \( \delta \) and in the arrival rate \( \lambda \). 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 \( \lambda \), it is necessary that the cost of posting jobs decline in \( \delta \). This is indeed what happens, as is apparent once one uses the first equation of (14) to obtain \( \theta \):

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

Since \( \theta \) is proportional to \( cv \), this shows that an increase in \( \lambda \) 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 one in which the cost per posting \( c \) is constant and one in which the total number of postings \( v \) is fixed while their cost \( c \) is determined in a competitive market.

### 2.2 Nonstationary Wages with Accurate Beliefs

Nonstationarity is introduced in this section by supposing that recently unemployed workers are both less likely to have access to viable competing job offers and more likely to see certain job ads which I call “prominent.” “Ordinary” ads, by contrast are never seen the first period of unemployment. Workers who see prominent rather than ordinary ads are less likely to identify two viable competing offers. This 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, and so 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 any event,
both the susceptibility to prominent ads and the low likelihood of finding two viable offers seem like natural consequences of inexperience and inattention. As unemployed workers gain experience they become less susceptible to prominent ads, and they have both a better chance of seeing ordinary ones and of finding two viable jobs.

I capture these ideas formally by supposing that workers can be associated with an indicator variable denoted by $\kappa$. When workers are susceptible to prominent adds, $\kappa$ equals $s$. When they are resistant, and see only ordinary ads, $\kappa$ equals $r$. Because of inexperience, a worker’s $\kappa$ is $s$ in the first period she is unemployed. Afterwards, $\kappa$ equals $r$ with probability $\eta$ and $s$ with probability $(1 - \eta)$. Because workers in state $s$ are inattentive also in the sense of being less likely to find two viable jobs, $\delta$ depends on $\kappa$ as well. In particular, the probability that an individual whose indicator variable is $\kappa$ finds two viable jobs in a period is $\lambda \delta_s$, the probability she finds only one is $\lambda(1 - \delta_s)$, and $\delta_r > \delta_s$.

Prominent ads reach only workers whose $\kappa$ is $s$ while ordinary ads influence only workers with a $\kappa$ of $r$. Supposing that firms specialize in the kind of ad they post, it is only a slight abuse of language to say that firms that post the former (or the ads that they post) are of type $s$ while the latter ads and the firms that post them are of type $r$. One can then refer to the number of ads of type $i$ and their cost by $v_i$ and $c_i$ respectively, where $i$ is either $r$ or $s$.

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 $w$ 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, $\lambda$ is once again the hazard rate of leaving unemployment. Total employment $m$ thus equals $\lambda/(\lambda + \sigma)$ and the number of people that are unemployed for the first time in any given period, $u_f$, equals $\sigma \lambda (\lambda + \sigma)$.

Let $Q_i$ denote the total number of viable jobs offered to unemployed individuals by firms
of type $i$. Recall that this is the same as the number of viable jobs found by all individuals whose $\kappa$ equals $i$. A single individual whose indicator variable is $\kappa$ can expect $\lambda (1 + \delta_\kappa)$ contacts with viable offers because she expects $\lambda (1 - \delta_\kappa)$ contacts with a single offer and $\lambda \delta_\kappa$ contacts with two. Therefore,

$$Q_s = (1 + \delta_s)\lambda(u_f + (1 - \eta)(U - u_f)) = \frac{\lambda \sigma (1 + \delta_s)(1 - \eta(1 - \lambda))N}{\lambda + \sigma}$$ (18)

$$Q_r = (1 + \delta_r)\lambda \eta (u - u_f) = \frac{\lambda \sigma (1 + \delta_r)\eta(1 - \lambda)N}{\lambda + \sigma}.$$ (19)

A firm of type $i$ with a wage of $w$ greater than or equal to workers’ reservation wage thus earns expected profits equal to

$$\frac{Q_i}{v_i} \frac{R - w}{D(1 + \delta_i)} \left(1 - \delta_i + 2\delta_i F_i(w)\right) - c_i,$$ (20)

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, (20) is zero for all wages that are actually offered so that we have

$$F_i(w) = \frac{\theta_i}{2\delta_i(R - w)} - \frac{1 - \delta_i}{2\delta_i},$$

where $\theta_i \equiv \frac{c_i v_i D(1 + \delta_i)}{Q_i}$. (21)

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

$$\frac{\theta_s}{1 - \delta_s} = \frac{\theta_r}{1 - \delta_r} = R - w^{*a}.$$ (22)

Using this in (21),

$$F_i(w) = \frac{1 - \delta_i}{2\delta_i} \left[\frac{R - w^{*a}}{R - w} - 1\right].$$ (23)

Since $\delta_s < \delta_r$, $F_s(w) \leq F_r(w)$ 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 consequence 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 $\kappa$ is

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

(24)

where the second equality is obtained using the logic that leads to (15). The average wage earned by such workers is thus

$$\bar{w}_\kappa = \int_{w^{\ast\kappa}} w dG_\kappa(w) = R - \theta_\kappa,$$

(25)

where the second equality follows from the argument that leads to (16).

With accurate beliefs, the average wage that unemployed workers can expect to earn from period 2 onward is $\bar{w}_2$ where

$$\bar{w}_2 = (1 - \eta)\bar{w}_s + \eta\bar{w}_r.$$

(26)

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

$$\bar{w} = \lambda\bar{w}_s + (1 - \lambda)\bar{w}_2 = \bar{w}_s + \eta(1 - \lambda)(\bar{w}_r - \bar{w}_s).$$

(27)

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 $V(w)$, which is given by (1). With accurate beliefs, the value $U$ under the strategy of using $w^{\ast a}$ as the reservation wage is

$$U^a = \gamma\bar{w} + \rho(\lambda V(\bar{w}_2) + (1 - \lambda)U^a) = \frac{D\gamma\bar{w} + \rho\lambda\bar{w}_2}{(1 - \rho)(D + \rho\lambda)}.$$  

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

$$w^{\ast a} = \frac{D\gamma\bar{w} + \rho\lambda\bar{w}_2}{D + \rho\lambda}.$$  

(28)

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

\[ w^{*a} = \frac{(D\gamma + \rho\lambda)[\delta_s + \eta(1 - \lambda)(\delta_r - \delta_s)] + \rho\lambda^2\eta(\delta_r - \delta_s)}{D(1 - \gamma) + (D\gamma + \rho\lambda)[\delta_s + \eta(1 - \lambda)(\delta_r - \delta_s)] + \rho\lambda^2\eta(\delta_r - \delta_s)}R. \]  

(29)

Moreover

\[ \frac{\bar{w}}{w^{*a}} \leq \frac{D + \rho\lambda}{D\gamma + \rho\lambda}, \]

(30)

with equality if \( \eta = 0 \) or \( \delta_s = \delta_r \) 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

\[ G(w) = [1 - (1 - \lambda)\eta]G_s(w) + (1 - \lambda)\etaG_r(w) \]

(31)

2.3 Single Group Naive Learning

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 \( \bar{w} \) 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 \( \tilde{w}_2 \) 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 case where everyone learns naively from all employees. The main conclusion from the subsection is that \( w^{sb} \), the common reservation wage when everyone is a naive group learner, is smaller than \( w^{sa} \). For certain parameters, the difference between the two can be substantial.

Most of 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 firms 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 \( w^{sb} \) implies that firms can raise their profits by undercutting this lowest wage. These two facts imply that the hazard of leaving unemployment remains \( \lambda \) for all workers. As a result, (3) implies that

\[
 w^{sb} = \frac{(D\gamma + \rho \lambda)\bar{w}}{D + \rho \lambda}.
\]  

Moreover,

**Proposition 5.** The unique reservation wage when every individual believes that the offer she receives with probability \( \lambda \) has a wage equal to the economy-wide average is

\[
 w^{sb} = \frac{(D\gamma + \rho \lambda)[\delta_s + \eta(1 - \lambda)(\delta_r - \delta_s)]}{D(1 - \gamma) + (D\gamma + \rho \lambda)[\delta_s + \eta(1 - \lambda)(\delta_r - \delta_s)]} R.
\]  

The uniqueness of this equilibrium contrasts with the multiplicity obtained when wages are exogenous in Section 1.2. The reason the equilibrium is unique here is that firms with prominent ads adjust their wages so they prevent equilibria in which individuals simply ignore them.
A comparison of (29) and (33) immediately indicates that $w^a$ is larger than $w^b$. This is to be expected because the unemployed workers who take jobs with probability $\lambda$ 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. With large values of $\lambda$ and $(\delta_r - \delta_s)$, these effects can be dramatic. Suppose that, as in HKV, $\sigma$, $\rho$, and $\gamma$ are set to .02, .9959, and .4 respectively, while $R$ is normalized to equal 1. If $\lambda$, $\delta_s$, $\delta_r$ and $\eta$ are set equal to .99, .01, .8, and .6 respectively, $w^a$ equals .96, while $w^b$ equals .48.

Equation (32) 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 $\lambda$ is assumed to be larger so that workers have even less cause to accept a wage below the mean wage. With accurate beliefs, these parameters yield a ratio of the mean to the minimum wage of 1.0006, which is smaller still.

Because the minimum wage is hard to measure accurately, HKV also report measures of the mean to the 10\textsuperscript{th} 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 (23) as long as one replaces $w^a$ with $w^b$. With the resulting values of $F_i(w)$, the formula in (24) gives the distribution of wages received by workers who accept an offer from a firm of type $i$. Finally, the formula in (31) gives the overall distribution of wages. Therefore the wage such that a fraction $x$ of workers earn less than this wage is given by

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

Using this formula, the ratio of the mean to the 10\textsuperscript{th} 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 10\textsuperscript{th} 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 demonstrate the stability of outcomes in which \( n^L \) workers have a reservation wage \( w^{*L} \), which is lower than the reservation wage \( w^{*H} \) of the rest of the population \((1 - n^L)\). This turns out to be possible if people of type \( L \) accept low wages in period 1 that people of type \( H \) turns down. For this to be true, people of type \( L \) must believe that the wage offer distribution is stationary. People of type \( H \) do not accept period 1 offers. Equilibria in which they do so turn out to exist whether they believe that offers are stationary from the beginning or whether they are aware that they are stationary only starting in period 2. The equations are different in the two cases, however.

The aforementioned requires that the exit rate from unemployment be higher for members of group \( L \). Let \( u^i_j \) and \( u^i_t \) denote, respectively, the number of people of type \( i \) that become unemployed in a period and the number that have been unemployed for \( t \) periods with \( t > 1 \). Then, this implies that \( u^H_t / u^H_j \) must exceed \( u^L_t / u^L_j \) for all \( t > 1 \). Since the total number of unemployed individuals of type \( i \), \( u^i \) equals \( u^i_j \) plus the sum over \( t > 1 \) of \( u^i_t \), we must have

\[
\frac{u^L - u^L_j}{u^L_j} < \frac{u^H - u^H_j}{u^H_j}.
\]

This inequality plays a key role in sustaining equilibria in which prominent ads offer wages below \( w^{*H} \) while ordinary ads offer wages above. It does so by ensuring that ordinary ads are seen mostly by people of type \( H \). Since these have high reservation wages, 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.

Let \( Q^H_i \) denote the viable total contacts between firms of type \( i \) and unemployed workers.
Using the analysis leading to (18) and (19), these $Q$’s are given by

\[ Q^H_s = \lambda(1 + \delta_s)(u^L_f + (1 - \eta)(u^L - u^L_f) + [u^H_f + (1 - \eta)(u^H - u^H_f)]) \]  

\[ Q^H_r = \lambda \eta(1 + \delta_r)(u^L - u^L_f + [u^H - u^H_f]). \]  

The reason for the $H$ superscript is that every one of these contacts find the job acceptable if its wage is greater than or equal to $w^*H$. 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. With this in mind, let $Q^L_i$ equal to the expressions for $Q^H_i$ when the terms in square brackets are set to zero so that they represent the total number of viable contacts of firms of type $i$ with members of group $L$.

The next proposition shows that, thanks to (34), wages below $w^*H$ are more likely to be offered by firms posting prominent ads.

**Proposition 6.** If

\[ \frac{Q^L_i(R - w^*L)}{Q^H_i(R - w^*H)} < 1 \]  

no firms of type $i$ offers a wage below $w^*H$. In this case, free entry implies that the distribution of wages paid by firms of type $i$ satisfies (21) with $Q_i$ replaced by $Q^H_i$. As a result, $\theta_i$ equals $(1 - \delta_i)(R - w^*H)$. If the inequality (37) is reversed, a positive fraction of ads 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 next proposition gives conditions under which wages above $w^*H$ are offered as well as conditions under which an interval of wages below $w^*H$ is not offered even though lower wages are. I construct an equilibrium of this sort below.

**Proposition 7.** If

\[ \frac{Q^L_i(R - w^*L)(1 - \delta_i)}{Q^H_i(R - w^*H)(1 + \delta_i)} > 1 \]  

27
no wage greater than or equal to \( w^{*H} \) is offered by firms of type \( i \). In this case, the distribution across postings of wages offered is given by (21) with \( Q_i \) replaced by \( Q^L_i \). The variable \( \theta_i \) is then equal to \( (1 - \delta_i)(R - w^{*L}) \).

Suppose (37) and (38) 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

\[
\frac{Q^L_i(R - w^{mL})}{Q^H_i(R - w^{*H})} = 1. \tag{39}
\]

For sufficiently low \( \delta_s \), (38) holds for prominent firms whenever (37) 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 (37) hold for \( i = r \) while (38) holds (so that (37) does not) for \( i = s \). The next proposition gives conditions for such equilibria in the case where people of type \( H \) are aware that acceptable offers commence in period 2 and are stationary from then on.

**Proposition 8.** Consider the reservation wages

\[
w^{*H} = \frac{\delta_r(D\gamma + \rho\lambda\eta)}{D(1 - \gamma) + \delta_r(D\gamma + \rho\lambda\eta)} \tag{40}
\]

\[
w^{*L} = \frac{(D\gamma + \rho\lambda)[(1 - (1 - \lambda)\eta)\delta_s + \frac{(1 - \lambda)\eta\delta_s}{D(1 - \gamma) + \delta_r(D\gamma + \rho\lambda\eta)}]}{D(1 - \gamma) + (D\gamma + \rho\lambda)((1 - (1 - \lambda)\eta)\delta_s + (1 - \lambda)\eta)}. \tag{41}
\]

If

\[
1 + \frac{(\lambda + \sigma)(1 - n^L)}{(\sigma + \lambda\eta(1 + \sigma))(1 - \lambda)n^L} \frac{R - w^{*H}}{R - w^{*L}} > 1 \tag{42}
\]

\[
1 + \frac{(\lambda + \sigma)(1 - n^L)}{(\sigma + \lambda\eta(1 + \sigma))n^L} \frac{R - w^{*H}}{R - w^{*L}} < 1, \tag{43}
\]

there exists a free entry equilibrium in which ads of type \( s \) lead to jobs that pay less than \( w^{*H} \) while wages above \( w^{*H} \) are offered by firms of type \( r \). The minimum wage \( w^{*L} \) is stable for the \( n^L \) workers of type \( L \) who learn from their group to set \( w^{*L} \) as their reservation wage.
As long as members of group H expect acceptable offers to arrive with probability \( \lambda \eta \) starting in the second period of unemployment, their reservation wage is \( w^{*H} \).

This proposition gives conditions on the parameters that are sufficient for a stable equilibrium to exist in which the two types have different reservation wages while firms behave optimally. It does not say whether parameters satisfying these conditions can be found. What is needed is that (42) and (43) hold when the reservation wages are given by (41) and (40) respectively. Inspection of these inequalities suggests that they are likely to hold if both \( \lambda \) and \( \delta_r \) are high while \( \delta_s \) is low. The reason is that a high \( \lambda \) implies that (42) holds regardless of the values of \( w^{*i} \). A high value of \( \lambda \) also ensures that \( w^{*L} \) in (41) is low when \( \delta_s \) is low. Lastly, because \( D(1 - \gamma) \) is small relative to \( (D\gamma + \rho \lambda \eta) \), a high value of \( \delta_r \) leads \( w^{*H} \) to be fairly high as well. The result is that \( (R - w^{*H})/(R - w^{*L}) \) is low so that (43) 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, \( \sigma, \rho, \) and \( \gamma \) were equal to .02, .9959, and .4 respectively, while \( R \) was normalized to equal 1. The parameters \( \lambda, \delta_s, \delta_r \) and \( \eta \) 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 belongs to the same group of naive learners while the second is slightly larger than the common reservation wage when everyone has accurate beliefs.

Because \( \delta_r > \delta_s, \eta \) is less than one and \( \lambda \) is between zero and one, the expression in (40) is smaller than the expression in (29), so that minimum wage for people of type \( H \) must exceed the minimum wage when everyone has accurate beliefs. One intuitive reason for this is that, because they wait until the second period, workers of type \( H \) are more likely to have competitive offers than do workers when they all have accurate beliefs. The fact that \( w^{*L} \) exceeds the wage when everyone is a naive group learner is also easy to rationalize. It comes about because some workers of type \( L \) are hired by ordinary ads, which carry high wages than those that prevail when everyone belongs to a single naive group. This raises the average wage of workers of type \( L \) and thereby raises the minimum wage they accept.
The overall average wage in this economy $\bar{w}$ is given by
\[
\bar{w} = \left( [1 - (1 - \lambda)\eta]\bar{w}_s + (1 - \lambda)\eta \bar{w}_r \right) n^L + \bar{w}_r (1 - n^L).
\]
(44)

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 (21), the distribution of wages received by workers who obtain them from postings of type $i$ is given by (24). Because the fraction of workers that obtains jobs from offers of type $s$ 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
\[
x = n^L[1 - (1 - \lambda)\eta]\frac{(1 - \delta_s)^2}{4\delta_s}\left[\left( \frac{R - w^{*L}}{R - w} \right)^2 - 1 \right],
\]
(45)

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.

For ease of comparison with the formulas in previous subsections Proposition 8 is based on the assumption that members of $H$ know that the offers they find acceptable are stationary only after period 1. Suppose instead that both groups believe that acceptable offers are stationary with a per-period probability of arriving equal to the inverse of the group’s average unemployment duration. There is then an equilibrium where everyone has a reservation wage of $w^{*b}$. But, there can also be an equilibrium in which group $H$ accepts only ordinary offers while group $L$ also accepts prominent ones. The average unemployment duration for members of $H$ is $(1 + 1/\lambda\eta)$ so they expect acceptable offers to arrive with probability $\lambda\eta/(1 + \lambda\eta)$. Equations (40) and (41) in Proposition 8 thus need to be modified so that
\[
\frac{w^{*H}}{R} = \frac{\delta_r(D\gamma(1 + \lambda\eta) + \rho\lambda\eta)}{D(1 - \gamma)(1 + \lambda\eta) + \delta_r(D\gamma(1 + \lambda\eta) + \rho\lambda\eta)},
\]
(46)

\[
\frac{w^{*L}}{R} = \frac{(D\gamma + \rho\lambda)[(1 - (1 - \lambda)\eta)\delta_s + \frac{(1 - \lambda)\eta\delta_r(D\gamma(1 + \lambda\eta) + \rho\lambda\eta)}{D(1 - \gamma)(1 + \lambda\eta) + \delta_r(D\gamma(1 + \lambda\eta) + \rho\lambda\eta)}]}{D(1 - \gamma) + (D\gamma + \rho\lambda)((1 - (1 - \lambda)\eta)\delta_s + (1 - \lambda)\eta)}.
\]
(47)
With this modification, a proof essentially identical to that of Proposition 8 demonstrates that, as long as (42) and (43) are met, there is an equilibrium in which ordinary ads post wages above $w^H$, prominent ads post wages strictly lower than $w^H$ but above $w^L$, and the two groups have reservation wages given by (46) and (47). For the parameters used above, these wages equal .95 and .52 respectively. Members of group $H$ now accept slightly lower wages than before because they think desirable jobs arrive more slowly. This effect is small enough that the minimum wage for members of $L$ is not materially affected.

3 Tests of Stationarity and Accurate Beliefs

According to the theory presented here, wage dispersion is enhanced because some people’s information leads them to have reservation wages that are too low. If these people raised their reservation wage, they would be surprised at how quickly they would receive acceptable offers. This naturally leads one to want to compare the average wage gain from having a slightly higher reservation wage to the corresponding average increase in the time it takes to find a job. If the former is 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. This comparison is possible only if people with identical job opportunities differ in their reservation wages. There is considerable evidence that such differences in reservation wages exist.\textsuperscript{10}

Consider an individual with have a constant probability $\lambda$ 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 $S$ if the individual has 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 (4). Differentiating these equations, the effect of changes in $w^*$ on the percentage changes

\textsuperscript{10}For evidence of low $R^2$ in regressions explaining reservation wages with observables, see Feldstein and Poterba (1984) and Krueger and Mueller (2014).
in $S$ and $\bar{w}$ is

$$
\frac{dS}{S} = \frac{g(w^*)}{1-G(w^*)} dw^* \quad \frac{d\bar{w}}{\bar{w}} = \frac{g(w^*)}{1-G(w^*)} \frac{\bar{w} - w^*}{\bar{w}} dw^*.
$$

(48)

Define $\alpha$ by\(^{11}\)

$$
\alpha \equiv \frac{d\bar{w}}{dS/S}.
$$

The ratio $\alpha$ is akin to a “return” for the extra time spent looking for a job. It can be computed from samples of individuals with the same job market opportunities and different reservation wages. Those that set higher reservation wages turn down more jobs so their unemployment duration is longer, and $\alpha$ measures the gain in wages they obtain by doing this. One benefit of computing this statistic is that it can be compared to the prediction that the stationarity assumption is correct. When it is, (48) implies that

$$
\alpha = \frac{\bar{w} - w^*}{\bar{w}},
$$

(49)

so that the distance between measures of $\alpha$ and measures of $\bar{w}/w^*$ allows one to detect departures from stationarity.

To carry out this comparison, one can use the estimates of Holzer (1986) and DellaVigna and Paserman (2005)\(^{12}\). 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.\(^{13}\) The ratio $\alpha$ is thus equal to the ratio of the primary coefficients obtained in these two regressions. When Holzer (1986) uses weighted least squares (based on sample weights), the resulting ratios are either 2.33 or 1.85 for whites depending on controls. For

---

\(^{11}\)Derivatives of the form of (48) can be found in Holzer (1986), but the ratio in (3) appears to be new to the literature.

\(^{12}\)Krueger and Mueller (2011, 2014) 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 $\alpha$. Krueger and Mueller (2014) argue that typical reservation wages are too large relative to a benchmark model with stationary wage offers, thereby showing that many workers do not act in accordance with this model. One possibility raised by the current paper is that these individuals realize that offers are not, in fact, stationary.

\(^{13}\)Holzer’s (1986) uses the same controls in both regressions. They are 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.
blacks, they are either .32 or .18. While not using exactly the same right hand side variables in both regressions, DellaVigna and Paserman (2005) report regressions explaining the log wage and the log hazard with the log reservation wage. Treating the duration as the inverse of the hazard, their estimates imply an $\alpha$ of 4.8.

Meanwhile, 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 $(\bar{w} - w^*)/\bar{w}$ are not out of line with the $\alpha$’s for blacks in Holzer (1986) but they are much lower than both Holzer’s (1986) $\alpha$’s for whites and the estimate based on DellaVigna and Paserman (2005). Indeed, the latter two values of $\alpha$ exceed one, and the right hand side of (49) 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” $\alpha$ 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. HKV show, for example, that on-the-job search can raise the ratio of mean to minimum wages. The more it does so, however, the stronger is the evidence against the stationarity of offers. The reason is that on-the-job search has no effect on the derivations in (48), except that $\bar{w}$ is the average wage obtained by people who search while unemployed. The lower is this $\bar{w}$ relative to the observed economy wide mean, the lower the implied $\alpha$ that is consistent with stationarity.

It might be thought that a simpler check on the model’s implications is to ask whether, controlling for individual characteristics, wages rise unusually fast with the duration of unemployment. Such a test would have the advantage of not requiring data on reservation wages. An obvious defect of this simpler test, however, is that some people take long time to find a (low-paying) job because they appear unattractive to potential recruiters. Outliers of this sort have a big effect on the estimated relation between duration and wage growth but are likely to have a much smaller effect on estimates of $\alpha$.

If one is armed not only with people’s reservation wages but also with their own estimates

---

14\textsuperscript{They include past wages in the reservation wage regression but not in the duration regression.}
of how much they expect to earn when they get a job, one can use (49) to see 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 $\alpha$ should equal .13, which remains smaller than any of the estimates based on Holzer (1986) and DellaVigna and Paserman (2005). While the samples in these studies are not the same, they broadly suggest that the benefits of increasing the time spent searching might well be larger than those expected by people who search for a job.

One can calculate a variant of $\alpha$ 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 $\bar{w}_s = w^{sL} + \delta_s(R - w^{sL})$ while the average wage of those that obtained their job through an ordinary ad is $\bar{w}_r = w^{sH} + \delta_r(R - w^{sH})$. The average wage earned by workers of type $H$ is simply $\bar{w}_r$ while that of people of type $L$ equals $\bar{w}_s + (1 - \lambda)\eta(\bar{w}_r - \bar{w}_s)$. 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 $\alpha$ 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. While lower than most empirical estimates reported above, these parameters do yield values of $\alpha$ that exceed the right hand side of (49). As discussed below equation (44), these
parameters imply that $\bar{w}/w^*$ equals 1.55. The right hand side of (49) thus equals .35, which is indeed smaller than $\alpha$.

4 Conclusions

The assumption that workers are correct in their belief that offers are stationary is ubiquitous in the literature. This assumption has not been seriously examined, however, and some of the existing evidence contradicts it by suggesting that some people would experience substantial gains by raising their reservation wage. This paper shows that this form of nonstationarity makes it easy for different groups to end up with different beliefs about the wages of available jobs, thereby fostering inter-group inequality in earnings.

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 can fit the wage distribution better.

The model implies that some people would benefit from staying unemployed longer. This need not 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 to lose marketability as a result. 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 members of groups that earn relatively little to quit more 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, as mentioned in the introduction, several empirical papers on job referrals suggest that jobs obtained through peer contacts sometimes carry low wages. 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.

Finally, the model has ignored on-the-job search even though there are many job-to-job transitions. The existence of such transitions need not, of course, eliminate either the role of beliefs or the transmission of information by groups stressed here. On the contrary, it would seem that on-the-job search has the potential for giving groups new roles. For example, people who are currently working have less time to look for new jobs than people who are unemployed, so that searching is more costly for them. One obvious question, then, is whether it is possible for equilibrium group experiences to differ so that groups differ in the advice they give their members regarding the desirability making the effort of searching on-the-job.
References


Feldstein, Martin and James Poterba, “Unemployment insurance and reservation wages,”


Montgomery, James D., “Social Networks and Labor-Market Outcomes: Toward an Eco-


Appendix: Proofs of Propositions

Proof of Proposition 1: Suppose a group has a reservation wage \( w^* \). Because wage offers are stationary, \( \bar{w} = E_G(w|w \geq w^*) \) and the duration of unemployment is such that its inverse \( \lambda \) equals \( \lambda(1 - G(w^*)) \). This means that, at a stable group learning outcome, \( U \) and \( V(w) \) represent the actual present discounted values of being unemployed and employed at a wage \( w \) respectively.

Now turn to the rational expectations optimum reservation wage \( w_o^* \). Optimality requires the true present value of having a job of \( w_o^* \) be equal to the true value of being unemployed when \( w_o^* \) is the reservation wage. Since the former equals \( V(w_o^*) \) ad the latter equals the value of \( U \) when the reservation wage is \( w_o^* \), it follows that the rational expectations optimum reservation wage \( w_o^* \) must also be a stable group learning outcome.

I now turn to the question of whether this is the unique stable group learning outcome. Since \( U \) must be accurate for any reservation wage \( w^* \), it must equal

\[
U = b + \rho \left[ U + \lambda \int_{w^*} (V(w) - U) dG(w) \right].
\]

Therefore any reservation wage that is a stable group learning outcome must satisfy

\[
w^* = b + \frac{\rho \lambda}{1 - \rho(1 - \sigma)} \int_{w^*} (w - w^*) dG(w).
\]

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: If, under accurate beliefs, workers did not accept offers below \( w_2 \), their \( U \) starting in from period 1 is given by (2) with \( \hat{\lambda} \) replaced by \( \lambda \eta \) and \( \bar{w} \) replaced by \( w_2 \). Inequality is equivalent to having \( V(w_1) \) exceed this value of \( U \). Therefore workers do indeed accept offers of \( w_1 \) if and only if this condition holds.

If, under group learning, workers do not accept offers lower than \( w_2 \), \( \hat{\lambda} \) is below \( \lambda \eta \) because no one takes a job in period 1. This means that \( U \) is lower, implying that these workers accept offers of \( w_1 \) whenever workers with accurate beliefs do so.

Proof of Proposition 3: The left hand side of (7), \( U_1 \) and \( W_1 \) are all convex combinations of \( b \) and \( w_2 \). The weight on \( b \) is unambiguous larger in \( U_2 \) than in \( W_1 \). Also, the weight of \( b \) is larger in the left hand term than in \( U_2 \) if \( \lambda \) and \( \eta \) are smaller than 1 because this implies that \( 1/(1 + \lambda \eta) > (1 - \lambda) \).

For there to exist a stable group learning outcome in which the reservation wage is \( w_1 \), it is necessary and sufficient that \( V(w_1) \geq U_1 \) where \( V(w_1) \) is computed assuming the value of being unemployed is \( U_1 \). Using the reasoning that leads to (3), this occurs when \( w_1 \geq (1 - \rho)U_1 \). This is equivalent to having \( w_1 \) exceed the left hand side of (7).

Even if all currently employed group members have accepted jobs that pay \( w_2 \), a newly unemployed worker will accept a job that pays \( w_1 \) if \( V(w_1) > U_2 \) when \( U_2 \) is used as the value of being unemployed in computing \( V(w_1) \). This occurs whenever \( w_1 \) exceeds \( (1 - \rho)U_2 \).

Proof of Proposition 4: Equations (22) and (25) imply that

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

Using this in (26) and (27), and plugging the results in (28), (29) follows. Moreover, for \( \eta = 0 \) or \( \delta = \delta_r \), both \( \bar{w} \) and \( \bar{w}_2 \) equal \( \bar{w}_1 \) so that (30) holds as an equality. Otherwise \( \bar{w} \) is smaller than \( \bar{w}_2 \) so that it holds as a strict inequality.
Proof of Proposition 5:
Because the hazard of leaving unemployment remains the same for all workers, the $Q_i$ are still given by (18) and (19), and the expected profits of a firm that uses an ad of type $i$ to offer a wage $w$ are given by (20). Therefore (21) holds at a zero profit equilibrium, though the values of $\theta_i$ can be different. It follows that both (22) and (23) hold when $w^a$ is replaced by $w^b$. Since (24) still defines the cdf of the wages earned by people who obtain their job using an ad of type $i$, (25) still gives the value of the average wages $\bar{w}_i$ as long as $w^a$ is replaced by $w^b$. This implies

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

Using this in (27), which remains a property of the economy-wider average wage, and substituting into (32) gives (33).

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

$$Q_i^L (R - w^L) \over (1 + \delta_i) Dv_i (1 - \delta_i) - c_i.$$  

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

$$Q_i^H (R - w^H) \over (1 + \delta_i) Dv_i (1 - \delta_i) - c_i,$$

and they are higher still if $F(w^H) > 0$. Thus offering a wage of $w^H$ strictly dominates offering any lower wage when (37) holds. With free entry, the lowest wage offered is $w^H$, which implies that the distribution of wages is given by (21) 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 (37) is reversed, expected profits are strictly higher by offering $w^L$ rather than $w^H$.

Using the definitions of $Q^j_i$ given in (35), (36) and the discussion below

$$Q_s^H = 1 + \frac{u^H}{u^H} + (1 - \eta)(u^H - u^H) \over u^H + (1 - \eta)(u^L - u^L),$$

$$Q_r^H = 1 + \frac{u^H}{u^H} - u^L \over u^L - u^L.$$

Since (34) implies that

$$\frac{u^H}{u^H} < \frac{u^H - u^H}{u^L - u^L},$$

it follows that $Q_s^H/Q_r^L < Q_r^H/Q_r^L$. Therefore (37) 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

$$Q_i^H (R - w^H) (1 - \delta_i + 2\delta_i F_i(w^H)) - c_i,$$

so they are bounded above by

$$Q_i^H (R - w^H) (1 + \delta_i) - c_i.$$  

(53)

The inequality in (38) implies that (37) is reversed so that ads of type $i$ include offers of $w^L$ and profits at this wage are given by the expression in (50). Therefore, these higher wages are not offered.

Free entry then ensures that the expression in (50) equals zero so that (21) 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 (37) and (38) are reversed. There must then exist an $0 < F_i(w^H) < 1$ such that the expression in (52) equals the expression in (50). Moreover, because (37) is reversed, there exists a value of $w^L < w^{mL} < w^H$ such that (39) 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 (50). Thus, for $w^L \leq w \leq w^{mL}$, 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 $w \geq w^H$.

**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 (42) and (43) 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 (20) to be zero when $Q_s$ is equated with $Q_i^L$ while $Q_r$ is equated with $Q_i^H$. These cdf’s must therefore satisfy (21) with these $Q$’s while, as in (22) the values of $c_i v_i$ ensure that

$$(R - w^L)(1 - \delta_s) = \theta_s \quad (R - w^H)(1 - \delta_r) = \theta_r.$$
It follows that the people who accept jobs from ads of type $i$ have a cdf of wages $G_i(w)$ given by (24) so that their average wage $\bar{w}_i$ satisfies (25). Together with the equation above, this implies that

$$\bar{w}_s = \delta_s R + (1 - \delta_s)w^{*L} \quad \bar{w}_r = \delta_r R + (1 - \delta_r)w^{*H}$$ (54)

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 (3) when $b = \gamma w^{*H}$. Using (54), this requires that

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

which is equivalent to (40). When combined with (54) this implies that

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

Now turn to individuals of type $L$. Suppose that they accept all viable offers above $w^{*L}$, so that their estimated and actual hazard of leaving unemployment is $\lambda$ while their average wage $\bar{w}^L$ is given by the expression in (27). 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 (3) with $b = \gamma \bar{w}^L$. Noting that $\bar{w}^L = \bar{w}_s + (1 - \lambda)\eta(\bar{w}_r - \bar{w}_s)$ and using (54) for $\bar{w}_s$, this requires that

$$(D + \rho \lambda \eta)w^{*L} = (D \gamma + \rho \lambda \eta)\left((1 - \eta(1 - \lambda))(\delta_s R + (1 - \delta_s)w^{*L}) + \eta(1 - \lambda)\bar{w}_r\right).$$

Using the solution for $\bar{w}_r$ above, this is equivalent to (41).

With this worker behavior the total number of employees and unemployed individuals of types $L$ and $H$ are given by (9) and (10) respectively. Given that the number of people of type $i$ who become unemployed in the current period $u'_i$ is given by $\sigma m^i$, we have

$$\frac{u^H - u^H_f}{u^L - u^L_f} = \frac{(\sigma + \lambda)(1 - n^L)}{(1 - \lambda)(\sigma + \lambda \eta(1 + \sigma))n^L} \quad \frac{u^H_f + (1 - \eta)(u^H - u^H_f)}{u^L_f + (1 - \eta)(u^L - u^L_f)} = \frac{(\sigma + \lambda)(1 - n^L)}{(\sigma + \lambda \eta(1 + \sigma))n^L}$$

As a result, (42) ensures that (37) holds for $i = 2$ while (43) ensures that (38) holds for $i = 1$. 

43