Equilibrium Directed Search with Multiple Applications*

James Albrecht
Department of Economics
Georgetown University
Washington, D.C. 20057

Pieter A. Gautier Free University Amsterdam Tinbergen Institute

Susan Vroman
Department of Economics
Georgetown University
Washington, D.C. 20057

January 30, 2006

1 Introduction

In this paper, we construct an equilibrium model of directed search in a large labor market in which unemployed workers make multiple job applications. Specifically, we consider a matching process in which job seekers, observing the wages posted at all vacancies, send their applications to the vacancies that they find most attractive. At the same time, each vacancy, when it chooses its wage posting, takes into account that its posted wage influences the number of applicants it can expect to attract. We assume that each unemployed worker makes a fixed number of applications, a. Each vacancy (among those receiving applications) then chooses one applicant to whom it offers its job. When a > 1, there is a possibility that more than one vacancy wants to hire the same worker. In this case, we assume that the vacancies in question can compete for this worker's services. The introduction of multiple applications adds realism to the directed search model, and, in addition, affects the efficiency properties of equilibrium. In the benchmark competitive

^{*}We thank the anonymous referees and our editor, Fabrizio Zilibotti, for detailed and helpful comments. We also thank Ken Burdett, Behzad Diba, Benoit Julien, Vladimir Karamychev, Ian King, Harald Lang, Dale Mortensen, Lucas Navarro, and Rob Shimer.

search equilibrium model (Moen 1997), equilibrium is constrained efficient. We show that changing the basic directed search model to allow workers to make more than one application results in equilibria that are not constrained efficient. This means there is a role for labor market policy in the directed search framework.

When a = 1, our model is essentially the limiting version of Burdett, Shi, and Wright (2001) (hereafter BSW) translated to a labor market setting. BSW derive a unique symmetric equilibrium in which (in the labor market version) all vacancies post a wage between zero (the monopsony wage) and one (the competitive wage). The value of this common posted wage depends on the number of unemployed, u, and the number of vacancies, v, in the market. Letting $u, v \to \infty$ with $v/u = \theta$, the equilibrium posted wage is an increasing function of θ . BSW do not consider normative questions. Moen's result is that in a large labor market, directed search implements what he calls competitive search equilibrium. Competitive search equilibrium is constrained efficient in the following sense. Assume there is a cost per vacancy created. A social planner would choose a level of vacancy creation - or, in a large labor market, a level of labor market tightness, θ , - to trade off the cost of vacancy creation against the benefit of making it easier for workers to match. Moen shows that the θ the social planner would choose is the same as the one that arises in competitive search equilibrium. Using a different approach, we also show that equilibrium in a directed search model is constrained efficient in a large labor market when a = 1. More importantly, however, we show that if each worker makes a finite number of multiple applications, that is, if $a \in \{2, ..., A\}$, where A is any arbitrary, finite integer, then equilibrium in a directed search model is **not** constrained efficient. Specifically, too many vacancies are posted (θ is too high) in freeentry equilibrium relative to the constrained efficient level. Equivalently, vacancies pay the workers who take their jobs too low a wage on average.

Our model is also related to Julien, Kennes, and King (2000) (hereafter JKK). JKK assume that each unemployed worker posts a minimum wage at which he or she is willing to work, i.e., a "reserve wage," and that each vacancy, observing all posted reserve wages, then makes an offer to one worker. If more than one vacancy wants to hire the same worker, then, as in our model, there is $ex\ post$ competition for that worker's services. This is equivalent to a model in which each worker applies to every vacancy, i.e., a=v, sending the same reserve wage in each application. Each vacancy then chooses one worker at random to whom it offers a job. If a worker has more than one offer, then there is competition for his or her services. In a finite labor market, JKK show that the unique, symmetric equilibrium

reserve wage lies between the monopsony and competitive levels. There is thus equilibrium wage dispersion in their model. Those workers who receive only one offer are employed at the reserve wage, while those who receive multiple offers are employed at the competitive wage. In the limiting labor market version of JKK, the symmetric equilibrium reserve wage converges to zero, and free-entry equilibrium is again constrained efficient.

In our model, when $a \in \{2, ..., A\}$, all vacancies post the monopsony wage in the unique symmetric equilibrium. As in JKK, this leads to equilibrium wage dispersion. Some workers (those who receive exactly one offer) are employed at the monopsony wage, and some workers (those who receive multiple offers) have their wages bid up to the competitive level. The key difference between our model and both BSW and JKK, however, is that free-entry equilibrium is inefficient. When $a \in \{2, ..., A\}$, there is excessive vacancy creation.

The inefficiency arises because when $a \in \{2, ..., A\}$, two coordination frictions operate simultaneously. The first is the well-known urn-ball friction; some vacancies receive no applications while others receive more than one. In addition, a new friction is introduced by multiple applications. Some workers receive multiple offers while others receive none. As a result, some vacancies with applicants fail to hire a worker. In BSW, only the urn-ball friction is present; in JKK, only the multiple-application friction applies. The market cannot correct both frictions at the same time. In our model, competition among vacancies, once applications have been made, can solve the multiple-application friction. This leads, however, to a posted wage that is too low to correct the urn-ball friction and that consequently generates too many vacancies.

The outline of the rest of the paper is as follows. In the next section, we derive our basic positive results in a single-period framework. Specifically, treating θ as given, we derive the matching function and the symmetric equilibrium posted wage. In Section 3, we endogenize θ by allowing for free entry of vacancies. This lets us compare the free-entry equilibrium level of θ to the constrained efficient level (the two values of θ are the same when a=1, different when $a\in\{2,...,A\}$, and the same once again as $a\to\infty$). In Section 4, we present a steady-state version of our model for the case of $a\in\{2,...,A\}$. The key to the steady-state analysis is that a worker who receives only one offer in the current period has the option to reject that offer in favor of waiting for a future period in which more than one vacancy bids for his or her services. This leads to a tractable model in which labor market tightness and the equilibrium wage distribution are determined simultaneously. The normative results that we derived in the single-period

model continue to hold in the steady-state setting. In Section 5, we consider three extensions. Specifically, (i) we allow workers to choose how many applications to make, (ii) we relax the assumption that each vacancy can consider only one worker's application, and (iii) we allow vacancies to follow strategies that rule out Bertrand competition. These extensions, while of interest in their own right, also serve as robustness checks – our basic result that the free-entry equilibrium value of θ is constrained inefficient when $a \in \{2, ..., A\}$ continues to hold. Finally, we conclude in Section 6.

2 The Basic Model

We consider a game played by u homogeneous unemployed workers and (the owners of) v homogeneous vacancies. This game has several stages:

- 1. Each vacancy posts a wage.
- 2. Each unemployed worker observes all posted wages and then submits a applications with no more than one application going to any one vacancy.
- 3. Each vacancy that receives at least 1 application randomly selects one to process. Any excess applications are returned as rejections.
- 4. A vacancy with a processed application offers the applicant the posted wage. If more than one vacancy makes an offer to a particular worker, then each vacancy can increase its bid for that worker's services.
- 5. A worker with one offer can accept or reject that offer. A worker with more than one offer can accept one of the offers or reject all of them.

Workers who fail to match with a vacancy and vacancies that fail to match with a worker receive payoffs of zero. The payoff for a worker who matches with a vacancy is w, where w is the wage that he or she is paid. A vacancy that hires a worker at a wage of w receives a payoff of 1-w. This is a model of directed search in the sense that workers observe all wage postings and direct their applications to vacancies with attractive wages and/or where relatively little competition is expected. We assume that vacancies cannot pay less than their posted wages.

Before we analyze this game, some comments on the underlying assumptions are in order. First, we are treating a as a parameter of the search technology; that is, the number of applications is taken as given. In general,

 $a \in \{1, 2, ..., A\}$. Second, we assume that it takes a period for a vacancy to process an application. This is why vacancies return excess applications as rejections. This processing-time assumption captures the idea that when workers apply for several jobs at the same time, firms can waste time and effort pursuing applicants who ultimately go elsewhere. Finally, we assume that a vacancy that faces competition for its selected applicant always has the option to increase its offer. This means that workers who receive more than one offer have their wages bid up via Bertrand competition to w = 1, the competitive wage. In Section 5, we consider the implications of relaxing each of these assumptions. We show that endogenizing a, allowing vacancies to process more than one application, and allowing vacancies that are competing for an applicant to pursue a different tie-breaking strategy do not reverse our main results.

We consider symmetric equilibria in which all vacancies post the same wage and all workers use the same mixed strategy to direct their applications. We do not consider equilibria in which workers follow asymmetric application strategies since this would require unrealistic implicit coordination. We do our analysis in a large labor market in which we let $u, v \to \infty$ with $v/u = \theta$ keeping $a \in \{1, 2, ..., A\}$ fixed. We show that for each (θ, a) combination there is a unique symmetric equilibrium, and we derive the corresponding equilibrium matching probability and posted wage. Assuming (for the moment) the existence of a symmetric equilibrium, we begin with the matching probability.

¹One might think of ruling out $ex\ post$ bidding by assumption, but then there would be no common equilibrium posted wage. To see this, suppose all vacancies post a wage of w. Then, assuming that a worker who has multiple offers accepts the highest one, it is in the interest of any vacancy to post a slightly higher wage so long as w is not too close to one. The reason is that if a vacancy posts a wage ε above the common wage, its probability of hiring a worker jumps discontinuously since it "wins" whenver the worker has multiple ofers. Once w is sufficiently close to one, a vacancy can profit by lowering its wage to the minimum level consistent with attracting one or more applicants with some positive probability. This is similar to the argument given in Burdett and Judd (1983) for nonexistence of a single-price equilibrium.

²One could alternatively consider symmetric equilibria in which vacancies follow mixed strategies, so that more than one wage is posted in equilibrium. This approach is taken in Galenianos and Kircher (2005), which combines elements of our paper and that of Chade and Smith (2004). In Galenianos and Kircher (2005), a vacancy whose chosen applicant has other offer(s) is precluded by assumption from increasing its initial offer, even though it would be in its interest to do so, given that other vacancies have committed to not changing their offers. The assumption that vacancies cannot engage in *ex post* bidding is restrictive, but without it, equilibrium in Galenianos and Kircher (2005) would not be subgame perfect.

Let M(u,v;a) be the expected number of matches in a labor market with u unemployed workers and v vacancies when each unemployed worker submits a applications. Then $m(\theta;a) = \lim_{u,v \to \infty, v/u = \theta} \frac{M(u,v;a)}{u}$ is the matching probability for an unemployed worker in a large labor market.

Proposition 1 Let $u, v \to \infty$ with $v/u = \theta$ and $a \in \{1, ..., A\}$ fixed. The probability that a worker finds a job converges to

$$m(\theta; a) = 1 - (1 - \frac{\theta}{a} (1 - e^{-a/\theta}))^a. \tag{1}$$

The proof is given in Albrecht et. al. (2004); see also Philip (2003). In Appendix A, we sketch the idea of the proof to clarify the relationship between our matching probability and the finite-market matching functions presented in BSW (the standard urn-ball matching function) and JKK (the urn-ball matching function with the roles of u and v reversed).

For use below, we note the following properties of $m(\theta; a)$:

(i) $m(\theta; a)$ is increasing and concave in θ ,

$$\lim_{\theta \to 0} m(\theta; a) = 0, \text{ and } \lim_{\theta \to \infty} m(\theta; a) = 1;$$

(ii)
$$\frac{m(\theta; a)}{\theta}$$
 is decreasing in θ , ³

$$\lim_{\theta \to 0} \frac{m(\theta;a)}{\theta} = 1, \text{ and } \lim_{\theta \to \infty} \frac{m(\theta;a)}{\theta} = 0.$$

The effect of a on $m(\theta; a)$ is less clearcut. Treating a as a continuous variable, we find that $m_a(\theta; a) \geq 0$ as $\frac{a}{1-q} \frac{\partial q}{\partial a} - \ln(1-q) \geq 0$ where $q = \frac{\theta}{a}(1-e^{-a/\theta})$ is the probability that any one application leads to an offer. For moderately large values of θ ($\theta > \frac{1}{2}$, approximately), $m(\theta; a)$ first increases and then decreases with a. This nonmonotonicity reflects the double coordination problem that arises when workers apply to more than one but not all vacancies. The first coordination problem is the standard one associated with urn-ball matching, namely, that some vacancies can receive applications from more than one worker, while others receive none.

³Interestingly, $\frac{m(\theta;a)}{\theta}$ is not convex in θ , as can be seen immediately by considering the case of a=1. The properties of $m(\theta;a)$ and $\frac{m(\theta;a)}{\theta}$ given in (i) and (ii) are the minimal ones required for our normative results in Sections 3 and 4 below.

With multiple applications, there is a second coordination problem, this time among vacancies. When workers apply for more than one job at a time, some workers can receive offers from more than one vacancy, while others receive none. Ultimately, a worker can only take one job, and the vacancies that "lose the race" for a worker will have wasted time and effort while considering his or her application. The matching function derived in BSW captures only the urn-ball friction, while the one derived in JKK captures only the multiple-application friction. Our matching probability incorporates both these frictions, and the interaction between these two frictions provides new insights.

Proposition 1 and its implications are only interesting if a symmetric equilibrium exists. We now turn to the existence question.

Proposition 2 Consider a large labor market in which $u, v \to \infty$ with $v/u = \theta$. There is a unique symmetric equilibrium to the wage-posting game. When a = 1, all vacancies post a wage of

$$w(\theta;1) = \frac{\frac{1}{\theta}e^{-1/\theta}}{(1 - e^{-1/\theta})}.$$
 (2)

When $a \in \{2, ..., A\}$, all vacancies post a wage of $w(\theta; a) = 0$, and the fraction of wages paid that are equal to one is

$$\gamma(\theta; a) = \frac{1 - (1 - \frac{\theta}{a}(1 - e^{-a/\theta}))^a - \theta(1 - e^{-a/\theta})(1 - \frac{\theta}{a}(1 - e^{-a/\theta}))^{a-1}}{1 - (1 - \frac{\theta}{a}(1 - e^{-a/\theta}))^a}.$$
(3)

The proof is given in Appendix B. The basic idea is as follows. To prove the existence of a symmetric equilibrium, we show that $w(\theta;1)$ has the property that if all vacancies, with the possible exception of a "potential deviant," post that wage, then it is also in the interest of the deviant to post that wage. When $a \in \{2, ..., A\}$, however, no matter what the common wage posted by other vacancies, it is always in the interest of the deviant to undercut that common wage. This forces the wage down to the monopsony level, which in our single-period model is w = 0.

The equilibrium wage for the case of a=1 is equal to one minus the price given in Proposition 3 in BSW – again with the appropriate notational change. The tradeoff that leads to a well-behaved equilibrium wage, $w \in (0,1)$, when a=1 is the standard one in equilibrium search theory. To see this, note that the profit for a deviant (D) from offering w' rather than the common posted wage, w, can be written as:

 $\pi(w'; w) = (1-w')P[D \text{ gets at least one application}]P[\text{selected applicant has no other offer}],$

where the third term equals 1 when a=1. As any particular vacancy increases its posted wage, holding the wages posted at other vacancies constant, the profit that this vacancy generates conditional on attracting an applicant, (1-w'), decreases. At the same time, however, the probability that it attracts at least one applicant increases. This tradeoff varies smoothly with θ ; so the equilibrium wage varies smoothly between zero and one. Thus, as emphasized in BSW (p. 1069), there is a sense in which frictions "smooth" the operation of the labor market.

When $a \in \{2, ..., A\}$, the posted wage collapses to the monopsony level (as in Diamond (1971)). The intuition for this result is based on the change in the tradeoff underlying equilibrium wage determination. This change – to be described below – has two implications. First, the equilibrium wage is lower than when a = 1. Second, when $a \in \{2, ..., A\}$, the lower is the putative common equilibrium wage w, the stronger is the incentive to deviate by posting w' < w. This second implication is what drives the wage down to the monopsony level.

Why is the equilibrium wage lower when workers make more than one application? Note first that the incentive to deviate from a common posted wage w comes from the first two terms in $\pi(w'; w)$ since the third term is unaffected by changes in w' when the labor market is large. That is, the incentive to deviate comes from the effect of w' on 1-w' and on the probability that the deviant receives at least one application. The effect of offering w' on 1-w' is obviously the same whether workers make one or multiple applications. However, a deviation has less effect on the probability that the deviant gets at least one applicant when workers make multiple applications. Consider a deviation w' > w. The higher wage makes the vacancy more attractive to a worker if w' is the only offer received. However, when $a \in \{2, ..., A\}$, workers have an interest in getting multiple offers in order to generate Bertrand competition for their services, and since the deviant vacancy is more attractive to all workers, applying to the deviant decreases the probability that this occurs. Thus, a deviation w' > w increases the probability that the deviant receives at least one application by less when workers make multiple applications than when a = 1. Similarly, a deviation w' < w decreases the probability that the deviant receives at least one application by less than when a = 1. In this case, a worker who applies to the deviant gets a lower wage if this is the only offer received. Just as when a=1, this makes the deviant less attractive. However, when $a\in\{2,...,A\}$, to increase the chance of getting multiple offers, workers have an incentive to apply to a "safe" job where others are less likely to apply. Relative to the case of a = 1, this reduces the decrease in the probability that the deviant

attracts at least one applicant. The fact that upward deviations are less attractive and downward deviations are more attractive explains why the equilibrium wage is lower when $a \in \{2, ..., A\}$ than when a = 1.

Why does the equilibrium wage fall to the monopsony level when workers make multiple applications? The potential benefit to a deviant of posting a wage that is ε below a common wage w is the same for all w, but the cost in terms of the reduction in the probability of receiving at least one application falls as the common wage falls. The probability that the deviant receives at least one application depends on the chance that workers are willing to take to try to generate multiple offers. The closer w is to zero, the greater is the benefit to a worker of receiving multiple offers; i.e., the greater is the difference between the competitive wage and w. Thus, the incentive for workers to apply to a vacancy offering ε below the common wage, w, rises as w falls, and the probability that a vacancy offering ε less than w receives at least one application rises as w falls. Thus, as w falls, the potential benefit of a downward deviation is constant, but the cost of such a deviation decreases. This is what drives the common wage to the monopsony level.

Interestingly, when $a \in \{2, ..., A\}$, the equilibrium outcome in our directed search model is the same as the outcome one would find in a random search model in which workers make multiple applications and vacancies engage in Bertrand competition when their candidates have multiple offers. If workers do not observe posted wages, they apply at random to a vacancies in symmetric equilibrium, and the matching rate is the same as in our model. In addition, vacancies pay the monopsony wage in this random search model, unless a worker has multiple offers, in which case Bertrand competition drives the wage to the competitive level. Thus, allowing for multiple applications in our model erases the difference between directed and random search in terms of outcomes in contrast to the case of a = 1. To the best of our knowledge, no random search model with multiple applications and Bertrand competition exists in the literature, but it would be straightforward to construct such a model. Postel-Vinay and Robin (2002) is the most closely related model. In their model, wage offers arrive at Poisson rates to both the unemployed and the employed. If a worker who is already employed receives another offer, then that worker's current employer and prospective new employer engage in Bertrand competition for his or her services. In the homogeneous worker/homogeneous firm version of their model, this leads to a two-point distribution of wages paid, namely, the monopsony wage and the competitive wage, as in our model.

Finally, despite the fact that the posted equilibrium wage in our model is zero when $a \in \{2, ..., A\}$, there is still a sense in which "the wage" varies smoothly with θ . The expected fraction of wages paid that are equal to one, $\gamma(\theta; a)$, has the following properties:

- (i) $\gamma(\theta; a)$ is increasing in θ and in a;
- (ii) $\lim_{\theta \to 0} \gamma(\theta; a) = 0$ and $\lim_{\theta \to \infty} \gamma(\theta; a) = 1$.

The fact that γ is increasing in θ is exactly as one would expect – as the labor market gets tighter, the chance that an individual worker gets multiple offers increases. To understand why γ is also increasing in a, it is important to remember that $\gamma(\theta;a)$ is the expected wage for those workers who match with a vacancy; in particular, those workers who fail to match are not treated as receiving a wage of zero. Finally, defining $\gamma(\theta) = \lim_{a \to \infty} \gamma(\theta;a)$, we can show

$$\gamma(\theta) = \frac{1 - e^{-\theta} - \theta e^{-\theta}}{1 - e^{-\theta}}.$$
 (4)

This is the expected wage in a large labor market when each worker sends out an arbitrarily large number of applications.

3 Efficiency

We now turn to the question of constrained efficiency. The result suggested by the efficiency of competitive search equilibrium holds in our setting when a=1; however, when workers make a fixed number of multiple applications, this result breaks down.

Suppose vacancies are set up at the beginning of the period and that each vacancy is created at cost c_v . The efficient level of labor market tightness⁴ is determined as the solution to

$$\max_{\theta>0} m(\theta; a) - c_v \theta.$$

The first-order condition for this maximization is

$$c_v = m_\theta(\theta^*; a). \tag{5}$$

⁴In a finite labor market with u given, the social planner chooses v to maximize M(u,v;a)-cv, i.e., expected output (equal to the expected number of matches since each match produces an output of 1) minus the vacancy creation costs. Dividing the maximand by u and letting $u,v\to\infty$ with $v/u=\theta$ gives the maximand in the text.

The equilibrium level of labor market tightness is determined by free entry. When a = 1, this means

$$c_v = \frac{m(\theta^{**}; 1)}{\theta^{**}} (1 - w(\theta^{**}; 1)), \tag{6}$$

whereas for $a \in \{2, ..., A\}$, the condition is

$$c_v = \frac{m(\theta^{**}; a)}{\theta^{**}} (1 - \gamma(\theta^{**}; a)). \tag{7}$$

Equations (6) and (7) reflect the condition that entry (vacancy creation) occurs up to the point that the cost of vacancy creation is just offset by the value of owning a vacancy. This value equals the probability of hiring a worker times the expected surplus generated by a hire – equal to 1 minus the posted wage when a = 1 and to 1 minus the expected wage when $a \in \{2, ..., A\}$.

Note that θ^* denotes the constrained efficient level of labor market tightness and θ^{**} denotes the equilibrium level of labor market tightness. At issue is the relationship between θ^* and θ^{**} .

Proposition 3 Let $u, v \to \infty$ with $v/u = \theta$ and $a \in \{1, ..., A\}$ fixed. For $a = 1, \theta^* = \theta^{**}$. For $a \in \{2, ..., A\}, \theta^{**} > \theta^*$.

Proof. Differentiating equation (1) with respect to θ gives

$$m_{\theta}(\theta; a) = (1 - \frac{\theta}{a} (1 - e^{-a/\theta}))^{a-1} (1 - e^{-a/\theta} - \frac{a}{\theta} e^{-a/\theta}). \tag{8}$$

For the case of a = 1, substituting this into equation (5) gives an implicit expression for θ^* ,

$$c_v = 1 - e^{-1/\theta^*} - \frac{1}{\theta^*} e^{-1/\theta^*}.$$

Using equations (1) and (2) in equation (6) gives an implicit expression for θ^{**} ,

$$\frac{m(\theta^{**};1)}{\theta^{**}}(1-w(\theta^{**};1)) = 1 - e^{-1/\theta^{**}} - \frac{1}{\theta^{**}}e^{-1/\theta^{**}}.$$

Thus, equations (5) and (6) imply $\theta^* = \theta^{**}$ when a = 1.

When $a \in \{2, ..., A\}$, substituting equation (8) into equation (5) implies that θ^* solves

$$c_v = (1 - \frac{\theta^*}{a} (1 - e^{-a/\theta^*}))^{a-1} (1 - e^{-a/\theta^*} - \frac{a}{\theta^*} e^{-a/\theta^*}), \tag{9}$$

whereas, using equations (1) and (3), θ^{**} (equation (7)) solves

$$c_v = \left(1 - \frac{\theta^{**}}{a} (1 - e^{-a/\theta^{**}})\right)^{a-1} (1 - e^{-a/\theta^{**}}). \tag{10}$$

The right-hand sides of both (9) and (10) are decreasing in θ . Since the right-hand side of (10) is greater than that of (9) for all $\theta > 0$, it follows that $\theta^{**} > \theta^*$.

Posting a vacancy has the standard congestion and thick-market effects in our model – adding one more vacancy makes it more difficult for the incumbent vacancies to find workers but makes it easier for the unemployed to generate offers. A striking result of the competitive search equilibrium literature is that adding one more vacancy causes the wage to adjust in such a way as to balance these external effects correctly. One way to interpret this result is that competition leads to a wage that satisfies the Hosios (1990) condition in a Nash bargaining model. Equivalently, one can say (Moen, 1997, p. 387) that the competitive search equilibrium wage has the property that the marginal rate of substitution between labor market tightness and the wage is the same for vacancies as for workers. The first part of Proposition 3 shows that this result holds when one uses an explicit urn-ball (a = 1) microfoundation for the matching function. When workers make multiple applications, however, the result that $\theta^{**} > \theta^*$ indicates that the equilibrium level of vacancy creation is too high. Equivalently, the equilibrium expected wage is below the level that would be indicated by the Hosios condition.

A first intuition for why we find constrained efficiency with a=1 but not with a fixed, finite number of multiple applications is that with a=1, only one coordination problem affects the operation of the labor market, whereas with a fixed $a \in \{2, ..., A\}$, the urn-ball and the multiple-applications coordination problems operate simultaneously. Adding a vacancy increases the number of matches by reducing the first coordination friction, the one that workers impose on each other, but at the same time increases the second coordination friction, the one that vacancies impose on each other. When each worker applies to only one vacancy, the second friction is absent, but with multiple applications there are two coordination problems that cannot be solved simultaneously. This intuition does not, however, address the question of why there is too much vacancy creation, as opposed to not enough. Accordingly, we now give a more detailed explanation of our inefficiency result.

The social planner opens vacancies as long as the marginal social benefit exceeds c_v , while the market opens vacancies as long as the marginal (=

average) private benefit exceeds c_v . When a=1, the private benefit of a new vacancy equals the social benefit. When $a \in \{2, ..., A\}$, the private benefit exceeds the social benefit. The social benefit of a new vacancy is simply $m_{\theta}(\theta; a)$; the private benefit is $\frac{m(\theta; a)}{\theta}(1 - \gamma(\theta; a))$. The key to understanding the discrepancy between the private and social benefits of a new vacancy when $a \in \{2, ..., A\}$ is to note that $m_{\theta}(\theta; a)$ can be expressed as

$$m_{\theta}(\theta; a) = \frac{m(\theta; a)}{\theta} (1 - \gamma(\theta; a)) p(\theta; a). \tag{11}$$

where

- $m(\theta; a)/\theta$ is the probability that a vacancy receives at least one application
- $1 \gamma(\theta; a)$ is the probability that the worker who has been offered the job has no other offers
- $p(\theta; a) = \frac{1 e^{-a/\theta} \frac{a}{\theta}e^{-a/\theta}}{1 e^{-a/\theta}}$ is the probability that a vacancy receives two or more applications conditional on receiving at least one.

One can derive equation (11) by differentiating $m(\theta; a)$, but this expression can also be derived using a straightforward economic argument. A vacancy has value to the social planner if it leads to an otherwise idle worker being employed and producing output. With probability $\frac{m(\theta; a)}{\theta}$, a vacancy is filled and produces one unit of output. However, with probability $\gamma(\theta; a)$, the worker who matches with the vacancy also receives another offer. In this case, the social benefit of the vacancy is zero; if it had not been opened, the number of matches would have been the same. The private benefit in this case is also zero – a vacancy receives nothing if it makes its offer to a worker who has other offer(s) since the wage is bid up to the competitive level. Social and private incentives are thus aligned with respect to the first two terms on the right-hand side of equation (11). The social planner considers an additional factor, which is given by $p(\theta;a)$. Consider the creation of a new vacancy. If $p(\theta; a) = 0$, then all vacancies have at most one applicant. Opening a new vacancy creates no social benefit because any applicant that it might attract would leave another vacancy unfilled. On the other hand, if $p(\theta;a)=1$, a new vacancy, if it is filled, does not leave another vacancy with no applicants. In general, the lower is $p(\theta; a)$, the more likely it is that a new

vacancy will cause an incumbent vacancy to fail to attract any applicants and hence the lower the social benefit.

To further understand our inefficiency result, we ask what wage, w^* , should have been posted in the first stage of the game in order to achieve efficiency? In other words, if the social planner could only determine the wage and not θ , what wage would she post? The efficient wage has to satisfy

$$m_{\theta}(\theta; a) = \frac{m(\theta; a)}{\theta} (1 - \gamma(\theta; a))(1 - w^*);$$

that is,

$$w^*(\theta; a) = 1 - p(\theta; a) = \frac{\frac{a}{\theta}e^{-(a/\theta)}}{1 - e^{-(a/\theta)}}.$$

When a = 1, w^* equals the posted wage given in equation (2) in Proposition 2. The fact that the posted wage is zero when $a \in \{2, ..., A\}$ is what leads to an inefficient outcome. The inefficiency problem when workers make multiple applications could thus be solved by an appropriately chosen minimum wage.

According to the Hosios condition, efficiency requires that the expected private benefit of opening a vacancy equals the marginal contribution of that vacancy to the matching process and that the expected wage equals the worker's marginal contribution to the matching process. The efficient wage w^* equals the probability that a vacancy receives exactly one application conditional on receiving at least one. This conditional probability is the marginal contribution of a worker to the matching process because output is only increased if the worker applies to a vacancy with no other applicants. When workers apply to more than one vacancy and there is $ex\ post$ Bertrand competition among vacancies, workers apply to vacancies even if they post a zero wage, and vacancies receive more surplus than their contribution to the matching process warrants. This is why there is excessive vacancy creation in equilibrium.

It is interesting to note that the equilibrium outcome is again Pareto efficient when we let $a \to \infty$. To see this, note that

$$m(\theta) = \lim_{a \to \infty} m(\theta; a) = 1 - e^{-\theta}$$

and

$$\gamma(\theta) = \lim_{a \to \infty} \gamma(\theta; a) = \frac{1 - e^{-\theta} - \theta e^{-\theta}}{1 - e^{-\theta}}$$

and substitute these into the efficiency and equilibrium conditions as in the proof of Proposition 3. Alternatively, following the route suggested by equation (11), note that as $a \to \infty$, $p(\theta; a) \to 1$, thus aligning the social and private benefits of vacancy creation.⁵ This result is Proposition 2.5 in JKK.

In a companion paper, Julien, Kennes, and King (2006) show that equilibrium in a finite labor market with a=v is also constrained efficient if one assumes a particular wage determination mechanism; namely, vacancies offering jobs to workers who have no other offers receive all of the surplus (w=0) but vacancies offering jobs to workers who do have other offers receive none of the surplus (w=1). Julien, Kennes, and King (2006) interpret this result in terms of what they call the Mortensen rule (Mortensen 1982) – that efficiency in matching is attained if the "initiator" of the match gets the total surplus. By mimicking our proof of Proposition 2, we can show that this assumed wage determination mechanism is in fact the symmetric equilibrium outcome in a directed search model with wage posting when a=v in a finite labor market.⁶

Could an adaptation of the Julien, Kennes, and King (2006) wage determination mechanism to a large labor market with $a \in \{2, ..., A\}$ deliver the constrained efficient outcome? In order to do this, we would have to assume that a worker receives w=1 if he or she (i) has multiple offers or (ii) has only one offer and is the only applicant to the vacancy making that offer but receives w=0 if he or she has only one offer but the vacancy making that offer has other applicants. The extra twist in the mechanism (setting w=1 in case (ii) above) is required because $p(\theta;a) < 1$ when $a \neq v$. This mechanism delivers an expected payoff to vacancy creation equal to the right-hand side of equation (11); thus, it implements the constrained efficient outcome.

We argue, however, that this wage determination protocol cannot be sustained as an equilibrium outcome in a large labor market. One reason is that it requires that when a worker is the sole applicant for a job, the vacancy has to reveal this, even though it is not in the vacancy's interest to

⁵The fact that the social planner cannot improve on the equilibrium outcome in this case does not mean that welfare increases as $a \to \infty$. To the contrary, $\gamma(\theta; a)$ increases and $m(\theta; a)$ decreases in a as $a \to \infty$. Increasing a makes the planner's problem more difficult. Similarly, even though equilibrium is constrained efficient when a = 1, welfare may increase by moving to a > 1.

⁶The intuition for constrained efficiency in a large labor market when a=1 is quite different from the intuition for the finite labor market case when a=v. In the former, constrained efficiency is a result of competition, and competition requires a labor market sufficiently large that individual vacancies have negligible market power. When a=v, constrained efficiency is a result of perfect monopoly power – the entire surplus goes to the vacancy if there is no competition for the applicant it selects and to the worker if he or she winds up having the monopoly power. The monopoly intuition does not require that the labor market be large.

do so. More fundamentally, even if a worker somehow knew that he or she was the only candidate for a job, this wage determination protocol would not survive if one allowed for competing mechanisms. The proposed mechanism gives an applicant an expected payoff of

$$\gamma(\theta; a) + (1 - \gamma(\theta; a))(1 - p(\theta; a)) = \gamma(\theta; a) + (1 - \gamma(\theta; a))w^*(\theta; a).$$

Note, however, that the proposed mechanism is equivalent in terms of expected payoff to one in which each vacancy posts $w^*(\theta; a)$ and pays that wage to its selected applicant unless that applicant has multiple offers, in which case the wage is bid up to one by Bertrand competition. However, Proposition 2 tells us that the proposed mechanism is not an equilibrium. If all vacancies were "in effect" posting $w^*(\theta; a)$, it would be in the interest of individual vacancies to post a slightly lower wage. Although we do not want to claim that it is "impossible" to find a mechanism that could implement the efficient outcome, the above argument suggests that Proposition 3 is more general than one might suspect at first glance. Specifically, this argument rules out any alternative mechanism that (i) has full $ex\ post$ competition (and, by equation (11), full $ex\ post$ competition is required for efficiency) and (ii) yields a positive expected payoff when a worker receives only one offer.

4 Steady State

We now turn to steady-state analysis for a labor market with directed search and multiple applications. We work with the limiting case in which $u, v \to \infty$ with $v/u = \theta$ and $a \in \{2, ..., A\}$ fixed. Since only the ratio of v to u matters in the limiting case, we normalize the labor force to 1; thus, u is interpreted as the unemployment rate.

In steady-state, workers flow into employment with probability $m(\theta; a)$ per period. We assume that matches break up exogenously with probability δ , giving the countervailing flow back into unemployment. Similarly, jobs move from vacant to filled with probability $\frac{m(\theta; a)}{\theta}$ and back again with probability δ . Steady-state analysis thus allows us to endogenize vacancies and unemployment. More importantly, moving to the steady state means

⁷ An applicant with multiple offers gets the full surplus (this occurs with probability $\gamma(\theta; a)$) as does an applicant who receives only one offer but does so from a vacancy that has no other applicants (this occurs with probability $(1 - \gamma(\theta; a))(1 - p(\theta; a))$). Otherwise, the applicant gets nothing.

that those unemployed who fail to find an acceptable job in the current period can wait and apply again in the future. In the case of a=1, this is not particularly interesting since, in equilibrium, there is no gain to waiting. However, with multiple applications, the ability of the unemployed to hold out for a situation in which vacancies engage in Bertrand competition for their services, albeit at the cost of delay, implies a positive reservation wage. This leads to a simple and appealing model in which labor market tightness and the reservation wage are simultaneously determined. On the one hand, the lower is the reservation wage of the unemployed, the more vacancies firms want to create. On the other hand, as the labor market becomes tighter, i.e., as θ increases, the unemployed respond by increasing their reservation wage. The steady-state equilibrium reservation wage is positive, thus suggesting that moving to the steady-state might restore efficiency. Our final result in this section shows that this is not the case – there is still excessive vacancy creation.

The analysis proceeds as follows. Suppose the unemployed set a reservation wage R. With multiple applications, the wage-posting problem for a vacancy is qualitatively the same as in the one-period game. Whatever common wage might be posted at other vacancies, each individual vacancy has the incentive to undercut. In the one-period game, this implies a monopsony wage of w=0; in the steady state, this same mechanism implies a dynamic monopsony wage of $w=R.^8$ To avoid complicated dynamics, we assume that a vacancy that fails to hire its candidate in period t cannot carry its queue of remaining applicants (if any) over to the next period. As a consequence, workers start with a new application round in each period since their earlier applications are no longer on file. This implies that the probability that an unemployed worker finds a job in any period and the probability that he or she is hired at the competitive wage, conditional on finding a job, are the same as in the single-period model; i.e., equations (1) and (3) for $m(\theta; a)$ and $\gamma(\theta; a)$ continue to apply.

We begin by examining the value functions for jobs and for workers. A job can be in one of three states – vacant, filled paying the competitive wage, and filled paying R. Let V, J(1), and J(R) be the corresponding values. The

⁸We restrict our attention to stationary strategies (as do JKK in their dynamic extension). That is, we rule out reputation mechanisms that might avert bidding wars. Since any two vacancies that might consider avoiding a bidding war today interact directly in any future period with probability zero, this seems reasonable. We consider a mechanism that rules out Bertrand competition in a static setting in Section 5.3 below.

value of a vacancy is

$$V = -c_v + \frac{1}{1+r} \{ \frac{m(\theta; a)}{\theta} [\gamma(\theta; a) J(1) + (1 - \gamma(\theta; a)) J(R)] + (1 - \frac{m(\theta; a)}{\theta}) V \}.$$

Maintaining a vacancy entails a cost c_v , which is incurred at the start of each period. Moving to the end of the period, and thus discounting at rate r, the vacancy has hired a worker with probability $\frac{m(\theta;a)}{\theta}$. With probability $\gamma(\theta;a)$, the worker who was hired had his or her wage bid up to the competitive level, thus implying a value of J(1). With probability $1 - \gamma(\theta;a)$ the worker was hired at w = R, thus implying a value of J(R). Finally, with probability $1 - \frac{m(\theta;a)}{\theta}$, the vacancy failed to hire, in which case the value V is retained.

Free entry implies V=0 so the analysis for vacancies remains the same; that is, free entry turns the dynamic game into one that is essentially static for vacancies. Given V=0, there is no incentive for vacancies competing for a worker to drop out of the Bertrand competition before the wage is bid up to w=1 (thus justifying the notation J(1)). This in turn implies that we also have J(1)=0. Inserting these equilibrium conditions into the expression for V gives

$$\frac{m(\theta; a)}{\theta} (1 - \gamma(\theta; a)) J(R) = c_v (1 + r).$$

At the same time, the value of employing a worker at w = R is

$$J(R) = (1 - R) + \frac{1}{1 + r} [(1 - \delta)J(R) + \delta V].$$

Again using V = 0, we have

$$J(R) = \frac{1+r}{r+\delta}(1-R).$$

Combining these equations gives the first steady-state equilibrium condition,

$$c_v = \frac{m(\theta; a)}{\theta} (1 - \gamma(\theta; a)) \frac{1 - R}{r + \delta}.$$
 (12)

A worker also passes through three states – unemployed, employed at the competitive wage, and employed at R. The value of unemployment is defined by

$$U = \frac{1}{1+r} \{ m(\theta; a) [\gamma(\theta; a) N(1) + (1 - \gamma(\theta; a)) N(R)] + (1 - m(\theta; a)) U \},$$

where N(1) and N(R) are the values of employment at w = 1 and w = R, respectively. These latter two values are in turn defined by

$$N(1) = 1 + \frac{1}{1+r} \{ (1-\delta)N(1) + \delta U \}$$

$$N(R) = R + \frac{1}{1+r} \{ (1-\delta)N(R) + \delta U \}.$$

The reservation wage property, i.e., N(R) = U, then implies

$$U = \frac{1+r}{r}R$$

$$N(1) = \frac{(1+r)}{r(r+\delta)}(r+\delta R).$$

Inserting these expressions into the expression for U and rearranging gives the second steady-state equilibrium condition,

$$R = \frac{m(\theta; a)\gamma(\theta; a)}{r + \delta + m(\theta; a)\gamma(\theta; a)}.$$
 (13)

The final equation for the steady-state equilibrium is the standard flow (Beveridge curve) condition for unemployment. Since the labor force is normalized to 1, this is

$$u = \frac{\delta}{\delta + m(\theta; a)}. (14)$$

Equations (13) and (14) show that, as is common in this class of models, once labor market tightness (θ) is determined, the other endogenous variables – in this case, R and u – are easily determined. Using equation (13) to eliminate R from equation (12) gives the equation that determines the steady-state equilibrium value of θ , namely,

$$c_v = \frac{m(\theta^{**}; a)}{\theta^{**}} \frac{1 - \gamma(\theta^{**}; a)}{r + \delta + m(\theta^{**}; a)\gamma(\theta^{**}; a)}.$$
 (15)

Using our results on the properties of $m(\theta; a)$ and $\gamma(\theta; a)$, we can show that the right-hand side of equation (15) equals $\frac{1}{r+\delta}$ as $\theta \to 0$, that it goes to zero as $\theta \to \infty$, and that its derivative with respect to θ is negative for all $\theta > 0$. Equation (15) thus has a unique solution for each $c_v \in (0, \frac{1}{r+\delta}]$.

The natural next step is to compare equilibrium steady-state labor market tightness with the constrained efficient value of θ . The planner's problem

is to choose the level of labor market tightness that maximizes the discounted value of output net of vacancy costs for an infinitely-lived economy.⁹ That is, the planner's problem is to maximize

$$\sum_{s=0}^{\infty} \left(\frac{1}{1+r} \right)^s \left(1 - u_s - c_v \theta_s u_s \right)$$

subject to

$$u_{s+1} - u_s = \delta(1 - u_s) - m(\theta_s; a)u_s$$

with u_0 given.

The Lagrangean for this problem is

$$\sum_{s=0}^{\infty} \left(\frac{1}{1+r} \right)^s \left[(1 - u_s - c_v \theta_s u_s) + \lambda_s (u_{s+1} - u_s - \delta(1 - u_s) + m(\theta_s; a) u_s) \right]$$

The necessary conditions for this problem evaluated at the steady state are

$$-c_v u + \lambda m_{\theta}(\theta; a) u = 0$$

$$-1 - c_v \theta + \lambda [r + \delta + m(\theta; a)] = 0$$

Eliminating λ gives

$$c_v = \frac{(1 + c_v \theta^*) m_\theta(\theta^*; a)}{r + \delta + m(\theta^*; a)}.$$
 (16)

Now we can compare the levels of labor market tightness implied by equations (15) and (16). Using equations (1) and (3), equation (15) can be rewritten as

$$c_v(r+\delta+m(\theta^{**};a)) = (1+c_v\theta^{**})(1-\frac{\theta^{**}}{a}(1-e^{-a/\theta^{**}}))^{a-1}(1-e^{-a/\theta^{**}}). (17)$$

Using equation (8), equation (16) can be rewritten as

$$c_v(r + \delta + m(\theta^*; a)) = (1 + c_v \theta^*) (1 - \frac{\theta^*}{a} (1 - e^{-a/\theta^*}))^{a-1} (1 - e^{-a/\theta^*} - \frac{a}{\theta^*} e^{-a/\theta^*}).$$
(18)

As in the single-period analysis, θ^* is the constrained efficient level of labor market tightness, i.e., the value of θ that solves equation (16), and θ^{**} is the equilibrium level of labor market tightness, i.e., the value of θ that solves equation (15). Comparing equations (17) and (18) yields the following:

⁹We consider only stationary solutions, but this is not likely to be restrictive in our model. There are two standard reasons why a nonstationary solution might be optimal. First, as shown in Shimer and Smith (2001), a nonstationary solution can be optimal in a matching model with two-sided heterogeneity when agents' characteristics are complements in production. A nonstationary solution may also be optimal if there are increasing returns to scale in the matching function. Neither of these features is present in our model.

Proposition 4 Let $u, v \to \infty$ with $v/u = \theta$ and $a \in \{2, ..., A\}$ fixed. Then in steady state, $\theta^{**} > \theta^*$.

Proposition 4 indicates that, as in the single-period analysis, when the unemployed make a fixed number of multiple applications per period $(a \in \{2, ..., A\})$, equilibrium is constrained inefficient. Specifically, there is too much vacancy creation. This result holds even though the ability of the unemployed to reject offers in favor of waiting for a more favorable outcome in some future period implies a dynamic monopsony wage above the single-period monopsony wage of zero. The intuition for the inefficiency result is the same as in the static model. As before, the social benefit of opening an additional vacancy, the right-hand side of equation (18), is $p(\theta; a)$ times the private benefit, the right-hand side of equation (17).

5 Extensions and Robustness Checks

In this section, we focus on three simplifying assumptions that we made in our basic model. These assumptions are: (i) that the number of applications sent out by each worker is a parameter of the search technology, (ii) that each vacancy can process at most one applicant per period, and (iii) that two or more vacancies competing for the same worker engage in Bertrand competition for that worker's services. Accordingly, we examine what happens to our results if (i) the number of applications per worker is a choice variable, (ii) each vacancy can process more than one applicant, and (iii) vacancies pursue strategies that rule out Bertrand competition. In all three robustness exercises, we confirm our result that equilibria in models of directed search with multiple applications are inefficient.

5.1 Endogenous a

We have assumed that each worker makes a applications, where $a \in \{1, 2, ..., A\}$ is exogenously given. Since the equilibrium level of labor market tightness is efficient when a = 1 but inefficient when $a \in \{2, ..., A\}$, it is natural to ask whether – and under what circumstances – workers would choose to make only one application or more than one. In addressing this question, we consider only pure-strategy symmetric equilibria in application strategies. That is, assuming that all other workers make a applications, under what conditions (taking into account how firms react to all workers choosing a) is it in the individual worker's interest also to choose a?

To make endogenizing a an interesting problem, there must be a cost associated with applications, so we assume that each application costs c_a to submit. In the one-shot game, there are then only 2 exogenous parameters, the cost of posting a vacancy, c_v , and the cost of submitting an application, c_a . We need only consider $0 \le c_v \le 1$ and $0 \le c_a \le 1$ since worker output equals 1 and if $c_v > 1$, no firm would post a vacancy, and if $c_a > 1$, no worker would make an application. Thus for each (c_v, c_a) in the unit square we can ask (i) what are the free-entry equilibrium values of θ and a and (ii) what values of θ and a would a social planner choose?

We start with the equilibrium problem and ask: For what values of (c_v, c_a) is a = 1 consistent with equilibrium? For what values of (c_v, c_a) is a = 2 consistent with equilibrium? Etc. We address this problem numerically as follows.

Consider a candidate equilibrium in which all workers make a applications. Then, for each θ , we know what wage vacancies choose to post (from equation (2) if a = 1; zero if $a \in \{2, ..., A\}$, and we know $m(\theta; a)$. We pick a value of c_v from a grid over (0,1). From the free-entry condition (equation (6) if a = 1; equation (7) otherwise), there is a corresponding implied value of θ . We then ask, using the value of θ implied by the free-entry condition, for what values of c_a is an individual worker's expected payoff maximized by choosing to send out the same number of applications as all other workers do? We answer this numerically by comparing the expected payoff associated with choosing a when all other workers also choose a with those associated with choosing a-1, a-2, ... and a+1, a+2, ..., etc.¹⁰ For the particular c_v that we chose, this gives us a range of values for c_a . We then repeat for the next value of c_v , etc. The outcome of this algorithm is the set of (c_v, c_a) combinations in the unit square that are consistent with a pure-strategy symmetric equilibrium in which all workers make a applications. We carry out this process for a wide range of values for a.

Next, we address the social planner's problem. Given (c_v, c_a) , the natural social planner's problem is

$$\max_{\theta} m(\theta; a) - c_v \theta - c_a a,$$

where $\theta \ge 0$ and $a \in \{0, 1, 2, ...\}$. We know this problem is concave in θ for a given a. Thus, if (θ^*, a^*) solves the social planner's problem, we must have

$$c_v = m_\theta(\theta^*; a^*),$$

¹⁰This comparison can be carried out in a finite number of steps since the maximum number of applications a worker might make is limited by the requirement that the total cost of submitting applications be less than one.

and $\theta^* = \theta^*(a^*; c_v)$ has a unique solution. We can plug this back into the social planner's objective and maximize numerically with respect to a. This gives a^* (and θ^*) as functions of (c_v, c_a) . We can then compare the equilibrium unit square with the social planner unit square.

The qualitative results of this exercise are as follows. First, although there are many parameter configurations for which the equilibrium number of applications, a^{**} , equals 1, this outcome requires relatively high values of c_a . Second, the equilibrium number of applications increases as c_a falls (as one would expect). Third, there are parameter configurations that admit multiple equilibria. This reflects a complementarity between workers' and firms' strategies. For example, if all workers choose a = 1, then vacancies post a positive wage, $w(\theta; 1) > 0$. For some values of θ (equivalently, for some values of c_v) it is not worthwhile for workers to submit a second application. On the other hand, if all workers choose a=2, then w=0, and it cannot be worthwhile for a worker to deviate to a = 1. Fourth, there are many parameter configurations for which no symmetric pure-strategy equilibrium exists. One parameter region in which this is the case is the set of (c_n, c_n) combinations in which individual workers would prefer not to send out any applications when all other workers choose a=1. This occurs when both c_v and c_a are relatively high. There are, however, other (c_v, c_a) combinations for which no symmetric pure-strategy equilibrium exists. Fifth, for relatively low values of c_a , there are parameter regions with unique equilibria at a^{**} 2, $a^{**} = 3$, etc.

In the parameter regions in which a symmetric pure-strategy equilibrium (or equilibria) exists, we find $a^{**} \geq a^*$. Specifically, there are parameter configurations for which $a^* = a^{**} = 1$ (where $a^{**} = 1$ may either be unique or one of two or more equilibrium possibilities). However, when $a^{**} \geq 2$, we find $a^{**} > a^*$. This occurs when c_v and c_a are low relative to the output produced by a match. That is, for what we view as reasonable values of c_v and c_a , the equilibrium number of applications exceeds the socially optimal value. The reason is simply that individual workers, when deciding how many applications to submit, fail to take into account the externality they impose on other workers. The countervailing effect that one might expect – that an increase in worker applications should make it easier for firms to fill their vacancies – is not sufficient to offset this externality and, indeed, may even be negative because of the coordination failure among vacancies.

Finally, endogenizing a does not affect our basic result that, while directed search with one application always leads to the efficient level of labor market tightness, this is not the case with multiple applications. For (c_v, c_a) combinations such that $a^* = a^{**} = 1$, we, of course, have $\theta^* = \theta^{**}$. For

almost all parameter configurations for which $a^{**} > a^*$, we find $\theta^{**} > \theta^*$ as we did before. There is a small set of parameter configurations, however, for which $\theta^{**} < \theta^*$.¹¹ This appears at first glance to be inconsistent with Proposition 3, but note that in that Proposition, we imposed the restriction that $a^* = a^{**}$.

The bottom line of this robustness exercise is that when a is endogenous and when workers choose a > 1, equilibrium may be inefficient in two ways. There are always too many applications per worker and labor market tightness is generally not at the level the social planner would choose. The assumption that a is an exogenous parameter of the search technology, which we made in order to make our basic model as transparent as possible, is not driving our results on the inefficiency effects of multiple applications.

5.2 Shortlisting

Our inefficiency result is based on a double coordination failure. Not only are workers unable to coordinate in terms of where they send their applications, but vacancies are unable to coordinate in terms of which applicants they try to hire. In our basic model, we represented the coordination failure among vacancies in a clean but extreme way. A natural question is the extent to which our results depend on our assumption that each vacancy can pursue at most one applicant.

To address this question, we now consider a version of the basic one-shot model in which each vacancy can make up to two offers. Specifically, we assume that vacancies form "short lists" as follows. If two or more workers apply to a vacancy, the vacancy selects two applicants at random and rejects the others. It selects one of its chosen applicants to receive its first-round offer. The other applicant, if she is not hired by another vacancy in the first round, gets a second-round offer in the event that the vacancy doesn't hire in the first round. If only one worker applies to a vacancy, then that worker gets the vacancy's first-round offer. To keep the algebra as simple as possible, we analyze this model for the case of a = 2.

This extension makes our model far more difficult. The basic reason is that when a worker thinks about applying to a vacancy that is deviat-

To understand why this can happen, recall that $m(\theta; a)$ is decreasing in a for sufficiently high a. Reducing the matching rate hurts both workers and vacancies. When workers choose $a^{**} > 1$, the planner can improve on the equilibrium allocation by reducing a. When $m(\theta; a)$ is decreasing in a, the reduction in a increases the matching rate and can, for some parameter values, increase the marginal benefit of opening a vacancy, i.e., $m_{\theta}(\theta; a)$, sufficiently so that the social planner would also raise θ . This happens despite the fact that were a fixed at either a^* or a^{**} , the social planner would want to decrease θ .

ing from the putative equilibrium wage, the indifference condition becomes considerably more complicated. A worker's application strategy affects the probabilities of being placed on 0, 1, or 2 short lists; the worker could be in first or second place on these short lists, etc. In addition, an intermediate wage arises in this model. Consider two vacancies competing for the same applicant in the first round. If either or both of these vacancies has a second-round candidate, then Bertrand competition in the first round stops before the competitive level.

Our analysis of shortlisting follows the same road map that we used for our basic model. We first derive the matching probability, assuming a symmetric equilibrium posted wage. Second, taking θ as given, we derive the symmetric equilibrium wage-posting strategy for vacancies. Finally, we characterize the free-entry equilibrium level of labor market tightness and the corresponding constrained efficient level and compare the two. The central result of our analysis still holds – the equilibrium level of θ exceeds the efficient level.

Because the details of the shortlisting extension are tedious, we present the derivations in the first section of the web supplements to this paper. Here, in the text, we simply summarize and comment on our results.

We begin with the matching probability. Assuming the existence of a symmetric equilibrium posted wage, that is, assuming that all vacancies are equally attractive *ex ante*, the probability that a worker finds a job is

$$m(\theta) = 1 - (1 - q_1)^2 (1 - q_2)^2,$$

where q_1 is the probability that an application leads to a first-round offer and q_2 is the probability that an application leads to a second-round offer given that it does not generate a first-round offer. An explanation of the form of $m(\theta)$ and expressions for q_1 and q_2 are given in Section 1 of the web supplements to this paper. Note that the probability that an application leads to a first-round offer is the same as the probability that the application would have generated an offer had there been only one round; i.e., $q_1 = q$ (from the basic model). The obvious result thus follows; namely, for each value of θ , shortlisting increases the probability that a worker finds a job.

From the social planner's perspective, the only effect of shortlisting is to change the form of $m(\theta)$. The effect on equilibrium is, however, much more complicated. For low values of θ , the equilibrium analysis is qualitatively similar to the one we carried out for our basic model. All vacancies post a wage of zero. Bertrand competition for an applicant who has two first-round offers either drives the wage to the competitive level (if neither of

the competing vacancies has a second-round candidate) or to the intermediate wage (if at least one of the vacancies has a second-round candidate). An applicant who, having failed to get any first-round offers, gets a single second-round offer receives the monopsony wage (zero). An applicant who gets two second-round offers receives a wage of one.

For higher values of θ (the cutoff value is approximately $\theta=0.42$), there are multiple equilibria. For example, when $\theta=1$, any wage in the interval [0.20,0.23] (approximately) is consistent with equilibrium. Multiple equilibria arise because the derivative of expected profit with respect to the potential deviant's wage is discontinuous at the equilibrium wage. The reason that w=0 is not an equilibrium posted wage for higher values of θ has to do with the change in application incentives implied by shortlisting. In our basic model, a worker whose application is accepted by more than one vacancy necessarily receives a wage of one, and workers are willing to apply to vacancies posting w=0 in hopes of hitting the jackpot. With shortlisting, however, a worker can wind up with the posted wage even if both of her applications are accepted – specifically, if she is first on one vacancy's short list and second on the other's. (When θ is low, w=0 arises even with shortlisting due to a lack of competition among vacancies.)

Whether θ is low, so w=0 is the unique posted wage, or θ is high, so there are multiple equilibria, workers can receive three different wages – the posted wage, the intermediate wage, and the competitive wage. The intermediate wage, s, is determined by

$$1 - s = (1 - q_1)(1 - q_2)(1 - w).$$

The left-hand side of this expression is the profit that a vacancy realizes if it hires its first-round candidate at wage s. The right-hand side is the expected profit for a vacancy that received two applications should it choose to proceed to the second round. With probability $1-q_1$ the vacancy's second-place candidate will still be available after the first round. Conditional on still being available, this candidate will fail to get a competing second round offer with probability $1-q_2$. The vacancy then realizes a profit of 1-w.

For each value of θ , the next step is to compute the expected profit of a vacancy, say $\pi(\theta)$. When there are multiple equilibria, we use the highest possible equilibrium wage. At this wage, $\pi(\theta)$ is at its lowest possible level; hence the incentive to create vacancies is as small as possible. The free-entry equilibrium value of labor market tightness, θ^{**} , is determined by

$$c_v = \pi(\theta^{**}),$$

which is analogous to equation (7) in our basic model. The efficient value of labor market tightness, θ^* , is determined by

$$c_v = m'(\theta^*),$$

precisely as in the basic model. The only effect of shortlisting is to change the form of $m(\cdot)$.

It is straightforward to compute $m'(\theta)$ and $\pi(\theta)$ numerically. Both of these functions are positive and decreasing in θ , and $\pi(\theta) > m'(\theta)$ for each $\theta > 0$. Equivalently, $\theta^{**} > \theta^*$. That is, the central result of our basic model, namely, that there is excessive vacancy creation in equilibrium, continues to hold when we extend our model to allow for shortlisting.

The fact that shortlisting reduces matching frictions does not necessarily mean that shortlisting makes equilibrium more efficient in the sense that θ^{**} "gets closer to" θ^* . Shortlisting affects both the social planner's problem and the market outcome so that both θ^* and θ^{**} change. The fact that $\theta^{**} > \theta^*$ continues to hold when we allow for shortlisting suggests that our result on the inefficiency of directed search equilibrium when workers make multiple applications is robust to our assumption that vacancies can consider at most one application. Even if vacancies could process all their applicants, some vacancies that receive applications would nonetheless lose all their candidates to rival vacancies. Shortlisting reduces the coordination problem among vacancies but does not eliminate it.

5.3 Offer-Beating Strategies

In our basic model, we assumed that if a worker receives offers from two or more vacancies, those vacancies then engage in Bertrand competition for the worker's services. Although the Bertrand assumption is standard in the literature, it can be debated in our environment. A vacancy that is about to lose a worker to a rival should be indifferent between letting the worker take the other job versus entering into Bertrand competition. After all, both policies, conceding or competing, lead to the same zero-profit outcome.

A natural alternative is to assume that each vacancy announces a wage and then commits not to engage in ex post bidding. However, as discussed in footnote 1, in this case, there is no equilibrium with a common posted wage. Moreover, simply assuming commitment is unsatisfactory because if all other vacancies were to follow the commitment strategy, any vacancy whose candidate has multiple offers could do better by deviating from that strategy and offering slightly more. This leads us to consider offer-beating strategies.

We define such strategies as follows:

- 1. Post w.
- 2. If all other vacancies pursuing the same applicant post w or less, continue to offer w.
- 3. If at least one other vacancy pursuing the same applicant posts w' > w or makes a counteroffer w' > w, make a counteroffer above w'. If one or more rivals makes a counteroffer to the counteroffer, respond in kind; i.e., engage in Bertrand competition.

Of course, these strategies only are relevant when workers make more than one application.

Offer-beating strategies are analogous to the price-beating strategies that have been used in the industrial organization literature to rule out Bertrand competition in prices. Price-beating strategies are sometimes used in that literature as a foundation for "kinked demand curves" (e.g., Tirole 1988, pp. 243-45). Typically, there is a continuum of price-beating Nash equilibria – absent any consideration of equilibrium refinements, there is a continuum of prices at which the demand curve can kink.

We begin our analysis of offer-beating equilibria taking θ as given. We first show that for each θ , there is a continuum of offer-beating Nash equilibria. We then show that when we endogenize θ , all of these equilibria are inefficient. The details of our analysis and the proofs of our results are given in Section 2 of the web supplements to this paper. Specifically, we prove the following:

Proposition 5 Let $\overline{w}(\theta; a) = \frac{\frac{a}{\theta}e^{-a/\theta}}{1 - e^{-a/\theta}}$. There exists a continuum of symmetric offer-beating Nash equilibria indexed by $w \in [0, \overline{w}(\theta; a)]$.

Proposition 6 There is excessive vacancy creation in any symmetric offerbeating Nash equilibrium.

This indicates that the inefficiency associated with multiple applications is not an artifact of *ex post* Bertrand competition for applicants.

To gain further insight into the inefficiency result in our basic model, it is useful to examine why offer-beating equilibria are also inefficient. Offer-beating strategies lead to implicit collusion among vacancies. This collusion shuts down all *ex post* competition, that is, the competition that can occur

after workers make their applications and vacancies select their candidates. Offer-beating strategies can also shut down ex ante competition among vacancies, that is, the competition among vacancies to attract applicants, to a greater or lesser extent, ranging from a complete absence of ex ante competition when w=0 to full ex ante competition when $w=\overline{w}(\theta;a)$. Note that $\overline{w}(\theta;a)=w^*$, the wage that a social planner would set if vacancies engaged in Bertrand competition rather than following offer-beating strategies. That is, were there full ex post competition, the social planner would want to implement full ex ante competition. If vacancies follow offer-beating strategies, the social planner would prefer a wage above $\overline{w}(\theta;a)$ to compensate for the lack of ex post competition. Absent ex post competition, the decision to post a vacancy neglects the externality that arises when a vacancy hires a worker with one or more other offers.

In our model, Bertrand competition fully implements $ex\ post$ competition but at the cost of eliminating $ex\ ante$ competition. Offer-beating strategies have the potential to achieve full $ex\ ante$ competition but by design shut down $ex\ post$ competition. The lesson we draw is that in directed search models in which workers make a finite number of multiple applications and vacancies post wages to attract workers, there is a fundamental tradeoff between $ex\ ante$ and $ex\ post$ competition.

6 Concluding Remarks

In this paper, we construct an equilibrium search model of a large labor market in which workers, after observing all posted wages, submit a fixed number of applications, $a \in \{1, ...A\}$, to the vacancies that they find most attractive. We derive the symmetric equilibrium matching probability and the common posted wage. When a = 1, our analysis is a large labor market version of BSW. However, when $a \in \{2, ...A\}$, i.e., when workers make multiple applications, the symmetric equilibrium of our model is radically different. With multiple applications, the matching probability in our model reflects the interplay of two coordination failures – an urn-ball failure among workers and a multiple-application failure among vacancies. In addition, when workers make more than one application, all vacancies post the monopsony wage, but there is dispersion in wages paid. Workers who receive only one job offer are paid the monopsony wage, but those who receive multiple offers get the competitive wage. When workers make a single application or when they apply to an arbitrarily large number of vacancies, equilibrium is constrained efficient; but when workers make a finite number of multiple applications, too many vacancies are posted. These results, both positive and normative, carry over from the single-period model to a steady-state framework and they are robust with respect to reasonable variations in our key assumptions.

Directed search is an appealing way to model equilibrium unemployment and wage dispersion. In reality, workers do direct their applications to attractive vacancies, but unemployment nonetheless persists as a result of coordination failures on both sides of the labor market. In addition, those workers who are lucky enough to generate competition for their services do in fact have their wages bid up. The contribution of our paper is to show that these realistic features can be captured in a tractable equilibrium model and, more importantly, that when these features are incorporated, equilibrium is not constrained efficient.

References

- [1] Albrecht, J., P. Gautier, S. Tan, and S. Vroman (2004), Matching with Multiple Applications Revisited, *Economics Letters*, 84, 311-314.
- [2] Burdett, K. and K. Judd (1983), Equilibrium Price Dispersion, *Econometrica*, 51, 955-970.
- [3] Burdett, K., S. Shi, and R. Wright (2001), Pricing and Matching with Frictions, *Journal of Political Economy*, 109, 1060-1085.
- [4] Chade, H. and L. Smith (2004), Simultaneous Search, mimeo.
- [5] Diamond, P. (1971), A Model of Price Adjustment, Journal of Economic Theory, 3, 156-168.
- [6] Galenianos, M. and P. Kircher (2005), Directed Search with Multiple Job Applications, mimeo.
- [7] Hosios, A. (1990), On the Efficiency of Matching and Related Models of Search and Unemployment, *Review of Economic Studies*, 57, 279-298.
- [8] Julien, B., J. Kennes, and I. King (2000), Bidding for Labor, Review of Economic Dynamics, 3, 619-649.
- [9] _____ (2006), The Mortensen Rule and Efficient Coordination Unemployment, Economics Letters, 90, 149-155.
- [10] Moen, E. (1997), Competitive Search Equilibrium, Journal of Political Economy, 105, 385-411.
- [11] Mortensen, D. (1982), Efficiency of Mating, Racing, and Related Games, American Economic Review, 72, 968-979.
- [12] Philip, J. (2003), Matching with Multiple Applications: A Note, mimeo.
- [13] Postel-Vinay, F. and J-M. Robin (2002), The Distribution of Earnings in an Equilibrium Search Model with State-Dependent Offers and Counter-Offers, *International Economic Review*, 43, 989-1016.
- [14] Shimer, R. and L. Smith (2001), Nonstationary Search, mimeo.
- [15] Tirole, J. (1988), The Theory of Industrial Organization, The MIT Press, Cambridge, MA.

Appendices

A Proof of Proposition 1

We now sketch the proof of Proposition 1. The full proof is given in Albrecht et. al. (2004). We compute $m(\theta; a)$ as follows. The probability that a worker finds a job is one minus the probability that he or she gets no job offers. Consider a worker who applies to a vacancies, and let the random variables $X_1, X_2, ..., X_a$ be the number of competitors that he or she has at vacancy 1, vacancy 2, ..., vacancy a. The probability that the worker gets no job offers can be expressed as

$$\sum \dots \sum \frac{x_1}{x_1+1} \frac{x_2}{x_2+1} \dots \frac{x_a}{x_a+1} P[X_1 = x_1, X_2 = x_2, \dots X_a = x_a].$$

In general, the random variables $X_1, X_2, ..., X_a$ are not independent, making the computation of the joint probability a difficult one. (Albrecht et. al. 2004 and Philip 2003 give an expression for the joint probability.) The intuition for dependence is straightforward. Consider, for example, a labor market in which u and v are small and in which each worker makes a=2applications. Then, if a worker has relatively many competitors at the first vacancy to which he or she applies, it is more likely that his or her second application has relatively few competitors. The key to Proposition 1 is that this dependence vanishes in the limit. In the limit, the fact that a worker has an unexpectedly large number of competitors at one vacancy says nothing about the number of competitors he or she faces elsewhere. The joint probability then equals the product of the marginals, and the probability that a worker gets at least one offer can be computed as 1 - $\left(\sum \frac{x}{x+1}P[X=x]\right)^a$. As $u,v\to\infty$ with $v/u=\theta$, the number of competitors an applicant faces at any particular vacancy, X, converges in distribution to a Poisson $(\frac{a}{\rho})$ random variable. A straightforward computation then gives equation (1)

If a=1, there is no problem of dependence. The number of competitors that a worker has at the vacancy to which he or she applies is a $bin(u-1,\frac{1}{v})$ random variable. The probability that a worker gets an offer is then

$$1 - \sum_{x=0}^{u-1} \frac{x}{x+1} \binom{u-1}{x} \left(\frac{1}{v}\right)^x \left(1 - \frac{1}{v}\right)^{u-1-x} = \frac{v}{u} \left[1 - \left(1 - \frac{1}{v}\right)^u\right].$$

With a change in notation, this result is the same as the one given in BSW. Taking the limit of this matching probability as $u, v \to \infty$ with $v/u = \theta$

gives $m(\theta;1) = \theta(1-e^{-1/\theta})$, as equation (1) implies. The case of a=v is the polar opposite. In this case, $X_1 = X_2 = \dots = X_a = u-1$ with probability one, so the probability a worker gets an offer is $1 - (\frac{u-1}{u})^v$, as in JKK. Taking the limit as $u, v \to \infty$ with $v/u = \theta$ gives

$$m(\theta) = 1 - e^{-\theta}.$$

The same expression can be derived by taking the limit of $m(\theta; a)$ as $a \to \infty$ in equation (1).

B Proof of Proposition 2

As discussed in the text, we need to show that when a=1, the wage $w(\theta;1)$ has the property that if all vacancies, with the possible exception of a potential deviant (D), post that wage, then it is also in D's interest to post $w(\theta;1)$. When $a \in \{2,...,A\}$, we need to show that no matter what common wage is posted by other vacancies, it is always in D's interest to undercut, thus driving $w(\theta;a)$ to zero.

Suppose D posts a wage of w' and that each nondeviant vacancy (N) posts w. Then D's expected profit is

 $\pi(w'; w) = (1 - w')P[D \text{ gets at least one application}]P[\text{selected applicant has no other offer}]$

Let k be the probability that any one worker applies to D. In symmetric equilibrium, k must be the same for all workers. As $u \to \infty$, k must go to zero; otherwise, any applicant to D would have an infinity of competitors and therefore would get the job at D with probability zero. We let $u \to \infty$ and $k \to 0$ in such a way that $ku = \xi$ stays constant; thus, in a large labor market, the number of applications sent to D is a Poisson (ξ) random variable. We therefore have

$$P[D \text{ gets at least one application}] = 1 - e^{-\xi}.$$

The parameter ξ depends on w' and w through an indifference condition, which we develop below. Finally, the last term on the right-hand side of $\pi(w'; w)$ can be written as

 $P[\text{selected applicant has no other offer}] = (1-q)^{a-1},$

where q is the probability that any one application to an N vacancy leads to an offer. We thus have

$$\pi(w';w) = (1-w')(1-e^{-\xi})(1-q)^{a-1}.$$

The parameter ξ determines the probability (call it q^D) that a worker who applies to D gets an offer from that vacancy, as follows:

$$q^{D} = \sum_{x=0}^{\infty} \frac{1}{x+1} \frac{e^{-\xi} \xi^{x}}{x!} = \frac{1}{\xi} (1 - e^{-\xi}).$$

To understand this expression, note that (i) a worker who has x competitors at D gets the offer from D with probability $\frac{1}{x+1}$ and (ii) the number of competitors faced by a worker who applies to D is Poisson (ξ). Similarly, the probability that an application to an N vacancy leads to an offer is

$$q = \sum_{r=0}^{\infty} \frac{1}{x+1} \frac{e^{-a/\theta} (\frac{a}{\theta})^x}{x!} = \frac{\theta}{a} (1 - e^{-a/\theta}).$$

Note that q was also defined in the discussion following Proposition 1 and does not depend on w'.

We now develop the indifference condition, which defines ξ as a function of w' given w and θ . Each worker must be indifferent between sending all a applications to N vacancies versus sending 1 application to D and the other a-1 to N vacancies. The expected payoff from sending all applications to N vacancies depends on neither ξ nor w' and can thus be treated as a constant. The expected payoff from sending one application to D and the others to N vacancies does, of course, depend on ξ and w'.

The possible payoffs for a worker who sends 1 application to D and the other a-1 applications to N vacancies are

(i) 1 if 2 or more applications are accepted. This occurs with probability

$$q^{D}(1-(1-q)^{a-1}) + (1-q^{D})(1-(1-q)^{a-1} - (a-1)q(1-q)^{a-2})$$

$$= 1 - (1-q)^{a-1} - (1-q^{D})(a-1)q(1-q)^{a-2}.$$

- (ii) w' if only the application to D is successful. This occurs with probability $q^D(1-q)^{a-1}$.
- (iii) w if the application to D is unsuccessful and only one application to N is successful.

This occurs with probability $(1 - q^D)(a - 1)q(1 - q)^{a-2}$.

(iv) 0 if no applications are successful.

This occurs with probability $(1-q^D)(1-q)^{a-1}$.

The expected payoff for a worker who sends 1 application to D and a-1 to N is thus

$$1 - (1-q)^{a-1} - (1-q^D)(a-1)q(1-q)^{a-2} + w'q^D(1-q)^{a-1} + w(1-q^D)(a-1)q(1-q)^{a-2}.$$

Equating the two expected payoffs implicitly defines $\xi(w'; w, \theta)$. Differentiating with respect to w', taking into account that $\frac{dq^D}{d\xi} = -\frac{1-e^{-\xi}-\xi e^{-\xi}}{\xi^2}$, and substituting for q^D and q gives

$$\frac{d\xi}{dw'} = \frac{\xi(1-e^{-\xi})(1-\frac{\theta}{a}(1-e^{-a/\theta}))}{(1-e^{-\xi}-\xi e^{-\xi})\left((a-1)\frac{\theta}{a}(1-e^{-a/\theta})(1-w)+w'(1-\frac{\theta}{a}(1-e^{-a/\theta}))\right)}$$

Since $1 - e^{-x} - xe^{-x} > 0$ for all x > 0 and $1 \ge w$, we have $\frac{d\xi}{dw'} > 0$ (as expected) and $\frac{d^2\xi}{dw'^2} < 0$.

Turning back to D's optimization problem, $\pi(w'; w)$ is proportional to $(1 - w')(1 - e^{-\xi})$. Maximizing with respect to w', the first-order (Kuhn-Tucker) condition is

$$-(1 - e^{-\xi}) + (1 - w')e^{-\xi} \frac{d\xi}{dw'} \le 0$$
 with equality if $w' > 0$.

If there is an interior solution, the second-order condition holds.

We are interested in the possibility of an interior solution at w' = w. Consider first the case of a = 1. If w' = w, then $\xi = 1/\theta$. Substituting and solving gives

$$w(\theta; 1) = \frac{e^{-1/\theta}}{\theta(1 - e^{-1/\theta})}.$$

Next consider the case of $a\in\{2,...,A\}$. Substituting the expression for $\frac{d\xi}{dw'}$ into the Kuhn-Tucker condition and evaluating at w'=w, where $\xi=a/\theta$, gives

$$\frac{(1-w)\xi e^{-\xi}(1-\frac{1}{\xi}(1-e^{-\xi}))}{(1-e^{-\xi}-\xi e^{-\xi})\left((a-1)\frac{1}{\xi}(1-e^{-\xi})(1-w)+w(1-\frac{1}{\xi}(1-e^{-\xi}))\right)} \le 1$$

This can be rewritten as

$$(1-w)e^{-\xi}(\xi^2-\xi(1-e^{-\xi})) \leq \left(1-e^{-\xi}-\xi e^{-\xi}\right) \left((a-1)(1-e^{-\xi})(1-w) + w(\xi-(1-e^{-\xi}))\right),$$

or

$$\frac{\xi^2 e^{-\xi} + (a-2)\xi e^{-\xi}(1-e^{-\xi}) - (a-1)^2(1-e^{-\xi})^2}{(1-e^{-\xi})} \leq w(\xi - a(1-e^{-\xi}) + (a-1)^2\xi(1-e^{-\xi})^2).$$

Only a corner solution exists with $w(\theta; a) = 0$ if this is a strict inequality.

To show that this inequality is in fact strict when $a \in \{2, ..., A\}$, we show that the RHS of the above expression is positive for any w > 0 and for all $\xi = a/\theta > 0$, while the LHS is negative. Note first that as $\xi \to 0$, the $RHS \to 0$ and, using a L'Hôpital's Rule argument, so does the LHS. Then note that

$$\frac{dRHS}{d\xi} = w(1 - ae^{-\xi} + (a - 1)^2(1 - e^{-\xi})^2 + 2(a - 1)^2\xi(1 - e^{-\xi})e^{-\xi}) > 0,$$

while

$$\frac{dLHS}{d\xi} = \frac{-e^{-\xi}((1-e^{-\xi})^2((a-1)(a-2) + \xi(a-2)) + (1-e^{-\xi} - \xi)^2}{(1-e^{-\xi})^2},$$

which is negative for $a \in \{2, ..., A\}$. Thus, in this case, we have a corner solution with $w(\theta; a) = 0$.

Finally to derive $\gamma(\theta;a)$, note that in symmetric equilibrium $q^D \equiv q = \frac{\theta}{a}(1-e^{-a/\theta})$. A fraction $1-(1-q)^a$ of all workers get a job. A fraction $1-(1-q)^a-a(1-q)^{a-1}$ of all workers receive multiple offers. Thus, a fraction

$$\frac{1 - (1 - q)^a - a(1 - q)^{a-1}}{1 - (1 - q)^a}$$

of the workers who find a job receive the competitive wage. Substituting for q gives equation (4). QED