Economics
Working Papers

2018-10

Complementarity and Advantage in the Competing Auctions of Skills

Alex Xi He, John Kennes and Daniel le Maire

Abstract:
We use a directed search model to develop estimation procedures for the identifi cation of worker and rm rankings from labor market data. These methods allow for a general speci cation of production complementarities and the possibility that higher ranked workers are not more productive in all fi rms. We also offer conditions for a positive/negative assortative matching that incorporate the possibility of a stochastic job ladder with on-the-job search. Numerical simulations relate the implications of the model to the implications of fi xed effect regressions and give further insights into the performance of our estimation procedures. Finally, we evaluate evidence for Denmark using our methods and we show that workers are highly sorted and that higher type workers are less productive than lower type workers while employed in lower type jobs.

AARHUS BSS
DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

Complementarity and Advantage in the Competing Auctions of Skills*

Alex Xi He, John Kennes and Daniel le Maire†

November 30, 2018

Abstract

We use a directed search model to develop estimation procedures for the identification of worker and firm rankings from labor market data. These methods allow for a general specification of production complementarities and the possibility that higher ranked workers are not more productive in all firms. We also offer conditions for a positive/negative assortative matching that incorporate the possibility of a stochastic job ladder with on-the-job search. Numerical simulations relate the implications of the model to the implications of fixed effect regressions and give further insights into the performance of our estimation procedures. Finally, we evaluate evidence for Denmark using our methods and we show that workers are highly sorted and that higher type workers are less productive than lower type workers while employed in lower type jobs.

Keywords: Directed search, sorting, wage dynamics, auctions.

JEL Classification: J64; J63; E32

*We wish to thank Preston McAfee, Giuseppe Moscarini, Dale Mortensen, Robert Shimer, and Randall Wright for helpful comments on earlier drafts. We thank the DTMC and the Danish National Research Foundation for financial support.

†Alex Xi He, Department of Economics, MIT, John Kennes, Department of Economics, Aarhus University, Denmark; Email: jkennes@econ.au.dk. Daniel le Maire, Department of Economics, University of Copenhagen, Denmark; Email: Daniel.le.Maire@econ.ku.dk

1 Introduction

In directed search models, the gains from trade within relationships are anticipated in advance of meetings and these opportunities direct search and help to determine who meets whom. For example, if workers and firms in the labor market know in advance of meetings that there are strong complementarities between particular types of workers and firms, directed search models predict that meetings will not be random (Acemoglu (2001), Eeckhout and Kircher (2010), Wright et al. (2018)). Directed search models also predict that productive matches can take time to form and that not all meetings yield matches. If this is the consequence of coordination frictions, Shimer (2005) and Shi (2001) demonstrate that matching will not only be assortative, but will also feature dispersion in the matching sets for individual agents types on each side of a market. Therefore, unlike the competitive benchmark model of sorting by Becker (1973), directed search models also predict that similar agents can expect different types of match partners over time.

Perhaps the simplest model of directed search with coordination frictions is McAfee's (1993) model of competing auctions. The simplicity of the equilibrium pricing mechanism – second price auctions with a reserve price equal to the seller's continuation value – is due to the fact that sellers compete on reserve prices and that the market utility of potential buyers in a large market is independent of the decision of each seller. Furthermore, if sellers are heterogenous, the distribution of buyer types bidding at each seller will depend on the seller's type. Therefore, in a labor context, the bidding game is simple Bertrand competition as in the Postel-Vinay and Robin (2002) model of employer competition for worker services, but in the competing auction model the distribution of offers facing each worker is endogenously determined and possibly different for each worker type. Another crucial difference between these two models is the assumption in McAfee (1993) that time is discrete. In this case, multilateral meetings are also possible for unmatched sellers.

The contribution of this paper is to use the directed search model of McAfee (1993) to develop methods to identify the ranking of workers and firms from labor market data. These methods seek to identify if there are complementarities between different types of workers and firms and whether the most advantaged workers are more productive in all firm types. We also use the model to characterize sorting and to give conditions for positive and negative assortative matching (PAM/NAM). Our methods can be derived from a simple static model. However, we also characterize a more complicated dynamic model where workers can climb a job ladder. The job ladder model is useful for understanding qualititative differences between the predictions of directed and undirected search models (Refer to Shimer and Smith (2000) and Postel-Vinay and Robin (2002)).

Our directed search model predicts that the equilibrium wage function is always monotonic in firm productivity.[1] This simple classification of the firm types is consistent with the assumption underlying fixed effect regressions, such as Abowd, Kramarz and Margolis (1999) (AKM), that wages are monotone in firm's productivity.[2] Random matching models do not offer a robust version of this prediction. For example, in Shimer and Smith (2000), wages are non-monotonic in firm productivity, because higher firm types will agree to hire a relatively unproductive worker only if this worker accepts a sufficiently low wage to compensate the firm for the option value of waiting for a more productive potential hire.[3] This option value does not apply to the competing auction environment, because a lower worker type offers an alternative form of compensation - an offer of a higher probability of trade at the stage where the firm chooses to approach a particular worker type. Wages are also not monotonic in firm types in Postel-Vinay and Robin's (2002) model of random matching without capacity constraints, because workers in high type firms may initially accept low wages in return for the future wage increases that occur by the arrival of future counter-offers from inferior firms. This second mechanism is also shut down in the competing auction environment, because the matching of firms to workers is directed. Therefore, on-the-job searchers can anticipate superior offers from better firms, but not bargaining power enhancing offers from worse firms.[4]

Our model does not predict that higher type workers are always paid better wages by every employer. One obvious reason is our assumption that higher worker types do not have global absolute advantage in all job types. Another reason is due to directed meetings. For example, an unmatched low type worker might expect many offers from low type jobs and few offers from medium job types while an unmatched high type worker might expect few offers from medium type jobs but many offers from high type jobs and no offers from low type firms. In this case, the high type worker will also expect to earn a lower wage than the low type worker in the medium type job. Certain random matching models, by contrast, predict that worker types can be ordered by their expected wage at each firm type. However, this prediction is driven by the assumption of a meeting function whereby all workers are given the same frequency of meetings.[5]

Since we do not assume global absolute advantage, we must address the crucial question

[1] The directed search model of Shimer (2005) does not feature this prediction, because some worker types may have higher probabilities of trade at higher type firms than at lower type firms. See Abowd, Kramarz, Perez-Duarte and Schmutte (2018) on the problem of empirical identification in Shimer's (2005) model.

[2] The AKM model assumes, more strongly, an additive separable firm fixed effect.

[3] See Hagedorn, Law and Manovskii (2017), for example.

[4] Bagger and Lentz (2018) consider a random matching model with no capacity constraints and where workers have relatively high bargaining power, which reduces the impact of future inferior counter-offers on equilibrium wages. Simulations show that wages are approximately monotonic in firm productivity.

[5] See Hagedorn, Law and Manovskii (2017) and Lopes de Melo (2018).

of how to formally rank different worker types. We rank workers by their expected market utility of engaging with the competing auction mechanism in the unmatched state. This method of ranking workers is essential for our analysis of higher and lower worker types in the competing auction environment, because the matching set of a higher type worker can generally include some job types, which are also in the matching set of lower type workers, and for which the lower type workers are more productive than higher type workers. Our method of ranking workers is also similar to how we rank firms. However, in the case of firms, our model predicts that vacancies with higher opportunity costs are also more productive when matched to any particular worker. This prediction follows from our assumption that each unmatched job is directed to a particular worker sub market. Therefore, if a higher opportunity cost firm has a lower productivity in a sub market than another firm type with a lower opportunity cost, the higher opportunity cost firm will not enter.

We give a characterization of PAM and NAM.[6] One complication regarding the characterization of assortative matching, which is shared by other stochastic matching models, is the equilibrium prediction that the matching sets of each worker type and each firm type are generally not singletons.[7] This is an issue in our model, because different worker types will generally have overlapping matching sets and these different worker types will also face a unique frequency of job offers within each of these matching sets. Therefore, we follow Lentz (2010) and define PAM/NAM by the condition that the offer distribution of a higher/lower worker types stochastically dominates the offer distribution of a lower/higher worker types. We also require, as in Shimer and Smith (2000), that the matching sets of the higher/lower worker types under PAM/NAM must dominate the matching sets of the lower/higher worker types.

We establish sufficient conditions for PAM/NAM in the competing auction equilibrium. We first derive the sufficient conditions on the production function between workers and firms in the base game of the competing auction model. We then extend this result to the dynamic stage game of the competing auction model where the sufficient condition for PAM/NAM then refers to the required properties associated with the equilibrium value function of a match. Interestingly, the sufficient conditions for PAM/NAM are stronger than super/submodularity of the production function. Instead, we offer an alternative sufficient condition for PAM/NAM, which we refer to as positive non-decreasing/negative non-increasing relative complementarity. This assumption requires that the complementarities

[6]Our analysis of sorting with directed search differs from Eeckhout and Kircher (2010) in two important respects: we assume meetings are multilateral instead of bilateral and we assume the continuation value of a worker is positive. The first assumption is premised on the existence of coordination frictions as in Shimer (2005). The latter assumption is crucial for the extension of our results to a dynamic environment.

[7]This is contrasted by Becker's (1973) seminal model of frictionless sorting.

in the value of the match are increasing/decreasing in firm type relative to the change in the marginal value of the match with respect to higher firm types.

We apply the competing auction model of sorting to the problem of determining a set of empirical methods, which can be used to identify the strength and sign of sorting, and a test of the assumption of global absolute advantage.[8] A key step in the development of the sign of sorting test is a prediction of the competing auction model concerning the mapping between an unmatched worker's expected utility (which is their ranking) and the distribution of wage offers that each unmatched worker type can expect to earn if they are hired by the highest ranking firm, which is within their matching set, out of unemployment. In this way, our identification of the sign of sorting is related to Bagger and Lentz (2018).

The assumption of no global absolute advantage is related to the recent literature on multidimensional skills and sorting (Lindenlaub (2017), Lise and Postel-Vinay (2016) and Lindenlaub and Postel-Vinay (2017)). In these models, no global absolute advantage arises because workers have different skill bundles and jobs have different skill requirements. This literature focuses on sorting within each skill, whereas we assume only a one-dimensional worker type that determines each worker's market utility. Our simple test for global absolute advantage exploits a crucial feature of our assumed local Bertrand competition for the worker's services. Namely, that the highest wage offered by each firm type to each worker type will always equal the worker's productivity, because the worker's wage equals productivity whenever the firm faces a rival offer from a similar type firm.

We use our empirical identification to analyze sorting and absolute advantage in the Danish labor market using register data for Denmark for 1995-2011. We find evidence of positive sorting and a strength of sorting of 0.3, which is in similar to the results of Bagger and Lentz (2018) and Lentz et al. (2018) for Denmark. We also reject the hypothesis of global absolute advantage of higher worker types.

The paper is organized as follows. In section 2, we first characterize a simple static model of competing auction with heterogenous workers and firms. In section 3, we use the results of the static model to characterize the solution of the stage game of an infinitely repeated competing auction game with random exogenous separations at the end of each period and time discounting. In section 4, we lay out the empirical identification strategy. We follow this with numerical simulations in section 5. We then use our methods of identification to infer the pattern of sorting and production using Danish data on wages and worker to firm transitions in section 6. The final section concludes the paper.

[8] The problem of identifying the sign of sorting owes much to the original discussions by Eeckhout and Kircher (2011) and Gautier and Teulings (2006). See also Bagger and Lentz (2018), Bartolucci, Devicienti and Monzón (2018), Hagedorn, Law and Manovskii (2017), Lise, Meghir and Robin (2016), and Lopes de Melo (2018).

2 The Static Model

In this section, we extend McAfee's (1993) competing auction model to an environment with many ***heterogenous*** competing sellers (workers) and heterogenous bidders (firms). We maintain the same assumptions as McAfee (1993) about the private information of buyer types, the public information of seller types, and the set of possible selling mechanisms used in each matching round.[9]

2.1 Environment

The players in a matching market consist of a continuum of workers and firms. The workers are risk neutral, expected utility maximizers and the firms seek to maximize expected profits. The workers are divided into types $h \in [0, \infty)$ with a fixed population of each type of worker equal to $n(h)$. The total size of the worker population is fixed and normalized to one. The firms are divided into types $k \in [0, \infty)$. If a type k firm employs a type h worker, the output of the match is determined by the production function, $y(k,h)$. If a worker is unmatched, the worker produces output $y(\underline{k}, h)$. We use $n(\underline{k}, h) = n(h)$ to denote a submarket consisting of all type h workers with home production $\underline{k}$. The opportunity cost of entry by a type k firm into any submarket is $c(k) = k$, which we normalize to be linear. We let $\phi(k \mid \underline{k}, h)\, n(\underline{k}, h)$ denote the number of firm types greater than type k who enter submarket $n(\underline{k}, h)$. We assume that $y(k,h) - y(\underline{k}, h) - k > 0$ for some $k > \underline{k}$. Finally, we assume that the production function $y(k,h)$ is increasing and concave with respect to firm type.

2.2 The competing auction game

The competing auction game has three stages. In the first stage of the game, firms choose to enter a submarket and each worker in the submarket then advertises a second price auction with a reserve price equal to the worker's continuation value, $y(\underline{k}, h)$. In the second stage of the game, there is initially $\phi(k \mid \underline{k}, h)\, n(\underline{k}, h)$ unmatched firms of type greater than $k \in [0, \infty)$ in submarket $n(\underline{k}, h)$. We assume a mixed strategy equilibrium where each of these heterogenous firm types play symmetric mixed strategies regarding their assignment to any worker in this submarket. Therefore, the distribution of firm types at each worker's

[9] A simplifying assumption of this model is to assume that buyers choose a submarket before the sellers compete in their choice over general direct revelation mechanisms. In this case, the economic environment is equivalent to McAfee (1993) because all sellers in a submarket are identical. We then borrow, without proving theorems 1 and 2 of McAfee (1993), that each seller will maximize his expected utility by the advertisement of a simple second price auction and that the reserve price of this auction is equal to the seller's continuation value.

auction will be determined by the Poisson distribution with parameter, $\phi(k \mid \underline{k}, h)$. By the theory of the Poisson distribution, the probability that the highest bidder in the auction for a type h worker with outside option $\underline{k}$ is a type k firm or below is given by

$$G^1(k \mid \underline{k}, h) = \exp(-\phi(k \mid \underline{k}, h)) \quad (1)$$

Similarly, the probability that the second highest bidder in the auction for a type h worker is a type k firm or below is given by

$$G^2(k \mid \underline{k}, h) = \exp(-\phi(k \mid \underline{k}, h)) + \phi(k \mid \underline{k}, h)\exp(-\phi(k \mid \underline{k}, h)). \quad (2)$$

We use $k^*(\underline{k}, h)$ and $\hat{k}(\underline{k}, h)$ to denote the highest and lowest type firms that approach workers in the type h worker submarket. We also use $G^1\left(\hat{k}(\underline{k}, h) \mid \underline{k}, h\right)$ to denote the probability that the worker gets no offer.

In the third stage of the game, the worker obtains a wage for his labor services by conducting his second price auction.[10] The auction awards the worker's services to the highest valuation firm that exceeds the worker's reservation value at a wage equal to the valuation of the second highest bidder. Letting k_1 and k_2 denote the highest and second highest firm type bidding for the worker's services, the wage earned by a worker is then given by,

$$w(k_1, k_2 \mid h, \underline{k}) = \begin{cases} y(k_2, h) & \text{if} \quad y(k_2, h) > y(\underline{k}, h) \\ y(\underline{k}, h) & \text{if} \quad \text{otherwise} \end{cases}, \quad (3)$$

Therefore, if $k_1 > \underline{k}$, the highest valuation firm hires the worker and earns $y(k_1, h) - w(k_1, k_2 \mid h, \underline{k})$ while all other firms bidding for this worker earn zero. If $k_1 \leq \underline{k}$, all firms earn zero revenue and the worker earns his continuation value, $y(\underline{k}, h)$. Applying equations (1) and (3), the expected profit for a type k firm entering submarket $n(\underline{k}, h)$ is given by[11]

$$\pi(k \mid \underline{k}, h) = \int_{\underline{k}}^{k} (y(k, h) - y(z, h))\, dG^1(z \mid \underline{k}, h) - k \quad (4)$$

The free entry condition is that a firm will enter if $\pi(k \mid \underline{k}, h)$ is positive and stay out of the

[10] By revenue equivalence, McAfee's (1993) competing auction equilibrium also includes the possibility of first price auctions. However, if sellers use first price auctions, the bidding distribution will be similar to that predicted by a buyer price posting model as in Burdett-Judd model (1983). Furthermore, if the number of bidder types is continuous, the buyer posting model will predict an identical distribution of prices as the seller price posting model of Shimer (2005). These results are explained in Kennes, le Maire and Roelsgaard (2018).

[11] Since the worker has a positive probability of no offer, the expected profit of the firm in terms of the density of highest competing offers should be written as $\pi(k \mid \underline{k}, h) = (y(k, h) - y(\underline{k}, h))\, G^1\left(\hat{k} \mid \underline{k}, h\right) + \int_{\hat{k}(\underline{k}, h)}^{k} (y(k, h) - y(z, h))\, dG^1(z \mid \underline{k}, h) - k$

submarket if $\pi(k \mid \underline{k}, h)$ is negative. The expected market utility of the worker in the static game is simply the expectation of their wage. Since workers never accept offers from job types below $\underline{k}$, the expected market utility of a worker with home production technology $\underline{k}$, which is both their best and second best current opportunity, is given by

$$W(\underline{k}, \underline{k}, h) = \int_{\underline{k}}^{k^*(h)} y(z, h) \, dG^2(z \mid \underline{k}, h) \tag{5}$$

where $dG^2(k \mid h)$ assigns the probability density that the worker's second highest bidder is type z and the bidding function gives the worker a payoff of $y(z, h)$ in such an event.

2.3 Equilibrium

For a set of worker submarkets summarized by the populations, $n(\underline{k}, h)$, an equilibrium is a list $\phi(k \mid \underline{k}, h)$, $k^*(\underline{k}, h)$, $\hat{k}(\underline{k}, h)$, $\pi(k \mid \underline{k}, h)$, and $W(\underline{k}, \underline{k}, h)$, which satisfies the free entry condition of each firm type $k \in [0, \infty)$ and equations (1), (2), (3), (4), (5). The following proposition characterizes the equilibrium solution.

Proposition 1. *In the static game, given any distribution of workers across $n(h, \underline{k})$ submarkets, the equilibrium exists, is unique, and*

1. The matching set upper bound $k^*(\underline{k}, h)$ is the solution to

$$y_1(k^*, h) = 1, \tag{6}$$

2. The matching set lower bound$\hat{k}(\underline{k}, h)$ is the unique solution

$$\hat{k}(\underline{k}, h) = \arg\max_k \left(\left(y(k, h) - y(\underline{k}, h) \right) / k \right), \tag{7}$$

3. The expected number of firms above k is given by $n(\underline{k}, h) \phi(k \mid \underline{k}, h) = n(\underline{k}, h) \phi(k \mid h)$ where

$$\phi(k \mid h) = \log\left(y_1(k, h)\right), \tag{8}$$

4. The expected profits of entering firms are zero and the expected market utility of each worker is given by equation (5).

Proof. See appendix. □

From equations (6) and (8) the functions $\phi(k \mid \underline{k}, h)$ and $k^*(h, \underline{k})$ are both independent of $\underline{k}$. Therefore, if two workers are otherwise identical but one worker has a higher outside opportunity, the only difference in their job offer function is the lowest entrant bidder type $\hat{k}(\underline{k}, h)$. A higher worker outside opportunity, $\underline{k}$, gives reduced incentives for entry by prospective employers. Thus the worker will get fewer bidders. However, this reduction in bidders happens at the bottom end of the job offer distribution leaving unchanged incentives for entry at the top of the job offer distribution.[12] This result greatly simplifies the analysis of the dynamic matching model where the static game described here is used as a stage game in the dynamic model. Note also that $\hat{k}(\underline{k}, h)$ is an increasing function of the workers outside option, $\hat{k}_1(\underline{k}, h) > 0$, and that the output of working at a type $\hat{k}(\underline{k}, h)$ firm is strictly greater than the value of the outside option, $\hat{k}(\underline{k}, h) > \underline{k}$.

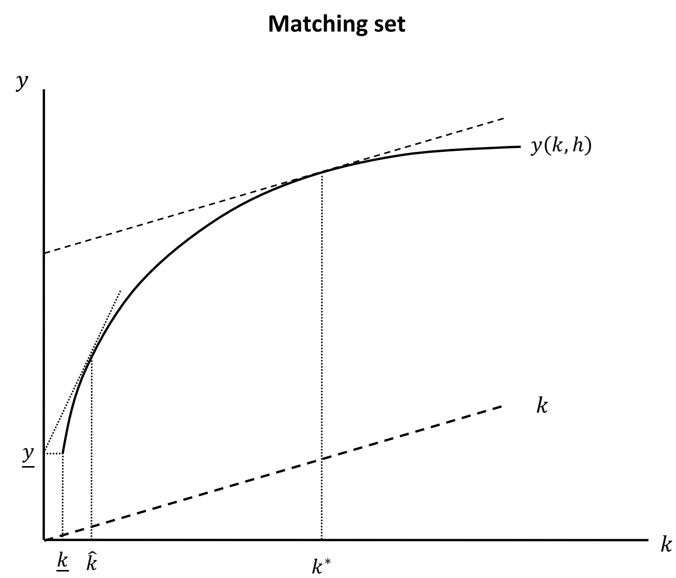


Figure 1: Worker production function and matching set

The lower and upper bounds of the worker's matching set are illustrated in figure 1. The equilibrium of this labor market departs from the Walrasian market solution because not all firms will offer $k^*(h)$ jobs. This is because coordination frictions can lead to more than one firm bidding for the same worker. Given that there are fewer $k^*(h)$ firms than what characterizes the Walrasian solution, there are opportunities for lower cost firms to

[12] Albrecht, Gautier and Vroman (2014) relate this 'offsetting business stealing' effect to the efficiency of the assignment. We extend this result to matching environments where seller heterogeneity includes differences other than differences in outside options. In Kennes and le Maire (2010), the result is derived as a decentralized competing auction game with a continuum of firm types and contrasted this solution to an equilibrium with a discrete number of firm types.

approach workers. The key condition for more than one job type to be offered in equilibrium is that there is a comparative advantage of low type jobs over high type jobs ($y(k,h)/k > y(k^*(h),h)/k^*(h)$) and an absolute advantage of the $k^*(h)$ job over low type jobs ($y(k^*(h),h) - k^*(h) > y(k,h) - k$).[13]

In each submarket, there is a trade-off between a higher type firm being more productive and having a higher probability of hiring a worker and the higher cost of participation. Equation (8) says that more firms will locate where the curvature of the production function is high since this means that the production value of having a slightly higher k is increasing less. We let $p^*(h)$ denote the value of $\hat{k}(p^*(h),h) = k^*(h)$. If $\underline{k} \geq p^*(h)$ the worker gets no offers.

2.4 Advantage

Firms are ranked by the market utility needed to ensure their entry to the competing auction mechanism. Therefore, a type k firm is ranked higher than a type k' firm if $k > k'$, because making a higher investment, the higher firm type must have higher expected returns. Otherwise, by the free entry condition, a k firm will never enter a submarket $n(h)$ where $y(k,h) < y(k',h)$. This implies that global absolute advantage of higher type firms is an equilibrium outcome.

The workers are also ranked by the market utility $W(\underline{k},\underline{k},h)$ which ensures their participation in the competing auction mechanism. Formally, workers are ranked according to the following definition.

Definition 1. If two workers h and h' are allocated the (home) production technology, $\underline{k}$, then worker h is said to be ranked higher than h' if the expected market utility of worker h is higher than worker h', i.e. $W(\underline{k},\underline{k},h) > W(\underline{k},\underline{k},h')$.

This market utility ranking of workers allows for the possibility that changes in the production function with respect to higher worker types, $y_2(k,h)$, may not be positive for all k. For example, a type h worker might be more productive in type a k job than a type $h' < h$ worker, while the type h' worker might be more productive in the type $k' < k$ job. For example, a professional engineer has higher expected earnings in the market place for skills than a less qualified service technician. However, the service technician may be more productive than the professional engineer in one of the job types. The market utility ranking captures whether the type h worker is more valued by the market and thus expects higher earnings.

[13] See also Julien, Kennes and King (2006). Shimer (2005) offers a related discussion for a posted price model of coordination frictions.

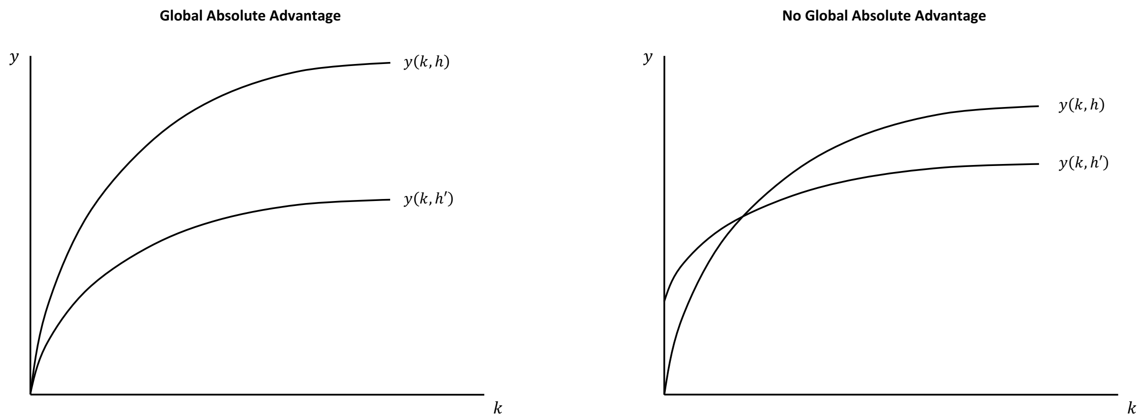


Figure 2: Technologies with and without global absolute advantage

The left panel of figure 2 illustrates the case where we assume that the type h worker is more productive in all jobs compared to a type h' worker and where the difference in productivity is increasing in firm types. This is a case of global absolute advantage. The second example illustrates the case, which we think is most relevant to a realistic description of the labor market, where neither worker h nor h' has global absolute advantage since the type h worker is more productive in high type jobs, whereas the type h' worker is more productive in the lower job types. The case without global absolute advantage allows for strong complementarities between high type workers and high type jobs. Moreover, this assumption also allows for a strong advantage in the expected earnings of high type workers.

We also note that the wage of the worker employed by the highest rank firm is given by $w\left(k^*\left(h\right), z, h\right) = y\left(z, h\right)$ where z is the worker's second highest offer. Integrating this wage over possible realizations of z also gives the worker's expected value. We have

$$\begin{aligned} W\left(\underline{k}, \underline{k}, h\right) &= \int_{\underline{k}}^{k^*(h)} y\left(z, h\right) dG^2\left(z \mid h\right) \\ &= \int_{\underline{k}}^{k^*(h)} w\left(k^*\left(h\right), z, h\right) dG^2\left(z \mid h\right) \end{aligned}$$

We will offer an extension of this result in the dynamic analysis and use it as means to rank workers, which utilizes information on the distribution of wages at the highest ranking firm type within each worker's matching set.

An interesting special case of global absolute advantage is the modular case where the difference in productivity between two workers over all firm types is constant. In this case, for two workers h' and $h < h'$, we have for all k that

$$y(k,h') = y(k,h) + A$$

where A is a scalar. The modular case is a useful benchmark since proposition 1 implies that the two workers h and h' have identical matching sets and identical distributions of job offers, i.e. $\hat{k}(k,h) = \hat{k}(k,h')$, $k^*(h) = k^*(h')$, and $G^1(k \mid h) = G^1(k \mid h')$ for all k. Therefore, in the special case of modular differences the assignment of jobs to workers will be random. Referring to the bidding function in equation (3), the workers' wages are simply scaled by the constant difference in worker productivities. Therefore, there is no advantage for any firm type to approach any worker type with a different frequency. However, if technology is not modular, the assignment will be directed.

2.5 Sorting

The GM game provides a model of the assignment of heterogeneous workers to heterogeneous firms. This section considers the restrictions on the technology that delivers either positive or negative assortative matching (PAM/ NAM) of workers to firms. Since the assignment is directed, the sorting will depend both on differences in the matching sets and differences in the expected relative frequency of jobs offers within overlapping matching sets. This is unlike random search models such as Shimer and Smith (2000), where two different workers have the same expected offer arrival rates for firms in the intersection of their matching sets. We use the following definition of sorting:

Definition 2. The conditions for positive/ negative assortative matching (PAM/ NAM) are:

1. The lower and upper bounds of the matching sets are weakly increasing/ decreasing in worker type.
2. First-order stochastic dominance of higher/ lower type workers in overlapping match sets.

For overlapping matching sets, it is easy to show that when the production function is super/ submodular, the assignment of higher/lower type workers stochastically dominates the assignment of lower/higher type workers. For example, if worker $h' > h$ and the production function $y(k,h)$ is supermodular, then $G^1(k \mid h') < G^1(k \mid h)$ for all k in the matching sets. This also implies that the higher type worker will be less likely to become unemployed. Furthermore, it is clear from equation (6) that supermodularity of the production function implies that the upper bound of the matching set, $k^*(h)$, is increasing in worker type. However, to establish PAM/ NAM we also have to ensure that the lower bound of the matching set, $\hat{k}(\underline{k}, h)$, is weakly increasing/ decreasing in worker type. This requires a stronger condition than supermodularity. The reason is that a higher degree of complementarity means more competition from higher type firms.

If we - similar Shimer and Smith (2000) - assume that the value of home production equals zero, $y(\underline{k}, h) = 0$, for all worker types, the proof of proposition 2 shows that the sufficient condition for PAM/ NAM is log-supermodularity/log-submodularity of the production function. However, we notice that under the assumptions of supermodularity and $y(\underline{k}, h) = 0$ for all h, we effectively assume global absolute advantage.

For the general case where we do not restrict the value of home production to be zero, we need another conditon on teh production technology.

Definition 3. The production function has *positive non-decreasing/negative non-increasing relative complementarity* (PRC/NRC) if $y(k, h)$ is super/submodular and the ratio $\frac{y_{12}(k,h)}{y_1(k,h)}$ is non-decreasing/non-increasing.

The proposition below provides the sufficient condition for PAM/ NAM and for the lower bound of the matching set to be increasing/ decreasing in the worker type.

Proposition 2. *A sufficient condition for PAM/ NAM is that production function $y(k, h)$ is PRC/NRC.*

Proof. See appendix. □

3 The Dynamic Model

The players in the matching market consist of a continuum of workers and firms. Time is infinite and discrete. All agents are infinitely lived with a common discount factor β. The workers are risk neutral, expected utility maximizers and the firms seek to maximize expected profits. The total size of the worker population is fixed and normalized to one. The workers are divided into types $h \in [0, \infty)$ with a fixed population of each type of workers equal to $n(h)$. The firms are divided into types $k \in [0, \infty)$. If a type k firm employs a type h worker, the output of the match is determined by the production function, $y(k, h)$. If a worker is unmatched, the worker produces output $y(\underline{k}, h)$. The opportunity cost of entry for a type k firm is exogenous and normalized to be linear $c(k) = k$. The population of each type of firm is determined by free entry. At the end of the period all matches are subject to an exogenous rate of destruction δ.

3.1 Regularity assumptions

We let the ***job value function*** $\Lambda(k,h)$ denote the combined present value of a match involving a type h worker and a type k firm at the time at which the firm negotiates a wage with the worker. We assume that this value function satisfies the following regularity conditions.

Definition 4. A ***job value function*** satisfies the regularity conditions if

$$\begin{aligned} \Lambda(k,h) - k &\geq 0 \text{ for some value of } k \geq 0 \\ \Lambda_1(k,h) &> 0 \\ \Lambda_{11}(k,h) &< 0 \end{aligned}$$

The regularity condition is posited as a convenient means to solve the equilibrium of the model by a block recursive system of equations.[14] When we solve the stage game for job entry using a job value function that satisfies the regularity condition, we will show that there exists a unique equilibrium solution for job entry for any initial distribution of workers into submarkets. We will also find that all possible equilibrium solutions for job entry can be attained by an appropriate specification of the job value function.

3.2 The stage game

At the start of each period there is an initial distribution of employed and unemployed workers. A submarket is a population of workers $n(k_1,k_2,h)$ where h is the worker's type, k_1 is the highest current bidder of the worker's services and k_2 is the second highest bidder. The assignment of new entrant firms into these submarkets is described by a three stage game, which is analogous to the static game of the previous section.

In the first stage, unmatched firms choose to enter a submarket $n(k_1,k_2,h)$ and each matched-worker pair in this submarket advertises a second price auction with a continuation value equal to $\mathit{\Lambda}(k_1,h)$.

In the second stage, there are $\phi(k \mid k_1,h)\, n(k_1,k_2,h)$ firm types greater than $k \in [0,\infty)$ in submarket $n(k_1,k_2,h)$. We assume a mixed strategy equilibrium such that each of group of unmatched heterogenous firm types play a symmetric mixed strategy regarding their assignment to any particular worker in this submarket. Therefore, the distribution of unmatched firm types bidding at each worker's auction will be determined by the Poisson

[14] Similar to Menzio and Shi (2011) the equilibrium is block recursive as value and policy functions do not depend on the equilibrium allocation of workers across firms.

distribution with parameter $\phi(k \mid k_1, h)$. We use $k^*(k_1, h)$ and $\hat{k}(k_1, h)$ to denote the highest and lowest type firms that approach workers in the type h worker submarket. We also use $G^1\left(\hat{k}(k_1, h) \mid k_1, h\right) = \exp\left(-\phi\left(\hat{k}(k_1, h) \mid k_1, h\right)\right)$ to denote the probability that the type h worker gets no offer. We let k_1' and k_2' denote the first and second highest unmatched firm bidding for the worker of type h with current employer k_1. Since recruitment is directed, any offer received by the worker will be better than the current employment. If the worker gets only one offer, we set $k_2' = \underline{k}$, and if no offer we set $k_1' = \underline{k}$.

In the third stage of the game, the set of unmatched firms bid for the worker's services against the worker's current employer k_1 who currently offers the worker an expected return equal to $\Lambda(k_2, h)$. In this second price auction, the worker is awarded a surplus equal to the second highest valuation. Therefore, the auction awards a value $Z(k_1', k_2' \mid k_1, k_2, h)$ to the worker according to the following bidding function.

$$Z(k_1', k_2' \mid k_1, k_2, h) = \begin{cases} \Lambda(k_2, h) & \text{if} \quad k_1' = k_2' = \underline{k} \\ \Lambda(k_1, h) & \text{if} \quad k_1' > k_1 \text{ and } k_2' = \underline{k} \\ \Lambda(k_2', h) & \text{if} \quad k_1' > k_1 \text{ and } k_2' > k_1 \end{cases}, \tag{9}$$

We understand this bidding function as follows. The contract at a type k_1 firm is initially set by the second highest offer that the worker has entertained in the past, k_2. On top of this, when the worker starts each period, he/she faces the possibility of getting additional job offers. The bidding function considers three cases: 1) the event that the worker gets no offer, in which case k_2 stays unchanged; 2) the event that the worker gets one offer, in which case the wage contract is now set by a second highest offer equal to k_1 - the worker's current employer; and 3) the event that the worker gets multiple offers, in which case the contract is set by the job value of the second highest of these offers.

If $k_1' > \underline{k}$, the highest valuation firm hires the worker and earns the job value $\Lambda(k_1', h)$ less the value paid to the worker $Z(k_1', k_2' \mid k_1, k_2, h)$, which depends on k_1 and the realization of k_2'. The equation for the expected profits of a type k firm entering a type $n(k_1, k_2, h)$ submarket if found by integrating over possible values of k_2' which are less than k. We have

$$\pi(k \mid k_1, k_2, h) = (\Lambda(k, h) - \Lambda(k_1, h))\, G^1\left(\hat{k} \mid k_1, h\right) + \int_{\hat{k}(k_1, h)}^{k} (\Lambda(k, h) - \Lambda(z, h))\, dG^1(z \mid k_1, h) - k \tag{10}$$

where the function for $G^1(z \mid k_1, h)$ has the same form as (1). The free entry condition is that a firm will enter if $\pi(k \mid k_1, k_2, h)$ is positive and stay out of the submarket if $\pi(k \mid k_1, k_2, h)$ is negative. We use $k^*(k_1, h)$ and $\hat{k}(k_1, h)$ to denote the highest and lowest type firms that approach workers in the type h worker submarket. The worker's current employer gives the worker's continuation value when setting wages with any future employers contacted by on-the-job search. At the start of each period, the expected present value of a worker in a type k_1 job with a second best offer k_2 from the time of the contract is given by

$$\begin{aligned} W(k_1,k_2,h) &= \Lambda(k_2,h)\,G^1(\widehat{k}(k_1,h)\mid k_1,h)+\\ &\quad \Lambda(k_1,h)\left(G^2(\widehat{k}(k_1,h)\mid k_1,h)-G^1(\widehat{k}(k_1,h)\mid k_1,h)\right)+\\ &\quad \int_{z=\widehat{k}(k_1,h)}^{k^*(h)}\Lambda(z,h)\,dG^2(z\mid k_1,h) \end{aligned} \tag{11}$$

where the formulas for $G^1(k \mid k_1,h)$ and $G^2(k \mid k_1,h)$ have the same form as equations (1) and (2). Again, as with the bidding function in equation (9), the three cases represent the events that the worker gets zero, one or multiple offers.[15] The present value of an unmatched worker is given by $W(\underline{k},\underline{k},h)$.

3.3 Stage game equilibrium

Consider any initial distribution of workers into submarkets $n(k_1,k_2,h)$ and a job value function $\Lambda(k,h)$ that satisfies the regularity conditions. An equilibrium to the stage game is a list $\phi(k \mid k_1,h)$, $k^*(k_1,h)$, $\hat{k}(k_1,h)$, $\pi(k \mid k_1,h)$, and $W(k_1,k_2,h)$, which satisfies the free entry condition of each firm type $k \in [0,\infty)$ and equations (1), (2), (9), (10), (11). The following lemma characterizes the equilibrium solution.

Lemma 1. *In the stage game, for any distribution of workers across $n(k_1,k_2,h)$ submarkets, the equilibrium exists, is unique, and*

1. The matching set upper bound $k^*(k_1,h)$ is the solution to

$$\Lambda_1(k^*,h)=1,$$

2. The matching set lower bound $\hat{k}(k_1,h)$ is the unique solution

$$\hat{k}(k_1,h)=\arg\max_k\left(\left(\Lambda(k,h)-\Lambda(k_1,h)\right)/k\right),$$

3. The expected number of firms above k is given by $n(k_1,k_2,h)\,\phi(k \mid k_1,h)=n(k_1,k_2,h)\,\phi(k \mid h)$ where

$$\phi(k\mid h)=\log\left(\Lambda_1(k,h)\right),$$

[15] Note that the probability of receiving exactly one offer above $\hat{k}$ is given by $\left(1-G^1(\widehat{k}\mid h)\right)-\left(1-G^2(\widehat{k}\mid h)\right)=G^2(\widehat{k}\mid h)-G^1(\widehat{k}\mid h)$.

4. The expected profits of entering firms are zero and the expected market utility of each worker is given by equation (11).

Proof. The proof of items 1-3 is the same as the proof of proposition 1. The only step is to substitute the production function $y(k,h)$ in the proof for the value function $\Lambda(k,h)$ □

As in the static game the $k^*(k_1,h) = k^*(h)$ is independent of the current employment. We let $p^*(h)$ denote the value of $\hat{k}(p^*,h) = k^*(h)$. If the worker's current employer type is $k_1 \geq p^*(h)$, then the worker gets no new offers from rival firms as long as the current match continues. Otherwise, workers will continue to move up the job ladder as better offers appear and are accepted by the worker.

A potential question regarding the properties of the competing auction equilibrium is whether some of these properties are associated with an inefficient assignment of workers to firms. We show in the appendix that the equilibrium meeting function, $\phi(k \mid k_1,h)$, is constrained efficient. Furthermore, the auction mechanism ensures that highest valuation firms always employ a worker. Therefore, both meetings and matchings are constrained efficient.

3.4 Wages and the ranking of firms

The wage contract of each worker is solved as follows. If the worker enters the period with wage contract $w(k_1,k_2,h)$ and then the worker realizes new bid values, $\{k_1',k_2'\}$ the wage contract is then set to $w(k_1',k_2',h)$. The workers are then contracted to be employed at the wage of the winning bid until either the match terminates exogenously or the worker gets an additional job offer from another employer. Given that the wage contract is determined by auction, the present value of a worker with a type k_1 employer and a type k_2 second best offer is $\Lambda(k_2,h)$. Therefore, after firms are assigned and contracts are set by bidding, the present value of the worker with a wage contract $w(k_1,k_2,h)$ will satisfy the following asset equation,

$$\Lambda(k_2,h) = w(k_1,k_2,h) + \beta\left[(1-\delta)W(k_1,k_2,h) + \delta W(\underline{k},\underline{k},h)\right]$$

where the present value of rewards is given by $W(k_1,k_2,h)$ if the worker is not displaced and by $W(\underline{k},\underline{k},h)$ otherwise. Rearranging, we can express the equilibrium wage as,

$$w(k_1,k_2,h) = \Lambda(k_2,h) - \beta\left[(1-\delta)W(k_1,k_2,h) + \delta W(\underline{k},\underline{k},h)\right] \tag{12}$$

Since $W(k_1,k_2,h)$ and $W(\underline{k},\underline{k},h)$ are characterized by equation (11), which is a function of $\Lambda(k,h)$ by lemma 1, the equilibrium wage is solved as part of a block recursive solution of the equilibrium.

In general the second best offer of a worker at any particular job type is disperse. Therefore, by equation (12), an important implication of the competing auction model is that the wages of workers at each job type are also disperse.[16] If $k_1 = k_2$, the worker 'owns' the job and is given a wage equal to output. Therefore, we have

$$w(k,k,h) = y(k,h). \tag{13}$$

We can then use equation (12) to express the production function $y(k,h)$ as a function of job value function $\Lambda(k,h)$. We

$$y(k,h) = \Lambda(k,h) - \beta\left[(1-\delta)W(k,k,h) + \delta W(\underline{k},\underline{k},h)\right] \tag{14}$$

We derive the following results. Firstly, there exist a job type k such that $\Lambda(k,h) > k$ if we assume $(y(k,h) - y(\underline{k},h)) / (1-\beta(1-\delta)) > k$. This condition is needed to ensure that an equilibrium with positive job entry exists. As noted with the wage equation (12), the RHS of equation (14) is a function of $\Lambda(k,h)$. Therefore, we differentiate this expression twice to obtain the following results.

Lemma 2. *If $\Lambda_1(k,h) > 0$ and $\Lambda_{11}(k,h) < 0$, then $y_1(k,h) > 0$ and $y_{11}(k,h) < 0$.*

Proof. See appendix □

One implication of lemma 2 concerns the ranking of firm. If a type k firm has a higher opportunity cost than a type k' firm, then it is higher ranked. However, by lemma 2 the higher type k firm is also more productive. Therefore, we can rank firm type k over firm type k' by the following condition.

$$y(k,h) > y(k',h)$$

Moreover, since each firm type's opportunity cost is common to all workers, this ranking applies to sub-markets involving all of the different worker types. Therefore, in equilibrium, higher ranked firms are more productive when matched to any worker type than less productive firms. Lemma 2 also shows that the assumed concavity of the job value function also implies concavity of the production function, which is assumed in the static game.

We derive an expression for the equilibrium wage at each labor auction by taking equation (12) for any pair $\{k_1, k_2\}$ and subtracting the same equation evaluated at $\{k_1, k_1\}$ where the wage is equal to $y(k_1, h)$. We have:

[16] Julien, Kennes and King (2006) and Kennes and le Maire (2010) apply the competing auction model as a tool for understanding observed wage dispersion between similar workers. The general conclusion is that the competing auction model can closely fit observed wage distributions.

$$w(k_1,k_2,h) = \begin{cases} y(k_2,h) & \text{if} \quad k \geq p^*(h) \\ y(k_1,h) - [1-\beta(1-\delta)/\Lambda_1(k_1,h)][\Lambda(k_1,h) - \Lambda(k_2,h)] & \text{if} \quad k < p^*(h) \end{cases} \tag{15}$$

This equation is indexed by whether or not the worker is in a sufficiently high type job such that the worker can expect not to receive rival offers. If the employed worker is in a type $k \geq p^*(h)$ job and does not expect rival offers, then the worker's wage is simply the productivity of the second highest valuation firm, which is the same as the static model. However, if the employed worker is in a lower type job and expects rival offers (i.e. the worker expects offers via on-the-job search), then the wage is reduced since the firm now expects that the worker will have a shorter tenure and thus is of less value to the bidding firm. We use equation (15) and lemma 1 to derive the following lemma.

Lemma 3. *Wages are monotonically increasing in the firm type.*

Proof. See appendix. □

By Lemma 3, wages can be used to rank firms.

3.5 Wages and the ranking of workers

The workers are ranked by their expected market utility when unmatched at the start of the period. An unmatched type h worker with home production technology $\underline{k}$ has an expected market utility given by

$$W(\underline{k},\underline{k},h) = \int_{z=\underline{k}}^{k^*(h)} \Lambda(z,h)\, dG^2(z \mid h)$$

where $\Lambda(z,h)$ is the present value of a worker and type z firm match, and $G^2(z \mid h)$ is the distribution of second highest valuations. Therefore, if $W_3(\underline{k},\underline{k},h) > 0$, the workers are ranked by h.

The following lemma provides the sufficient condition on the job value function for the expected market utility $W(\underline{k},\underline{k},h)$ to be monotonically increasing in h.

Lemma 4. *When* $\Lambda_2(k,h) > 0$ for all k, the workers are ranked by h.

Proof. See appendix. □

By lemma 4 and equation (14), we notice that even though $\Lambda_2(k,h) > 0$, this does not imply that $y_2(k,h) > 0$. This simply says that even though a worker in expectation is more productive than another, there might be some jobs where he is not. Therefore, the case of no global absolute advantage in figure 2 (b) is still possible under the assumption of lemma 4.

As in the analysis of the static game, we can relate how to use the distribution of wages at the highest type employer as a means to solve for the expected market utility.

Lemma 5. *The expected market utility of a type h worker in the unmatched state is given by*

$$W\left(\underline{k},\underline{k},h\right)=\frac{1}{1-\beta}\int_{z=\underline{k}}^{k^*(h)}w\left(k^*\left(h\right),z,h\right)dG^2\left(z\mid h\right)$$

Proof. See appendix. □

We use lemma 5 in the empirical identification section as a means to rank workers using data on the distribution of wages for the worker's highest type employer.

3.6 Steady state equilibrium

The block recursive solution for equilibrium employment transitions, wages, and profits applies to any initial distribution of jobs to worker types. This section solves for the steady state. A steady state equilibrium imposes the additional requirement that the distribution of workers to different jobs and unemployment in any period is equal to the distribution of workers to jobs and unemployment in the next period. The steady state equilibrium is solved as follows. The probability that an unmatched worker moves into employment at the start of the period is equal to the probability that the worker has at least one offer, $1-G^1\left(\widehat{k}\left(\underline{k},h\right)\mid h\right)$. At the end of the period there is a probability δ of moving into the unmatched state. The steady-state unemployment rate $u(h)$ of a type h worker after the bidding stage is found by equating the flows into and out of unemployment. We have

$$u\left(h\right)=\frac{\delta G^1\left(\widehat{k}\left(\underline{k},h\right)\mid h\right)}{1-\left(1-\delta\right)G^1\left(\widehat{k}\left(\underline{k},h\right)\mid h\right)} \tag{16}$$

We let $n\left(k\mid h\right)$ denote the density of type h workers employed in a type k job where $n\left(k\mid h\right)=\int_{\underline{k}}^{k}\left(n\left(k,k_2,h\right)\right)dk_2/\int_{\underline{k}}^{k^*}\left(\int_{\widehat{k}(\underline{k},h)}^{k^*}\left(n\left(k_1,k_2,h\right)\right)dk_1\right)dk_2$ and $N\left(k\mid h\right)$ denote the distribution of type h workers with job types less than k. In equilibrium, inflow to the mass

$N(k \mid h)$ must equal outflow

$$
\begin{aligned}
& u(h)\left(G^1(k,h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) \\
& + (1-u(h))(1-N(k \mid h))\,\delta\left(G^1(k \mid h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) \\
= & (1-u(h))\,N(k \mid h)\,\delta\left[1-\left(G^1(k \mid h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right)\right] \\
& + (1-\delta)(1-u(h))\int_{\hat{k}(\underline{k},h)}^{k} n(z \mid h)\left[1-G^1\left(\max\left(\hat{k}(z,h),k\right) \mid h\right)\right] dz
\end{aligned}
$$

There are two types of inflow to the mass $N(k \mid h)$. First, we have the inflow from unemployment. Second, some jobs with a higher type than k are exogenously destructed and some of the workers are approached by firms of type k or less. There are two types of outflow. First, jobs in the mass $N(k \mid h)$ are destructed and some stay unemployed whereas other find employment at firms of k or above. Second, some workers whose jobs are not destructed leave firms of type k or less to join firms above type k. In the appendix we show that

$$
\begin{aligned}
N(k \mid h) \;=\; & \left[\delta + \frac{u(h)}{1-u(h)}\right]\left(G^1(k,h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) \\
& + (1-\delta)\int_{\hat{k}(\underline{k},h)}^{k} n(z \mid h)\,G^1\left(\max\left(\hat{k}(z,h),k\right) \mid h\right) dz \qquad (17)
\end{aligned}
$$

3.7 Sorting

In the analysis of the static game we considered sufficient conditions to ensure that the distribution of offers is assortative across different types. In this section, we apply our definition of PAM/NAM to the analysis of the steady state distribution of firms to workers. The steady state distribution of workers is complicated by the possibility of on-the-job search in the stage game whereby workers climb a firm type ladder. Therefore, in order to analyze sorting in the dynamic competing auction game, we evaluate the steady state equilibrium allocation $N(k \mid h)$ and the lower and upper bounds of the matching set. By the definition of sorting, if worker h' is a higher type than worker h, PAM is characterized by $N(k \mid h') < N(k \mid h)$, $\hat{k}(\underline{k},h') > \hat{k}(\underline{k},h)$, and $k^*(h') > k^*(h)$, and NAM is characterized by $N(k \mid h') > N(k \mid h)$, $\hat{k}(\underline{k},h') < \hat{k}(\underline{k},h)$, and $k^*(h') < k^*(h)$. We then derive the following sufficiency condition for PAM/NAM in the steady state equilibrium.

Proposition 3. *A sufficient condition for PAM/ NAM in the steady state equilibrium is that the job value function $\Lambda(k,h)$ is PRC/NRC.*

Proof. See appendix □

The role of PRC/NRC (See definition 3) is due to the complicating factor that the continuation value of the worker, $\Lambda(\underline{k}, h)$ is not generally zero, which means that log-supermodularity/log-submodularity of $\Lambda(k, h)$ is not a sufficient condition for PAM/NAM. The stronger assumption of a PRC/NRC job value function is a means to ensure that the steady state equilibrium assignment is PAM/NAM.

We can also relate complementarities in the job value function between higher/lower worker and firm types to the complementarities between these worker and firm types in the primitive production function. The functions $\Lambda_{12}(k, h)$ and $y_{12}(k, h)$ are related (with some manipulation given in the appendix) as follows:

$$\frac{y_{12}(k,h)}{y_1(k,h)} = \frac{\Lambda_{12}(k,h)}{\Lambda_1(k,h)} + \begin{cases} 0 & \text{if } \quad k \geq p^*(h) \\ \frac{\beta(1-\delta)}{\Lambda_1\left(\hat{k}(k,h),h\right)-\beta(1-\delta)} \frac{\int_k^{\hat{k}(k,h)} \Lambda_{12}(z,h)dz}{\int_k^{\hat{k}(k,h)} \Lambda_1(z,h)dz} & \text{if } \quad k < p^*(h) \end{cases} \tag{18}$$

where all term of the r.h.s. are positive/negative if and only if $\Lambda(k, h)$ is super/sub-modular. Furthermore, If $\frac{\Lambda_{12}(k,h)}{\Lambda_1(k,h)} > 0$ and nondecreasing in k, $\frac{y_{12}(k,h)}{y_1(k,h)} > 0$ will be nondecreasing in k, since $\frac{\beta(1-\delta)}{\Lambda_1\left(\hat{k}(k,h),h\right)-\beta(1-\delta)}$ is increasing in k. If $\frac{\Lambda_{12}(k,h)}{\Lambda_1(k,h)} < 0$ and decreasing in k, $\frac{y_{12}(k,h)}{y_1(k,h)} < 0$ will be decreasing in k, since $\frac{\beta(1-\delta)}{\Lambda_1\left(\hat{k}(k,h),h\right)-\beta(1-\delta)}$ is increasing in k.

Proposition 4. *If the job value function $\Lambda(k, h)$ is PRC/NRC, the production function $y(k, h)$ is PRC/NRC.*

Proof. Follows directly from equation (18) and definition 3. □

Therefore, if the steady state assignment is assortative, then there are also complementarities in production between workers and firms.

3.8 The non-monotonicity of worker wages

As we noted in section 3.5, the wage equation (15) is not generally informative about the worker ranking. From equation (15), there are two ways that this non-monotonicity in the wages can arise. First, no absolute advantage implies that the production function $y(k_1, h)$ is not monotonic in h, whereby $w(k_1, k_2, h)$ may be non-monotonically in h. Second, the second term of equation (15) is negative and is numerically largest with a high degree of submodularity or supermodularity of the job value function. This may lead to the wage being non-monotonic under PAM, but not under NAM. We summarize the different cases by the table below.[17]

[17] The results are derived in the appendix.

	Global absolute advantage	No global absolute advantage
NAM	Monotonically increasing in h	Non-monotonic in h
PAM	May be non-monotonic in h	Non-monotonic in h

Table 1: PAM/NAM and the monotonicity of wages in worker type

The non-monotonicity of wages under PAM and global absolute advantage can be easily understood by a simple example. In this example, a higher type worker h frequently gets offers from higher type firms but infrequently gets offers from medium type firms while a lower type worker h' frequently gets offers from lower type firms but infrequently gets offers from medium type firms. This gives the following implications for wages. The high type workers expects a low wage offer from a medium type firm, because there are no counter offers from lower type firms when the worker's best offer is a medium type firm (the case of a medium type firm counter offer is also rare). However, the low type worker expects a high wage from a medium type firm, because the worker also expects offers from low type firms when the worker's best offer is a medium type firm. Therefore, the medium type firm pays the lower type worker higher wages.

It is also interesting to note that the results summarized in Table 1 do not imply a similar logic when higher type workers have global absolute advantage and the assignment is NAM. In this special case, the expected wage paid by each firm type will increase monotonically with respect to higher worker types. In the example, we would have a NAM assignment if $h' > h$. The wage will be increasing in firm type because the higher type worker is more productive in each firm type and the higher type worker has higher expected counter offers when matched to each firm type.

Figure 3: Sorting example

Consider the example with two worker types and three firm types, which is illustrated in figure 3. Relative to the high type worker, the low type worker is more productive in the low type firm, equally productive in the medium type firm, and less productive in the high type firm. The mapping of productivities into expected wages and matching sets is summarized in the right panel of figure 3. The example illustrates the general feature of the model that higher firm types pay higher expected wages to each worker type. The example also illustrates the possibilibility that within the intersection of the matching sets of the two worker types, the higher type worker is paid lower wages. The key problems of empirical identification here are to ensure that the high type workers are correctly ordered and that we can identify firm types for which the high type workers are less productive than low type workers.

4 Empirical Identification

This section derives the predictions of the model that are used to identify worker and firm rankings, the sign and strength of sorting, and a test of global absolute advantage in production.

4.1 Identifying the ranking of firms

By Lemma 3, wages can be used to rank firms. This is in contrast to some prominent other models. One of important reason for this difference is the assumption of directed matching. In the related sequential auctions model with undirected search, Postel-Vinay and Robin (2002), higher type firms may initially offer workers lower wages than lower type firms, because in the future they will be able to match outside offers by bidding up the wage. With directed recruitment, less productive firms never poach more productive firms and, hence, the wage is not bid up during the worker's tenure. Even without on-job search, wages may not be monotonic with random search. If the number of participants in the matching market is fixed, as in Shimer and Smith (2000), a high type firm may agree to hire a relatively low type worker if the worker accepts a sufficiently low wage to compensate the firm for the option value of waiting for a more productive potential hire. This does not occur in our model, because the number of firms for each worker submarket is determined by free entry. The general problem of non-monotonic wages is described by Eeckhout and Kircher (2011).

4.2 Identifying the strength of sorting

We follow Eeckhout and Kircher (2011) and Bartolucci et al. (2018) and estimate the strength of sorting (degree of complementarity) by examining whether each worker is employed in similar ranked firms. We only use workers with at least one unemployment spell and exploit that observations before and after the unemployment spell are not related due to on-the-job search. While the correlation of the average firm rank before and after an unemployment spell measures the strength of sorting, it will not determine the sign of sorting since a high degree of PAM or NAM both lead to high correlations in within-worker firm ranks. Under PAM/NAM, the distribution of higher/lower type workers stochastically dominates the distribution of lower/higher type workers. Therefore, when letting the number of time periods go to infinity, measuring the worker type by the average firm ranking will approximately give us the correct ranking under PAM, but the inverse ranking under NAM.

4.3 Identifying the sign of sorting

It is a difficult task to identify the ranking of workers in particular when we are not assuming global absolute advantage. For example, high type workers may earn lower wages at low type firms than lower type workers simply because they are less productive in these jobs. To isolate the worker ranking we will use lemma 4 and rank workers based on their expected market utility. Obviously, expected market utility is not observed in data and we need to relate this to the workers' assignment and observed wages. Lemma 5 shows how to use the distribution of wages at the highest type employer as a means to solve for the expected market utility. We have

$$W\left(\underline{k}, \underline{k}, h\right) = \frac{1}{1-\beta} \int_{z=\underline{k}}^{k^*(h)} w\left(k^*\left(h\right), z, h\right) dG^2\left(z \mid h\right) \tag{19}$$

Our identification result is related to the worker ranking of Bagger and Lentz (2018). They show that they can use wages in the highest firm type immediately after an unemployment spell to rank workers. In our setting with multiple bidders, we need isolate the distribution of second best offers in order to calculate the expected market utility. Therefore, we need to reweight the wages by worker type using $G^2\left(k \mid h\right)$. Given the definitions of $G^1\left(k \mid h\right)$ and $G^2\left(k \mid h\right)$, it straightforward to show that $g^2\left(k \mid h\right) = -g^1\left(k \mid h\right) \ln G^1\left(k \mid h\right)$. In the algorithm below, we lay out how to rank workers and determine the sign of sorting.

Algorithm 1: *Ranking workers and identifying the sign of sorting.*

1. Rank workers by using the average firm rank a worker is employed by and group workers into groups according to this ranking.

2. Consider each group $l = 1, 2, ..., L$ of workers separately and estimate the distribution of wages for those leaving unemployment to join their highest ranking firm, i.e. $G^{1,est}\left(w \mid h_l\right)$ and $g^{1,est}\left(w \mid h_l\right)$.

3. Calculate
$$g^{2,est}\left(w \mid h_l\right) = -g^{1,est}\left(w \mid h_l\right) \ln G^{1,est}\left(w \mid h_l\right)$$

4. Calculate
$$\widetilde{W}^{est}\left(\underline{k}, \underline{k}, h_l\right) = \int_{-\infty}^{\infty} w\left(w \mid h_l\right) g^{2,est}\left(w \mid h_l\right) dw$$
where we have omitted the term $1/\left(1-\beta\right)$ from equation (19).

5. Repeat 2-4 for each worker group l.

6. Examine the relationship between $W^{est}\left(\underline{k}, \underline{k}, h_l\right)$ and h_l. If we have an increasing relationship, workers employed in the highest type firms have the highest market value, so the labor market features PAM. If we have a decreasing relationship, workers employed in the highest type firms have the lowest market value, so we have NAM.

4.4 Identifying global absolute advantage

Finally, we need to be able to identify whether the labor market is characterized by global absolute advantage in production. If the assignment is assortative, Proposition 4 established that when $\Lambda_{12}\left(k,h\right) > / < 0$ we know that $y_{12}\left(k,h\right) > / < 0$. However, in either case, it is also possible that $y_2\left(k,h\right) < 0$ even when $\Lambda_2\left(k,h\right) > 0$. Thus under both PAM and NAM, we do not necessarily assume global absolute advantage.

From equation (15), it is clear that when $k_1 = k_2 = k$ we have that the wage is equal to the productivity of the match $w\left(k,k,h\right) = y\left(k,h\right)$. If using only worker-firm observations where $k_1 = k_2 = k$, it would be trivial to examine the sign of $w_3\left(k,k,h\right) = y_2\left(k,h\right)$ and tell whether the labor market features global absolute advantage as this would imply that $w_3\left(k,k,h\right) > 0$ for all $\left(k,h\right)$. Alternatively, if we observe that $w_3\left(k,k,h\right) > 0$ for some firms whereas $w_3\left(k,k,h\right) < 0$ for other firms, we would conclude that the labor market is not characterized by global absolute advantage. Unfortunately, as k_2 is unobserved we can never tell when $k_1 = k_2$. Furthermore, we do not observe each worker enough times to be sure that the highest wage for each worker in each firm correspond to $k_1 = k_2$. Therefore, for the identification strategy for global absolute advantage we need to group both similar firms and similar workers such that we can hope that k_2 at least is close to k_1.

Algorithm 2: *Identifying global absolute advantage*

1. Aggregate workers in L groups according to the average firm rank the worker has visited as in algorithm 1.

2. Aggregate firms in M groups according to the firm ranking.

3. Estimate the following equation by quantile regression using a high quantile, say the 95th quantile, separately for the observations for each firm group $m = 1, 2, .., M$

$$\ln w_{it} = \gamma_0 + \gamma_1 l_{it}^{est} + \xi_{j(i,t)} + \varepsilon_{it}$$

where the worker ranking from algorithm 1,l_{it}^{est}, is included linearly and $\xi_{j(i,t)}$ are firm fixed effects for the firms in group $m = 1, 2, .., M$.

4. If $\gamma_1^{est} > 0$ for all groups of firms $m = 1, 2, .., M$, then the labor market features global absolute advantage. If $\gamma_1^{est} > 0$ for low type firms and $\gamma_1^{est} < 0$ for high type firms, this is evidence against global absolute advantage. Both PAM and NAM implies that γ_1^{est} is increasing in firm type since the workers' average firm rank will be increasing/ decreasing in worker type under PAM/ NAM.

5 Simulations

In this section, we simulate the dynamic model to gain additional insights on how parameters controlling complementarity and advantage affect the estimation in finite samples using our strategies for the identification of sorting and the identification of global absolute advantage. For the simulations it is useful to specify the job value function $\Lambda(k, h)$ since by lemma 4 workers are ranked by h when restricting focus to job value functions where $\Lambda_2(k, h) > 0$. We use the following specification of the job value function $\Lambda(k, h)$

$$\Lambda(k, h) = A_0 + A_1 k^{\alpha_1} + A_2 h^{\alpha_2} + A_3 k^{\alpha_1} h^{\alpha_2} \tag{20}$$

where we set $A_0 = 25$, $A_1 = 1.5$, $\alpha_1 = \alpha_2 = 0.3$.[18] We vary A_2 between 0.0 and 2.1 and A_3 between -0.7 and 1.9 and in total we have 112 combinations of A_2 and A_3. However, we only consider the 108 cases where the regularity conditions are met, that is $\Lambda_1(k, h) > 0$, $\Lambda_2(k, h) > 0$ and $\Lambda_{11}(k, h) \leq 0$ for all k, h. The remaining four cases for which these regularity conditions are not met correspond to the white areas in figures 4-10.

We assume 40 different worker types with h_l for $l = 1, 2, ..., 40$ being evenly distributed between 0.5 and 3.0 and that there exist 1000 persons of each type. We assume 800 firms. This gives us an average number of workers per firm per time period of 50 (abstracting from unemployment). Given equation (20) and the assumed parameter values, we can determine $k^*(h_l)$ for each of the worker types and fix $\underline{k}$ such that all worker types can search on the

[18] For each simulation, we initially set $A_0 = 25$, but then subsequently adjust A_0 upwards if any wages are negative such that we can take the log of the wages.

job.[19] We set the exogenous job destruction rate $\delta = 0.16$ and the discount factor $\beta = 0.96$. In each of the 108 simulation, we observe each worker for 25 time periods.

Figure 4 shows the true rank correlation between worker and firm types for each of the 108 simulated economies. It is clear that the sign of A_3 controls the complementarity of the job value function and, hence, determines whether the economy exhibits PAM/NAM and also the strength of sorting. A positive/negative A_3 gives PAM/NAM. The larger positive/ negative A_3, the stronger strength of sorting. A_2 impacts solely the advantage of different workers but does not induce more or less complementarity and, hence, has no effect on the sorting.

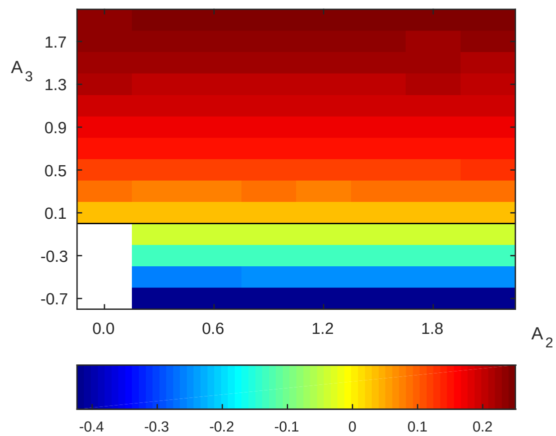


Figure 4: True rank correlations

In figure 5, we depict the rank correlations between AKM firm and worker fixed effects against their true values. From lemma 3, we know that wages are monotonically increasing in firm type. Panel (A) of figure 5 shows that the rank correlations between the estimated AKM firm fixed effects and the true firm types are very close to 1 in the simulations. In contrast to this, panel (B) shows that the wage may not be monotonic in the worker type and that the correlation between the estimated AKM worker fixed effect and the true worker type can easily be negative in particular when the advantage parameter A_2 is small and the complementarity parameter A_3 is large and positive. Below, we will show that these cases with negative correlations between the true worker type and the estimated AKM worker fixed effect do not feature global absolute advantage.

[19] We set $\underline{k} = \min_l \frac{1}{2}\left(\hat{k}^{-1}\left(\hat{k}^{-1}\left(k^*\left(h_l\right)\right)\right)\right)$.

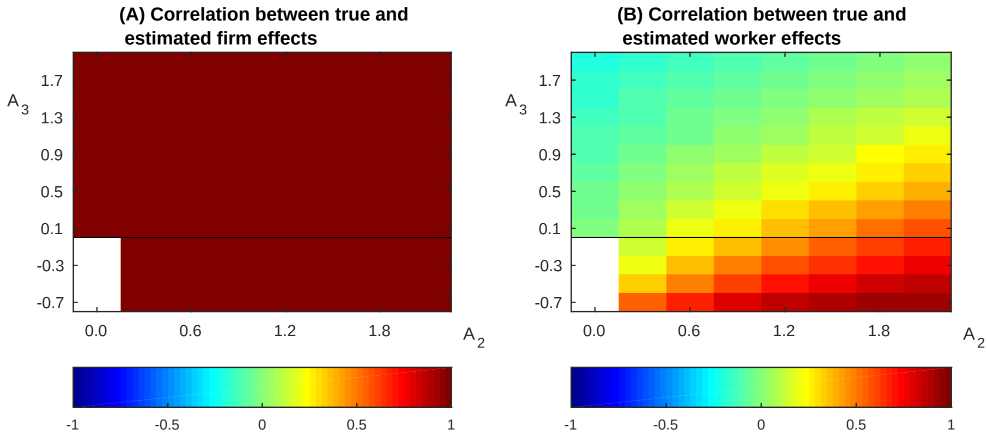


Figure 5: Rank correlations between estimated AKM fixed effects and true worker and firm types

Under PAM/NAM, the allocation of higher/lower type workers across firms stochastically dominate the allocation of lower/higher type workers. Therefore, we compute the worker ordering as the average firm rank the worker obtains in the sample window. Figure 6 shows how well the average firm rank (using AKM firm fixed effects) correlate with the true worker ranking. We report the absolute value of the rank correlations. In panel (A) we use all observations. It is clear that when the degree of complementarity is high, i.e. when A_3 is not too close to zero, the matching sets of different workers are quite different and we obtain high rank correlations between the average firm rank visited and the true worker type. However, when the matching sets of different worker types are very much alike, the average firm rank is a poor measure of worker type and in the special case of a modular production function, where all workers have identical matching sets and identical job offer distributions, our measure of worker type is uninformative.

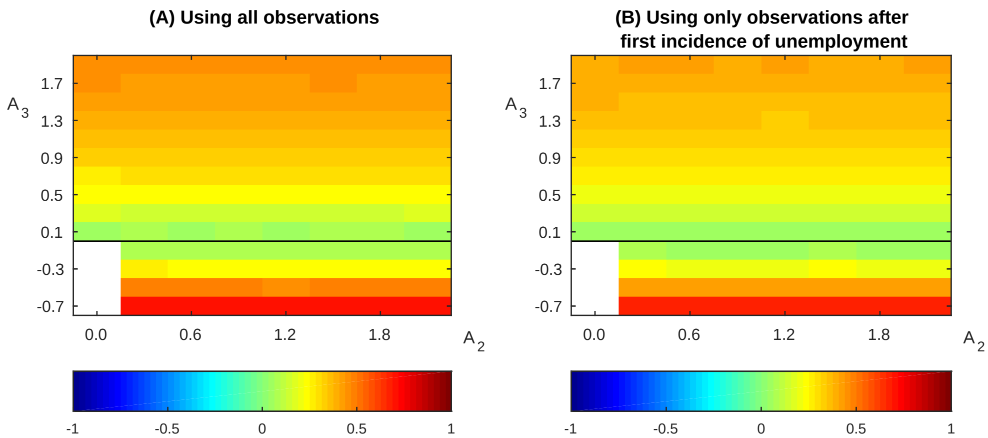


Figure 6: Rank correlations between true worker effect and estimated average firm rank

To assess the strength of sorting, we simply correlate a worker's estimated firm rankings before and after an unemployment spell. In panel (B) of figure 6, we only used observations following an unemployment spell to estimate the worker ranking using the average firm ranking visited. With fewer observations used, we naturally obtain a slightly lower correlation with the true worker ranking than when using all observations as in panel (A).

Using the estimate of the worker type in panel (B) of figure 6, we obtain an estimate of the strength of sorting by correlating this with the estimated firm ranks of the firms the worker visits before the unemployment spell. In panel (A) of figure 7, we show the estimated strength of sorting. As expected, the lowest and highest values of A_3 give the highest rank correlations. However, panel (B) also shows that the estimated correlations are generally lower than the true absolute correlations. The differences between the estimated strength of sorting and the absolute value of the true correlation are between -0.15 and -0.02.

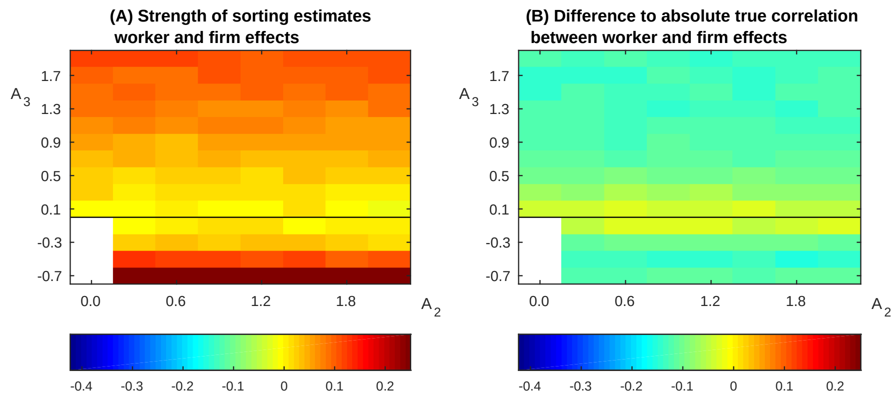


Figure 7: The strength of sorting

In figure 8 we examine how well we are able to determine the sign of sorting. For this, we need to figure out whether the estimated worker ranking in figure 6 is positively or negatively related to the estimated expected market utility ranking based on $\widetilde{W}(\underline{k}, \underline{k}, h)$ using algorithm 1. This identification strategy works well in all but the few cases where the advantage parameter, A_2, is low and the complementarity parameter, A_3, is negative. In these cases the small advantage and the low submodularity works against each other, leaving the true expected market utility to only differ marginally between different worker types. The reason is that high type workers will tend to have slightly higher wages in most jobs, but the distribution would be slightly more favorable for low type workers. Notice that when $A_3 > 0$, the effect of advantage and supermodularity work in the same direction by both increasing the expected market utility for high type workers relative to low type workers.

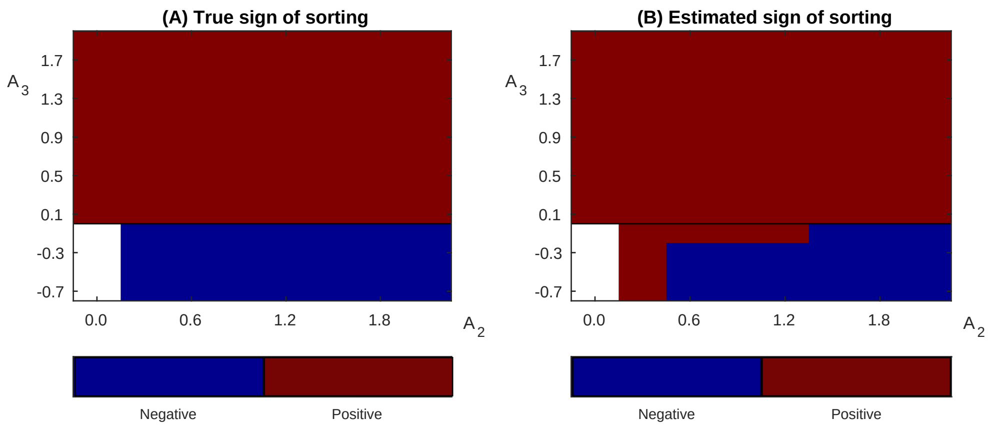


Figure 8: The sign of sorting

The test of global absolute advantage is computed using algorithm 2. We divide firms into 25 groups based on the AKM firm fixed effect. For each group we regress the wage on the worker's average firm rank using quantile regression in the 95th quantile. If all these 25 estimated γ_1 coefficients have the same sign, this is evidence of absolute advantage. However, estimation noise can imply that a few of the γ_1 estimates are significant with the opposite sign even under global absolute advantage. We need a simple rule to avoid any estimation noise leading us to mistakenly conclude that such cases do not feature global absolute advantage.[20]

We know that under no global absolute advantage, the true γ_1 will have opposite signs for low type and high type firms, so consequently opposite signs for the estimated γ_1 for only a few intermediate firm types is likely to be due to estimation noise. Hence, we also classify the simulated economy as featuring global absolute advantage if less than 10 percent of the estimated γ_1 coefficients are significant positive (negative) and the lowest firm group with a significant positive (negative) γ_1^{est} is not among the bottom or top 10 percent of the firm groups. Our test of global absolute advantage works quite well as shown in figure 9 and in most cases we correctly detect whether a simulated economy features global absolute advantage or not.

[20] In an empirical application, plotting the estimated γ_1 coefficients against the firm groups is likely to be more useful, but with 108 cases, we need a simple and general rule.

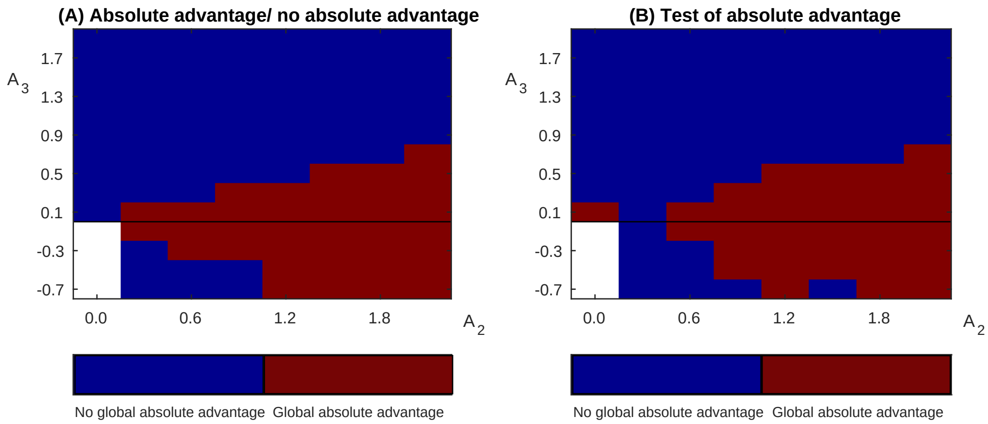


Figure 9: Absolute advantage

Finally, it is relevant to examine how the deviation from the assumption of absolute advantage impacts the estimated AKM correlation between worker and firm fixed effects. Figure 10 compares the true correlations between worker and firm types (panel A) to the estimated estimated AKM correlations between worker and firm fixed effects (panel B). Since the AKM estimation estimates the firm effects very well, the AKM correlation between worker and firm fixed effects is only approximately correct when the worker effect is precisely estimated and this is essentially only true for NAM cases. Panel (C) reveals that the AKM correlation is wrongly signed for all cases with PAM and no absolute advantage. However, panel (C) also shows that an estimated positive AKM correlation implies PAM since the estimated correlations are downward biased.

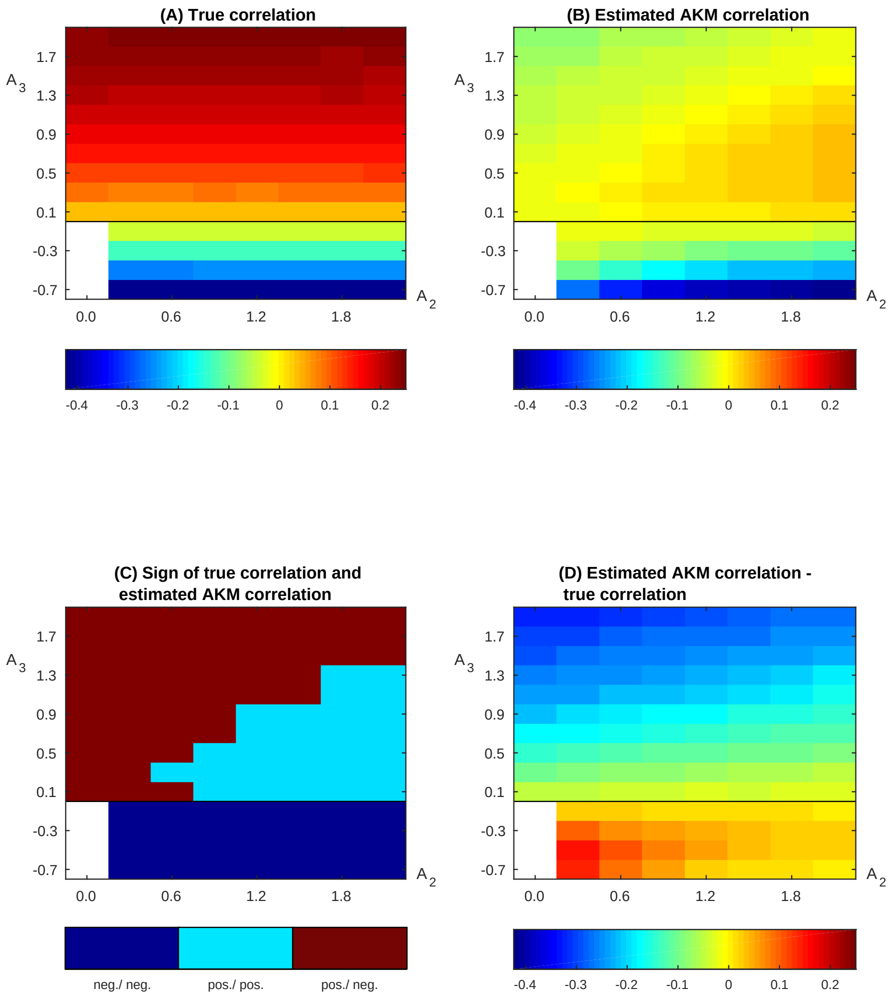


Figure 10: True and AKM estimated correlations

6 Estimation

In this section we apply our methodology to Danish register data for 1995-2011. We find that the Danish labor market features PAM and that the production function does not satisfy global absolute advantage.

6.1 Data

The data set is extracted from the Integrated Database for Labor Market Research (IDA), which covers the entire Danish population aged 15–74. IDA associates each person with his/her unique identifier, and provides annual data on many of the individual's socioeconomic characteristics, such as hourly wage, education, and occupation. We measure the hourly wage rate as annual labor income plus mandatory pension fund payments divided by annual hours.[21] To match our firm data with our worker data we draw on the Firm-Integrated Database for Labor Market Research, or FIDA, which links every firm with every worker in IDA who is employed by that firm in the last week of November of each year. Using this matched worker-firm data, we can consistently track virtually every person in the Danish economy over time regardless of his/her employment status or employer identity.

We make two sample restrictions on the dataset. First, we only consider workers and firms in the private sector. About 40% of workers in Denmark work in the public sector. Second, we only focus on the largest connected set of firms. This is because in AKM regressions person fixed effects and firm fixed effects can be identified separately only for firms in the largest connected sets. The largest connected set in our sample includes over 99.4% of the workers and over 99.6% of the observations, so this restriction has little effect on our sample size. We end up with a sample of 19,161,478 observations with 2,232,547 workers and 242,659 firms.

6.2 Ranking firms

We rank firms using the firm fixed effects from the two-way fixed effects as in Abowd, Kramarz and Margolis (1999) and more recently in Card, Heining and Kline (2013). We estimate the following regression:

$$\ln w_{it} = \psi_{J(i,t)} + \alpha_i + \theta_t + x'_{it}\beta + \epsilon_{it}$$

where $\psi_{J(i,t)}$ is a vector of firm fixed effects, α_i is a vector of worker fixed effects, and θ_t and x_{it} are year fixed effects and time varying individual characteristics, respectively. As in Card, Heining and Kline (2013), we include in x_{it} an unrestricted set of year dummies as

[21] Information about annual hours comes from the pension fund, Arbejdsmarkedets Tillægspension (ATP), which collects a relatively modest mandatory pension fund payment from all workers in the Danish labor market. The payment depends on the number of hours worked in the following way: (i) no payment if working 0–9 hours per week, (ii) 1/3 of full-time payment if working 9–18 hours per week, (iii) 2/3 of full-time payment if working 18–27 hours per week, and (iv) full-time payment if working 27 or more hours per week. The hours worked are then imputed from knowledge about a worker's ATP group. If the worker is registered as having paid full-time ATP, then the hours worked are measured as the number of hours corresponding to the standard 37-hour workweek.

well as quadratic and cubic terms in age fully interacted with educational attainment. Firms are ranked using the estimated firm fixed effects $\psi_{J(i,t)}$.

As a robustness check, we also rank firms using the poaching index as in Bagger and Lentz (2016). The poaching index is defined by the fraction of hires that is poached from other firms, and a higher-ranked firm hires relatively more from other firms instead of from the unemployment pool. We do not use poaching index in the main analysis over the period 1995-2011 because we only have monthly employment data to precisely measure job flows after 2007. The two firm rankings are positively correlated: correlation between AKM firm fixed effects and the poaching index is 0.27.

6.3 Strength of sorting

We measure the strength of sorting using the correlation between the firm rank before the first observed unemployment for each person and the average firm rank after this unemployment spell. We use the unemployment spell to reset the climbing of the firm productivity ladder to avoid the spurious correlation due to on-the-job search. The estimated correlation is 0.30, which is much higher than the correlation between AKM firm fixed effects and worker fixed effects, 0.08. This is consistent with studies showing that the correlation between AKM fixed effects underestimate the true correlation between worker type and firm type (see for example, Andrews et al. 2008; Borovicková and Shimer 2017; Woodcock 2015). This correlation between worker type and firm type is close to the estimate of 0.37 in Bagger and Lentz (2018), who also use Danish register data but use different methods to measure firm types and worker types. Our estimate is also close to the estimate of 0.28 in Lentz, Piyapromdee and Robin (2018).

6.4 Worker ranking and sign of sorting

We follow Algorithm 1 to rank workers and determine the sign of sorting. First, we put workers in 50 groups based on the *average* ranked firm each worker is matched with. Second, we estimate for each worker group the expected market utility.

Figure 12 shows the expected market utility of the worker groups. The horizontal axis is worker group ranked by the average matched firm type. The expected utility is almost linearly increasing in the worker type, therefore higher-ranked workers are also employed by higher-ranked firms. This indicates that the Danish labor market features PAM.

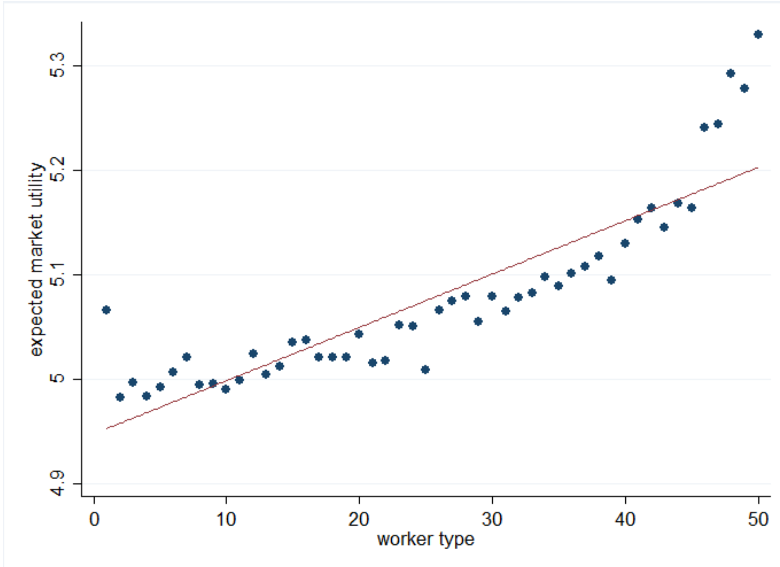


Figure 11: Sign of Sorting

6.5 Testing global absolute advantage

Finally, we follow Algorithm 2 to test global absolute advantage. We put all the firms in 25 bins based on AKM firm fixed effects, and Figure 13 plots the estimated coefficient γ_1^{est} in the quantile regression of the 95th percentile log wage for each firm bin.[22] In the regression, we include year fixed effects, industry fixed effects, as well as quadratic and cubic terms in age fully interacted with educational attainment.

The graph shows that global absolute advantage fails to hold in the Danish labor market. The coefficients are negative and significant for the lowest 7 firm bins, which is about 28% of the firms. For instance, for the lowest firm type, a coefficient of -0.0006 means that the productivity of workers at the 75th percentile is 3% lower than the productivity of workers at the 25th percentile. Since 95th wage percentile approximates the productivity, this suggests that for the lowest ranked firms higher ranked workers are less productive.

In the Appendix we apply the test of global absolute advantage separately for different observable characteristics, including region, years, gender, age and industry. We find that the result of no global advantage is pervasive across the different breakdowns besides for the industries of finance and knowledge services. Furthermore, we also find evidence of no global absolute advantage using the poaching index of Bagger and Lentz (2018) as a means to rank firms.

[22] We choose 95th percentile rather than 99th percentile because it's less noisy.

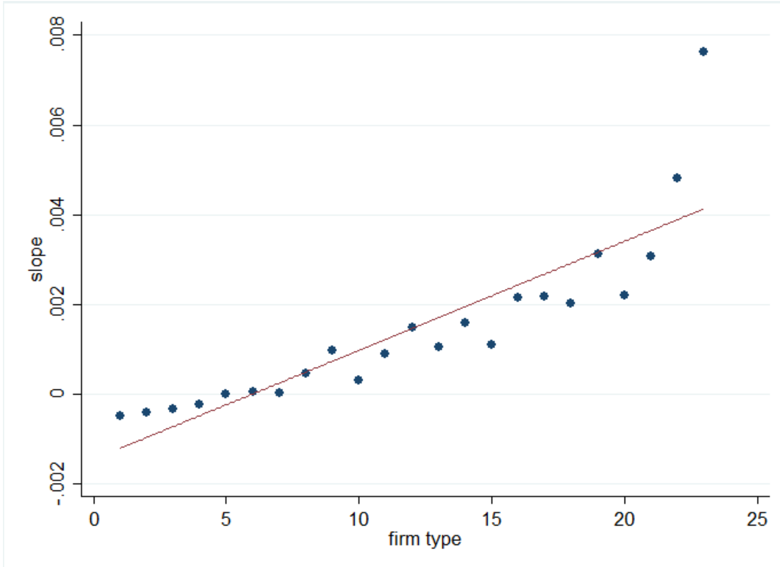


Figure 12: Testing Global Absolute Advantage

7 Conclusion

McAfee's (1993) model of mechanism design by competing sellers is a foundational model of directed matching (Wright et al., 2017). The application of this model to the sale of labor offers a microeconomic foundation for the analysis of unemployment, wage dispersion, on-the-job search and sorting. The model predicts that higher valuation bidders (firms) will always pay higher wages to the sellers of labor (workers) than lower valuation bidders. This prediction is driven by the directed assignment of the bidders to different worker submarkets and the equilibrium bidding strategy of the higher valuation firms to always win the worker's services in head to head competition with lower valuations firms. The model also predicts, with the exception of the case of NAM and global absolute advantage, that higher worker types will not generally earn higher wages than lower worker types in all job types. This result is driven not only by the endogeneity of the number of bidders at each submarket, but also the possibility that higher type workers are less productive than low type workers in low type jobs.

While firms can be ranked by wages (and/or by a poaching index), the key challenge for the empirical assessment of labor market sorting is to rank workers. If there are strong complementarities between high type workers and high type firms and the high type workers do not have a global absolute advantage (features of the labor market that we suggest are important), the wages of workers will generally not be monotonic in worker type at each firm, and the more capable workers may not always outshine less capable workers in all jobs that employ them. Since these workers cannot be ranked by either their productivity or their wage at any particular firm, our proposal was to rank the workers by their expected market utility. We then use the predictions of the competing auctions model to identify how this worker expected market utility ranking can be empirically identified from the distribution of wages the workers obtain in the best jobs that they can find. Using the estimated expected market utility ranking, we can estimate the sign of sorting.

We propose a simple test of global absolute advantage. This test assumes that the highest wages a given worker type receives in a given firm type approximate the productivity of this worker type. As input to this test we essentially only need estimated firm rankings and wage observations for a matched employer-employee data set. Therefore, the test can also serve as specification test of the AKM model using a firm ranking based on AKM firm fixed effects.

We find evidence that the labor market is both highly sorted and that high type workers are sometimes poorly assigned to jobs where they are less productive than less capable workers. Therefore, the directedness of matching firms to workers is clearly important and also imperfect. We have considered only some of the challenges associated with empirical modelling the directedness of matching in the labor market. We have used the predictions of competing auction theory to address these challenges and we can speculate that this theory will offer additional solutions to other problems of empirical identification both in the labor market and in other contexts.

References

[1] Abowd, John, Francis Kramarz, and David N. Margolis (1999). "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251-333.

[2] Abowd, John M. Abowd, Francis Kramarz, Sébastien Pérez-Duarte and Ian M. Schmutte (2018). "Sorting Between and Within Industries: A Testable Model of Assortative Matching", *Annals of Economics and Statistics*, No. 129, pp. 1-32

[3] Acemoglu, Daron (2001). "Good Jobs Versus Bad Jobs", *Journal of Labor Economics*, 19, pp. 1-22.

[4] Albrecht, James, Pieter Gautier, and Susan Vroman (2014). "Efficient Entry and Competing Auctions," *American Economic Review,* 104(10), 3288-96.

[60] Andrews, M.J., Gill, L., Schank, T., Upward, R. (2008). "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?," *Journal of the Royal Statistical Society Series,* A171 (3), 673–97.

[5] Bagger, Jesper and Rasmus Lentz (2018). "An Empirical Model of Wage Dispersion with Sorting," *Review of Economic Studies (forthcoming).*

[59] Bartolucci, Cristian, Francesco Devicienti and Ignacio Monzón (2018). "Identifying Sorting in Practice," *American Economic Journal: Applied Economics (forthcoming).*

[6] Becker, Gary (1973), "A Theory of Marriage: Part I," *Journal of Political Economy*, 81(4), 813-46.

[7] Borovicková, Katarina, and Robert Shimer. 2017. "High Wage Workers Work for High Wage Firms." Mimeo.

[8] Burdett, Kenneth, Shouyong Shi and Randall Wright (2001) "Pricing and Matching with Frictions," *Journal of Political Economy,* 109(5), 1060-85.

[9] Eeckhout, Jan, and Philipp Kircher (2010). "Sorting and Decentralized Price Competition," *Econometrica*, 78(2), pp. 539-574.

[10] Eeckhout, Jan, and Philipp Kircher (2011). "Identifying Sorting - in Theory," *Review of Economic Studies* 78(3), 872–906.

[11] Hagedorn,Marcus, Tzuo Hann Law, and Iourii Manovskii (2017). "Identifying sorting," *Econometrica,* 85(1), 29-65.

[12] Haltiwanger, John, Henry Hyatt and Erika McEntarfer (2017). "Who Moves up the Job Ladder", NBER working paper no. 23693.

[13] Julien, Benoît, John Kennes and Ian King (2000)."Bidding for Labor," *Review of Economic Dynamics*, vol. 3(4), 619-49.

[14] Julien, Benoît, John Kennes and Ian King (2006a). "Residual Wage Disparity And Coordination Unemployment," *International Economic Review*, vol. 47(3), pages 961-89

[15] Kennes, John and Daniel le Maire (2010). "Coordination Frictions and Job Heterogeneity: A Discrete Time Analysis," Aarhus University working paper WP10-05.

[16] Lentz, Rasmus (2010). "Sorting by search intensity," *Journal of Economic Theory* 145(4), pp. 1436-52.

[17] Lentz, Rasmus, Suphanit Piyapromdee and Jean-Marc Robin (2018), "On Worker and firm Heterogeneity in Wages and Employment Mobility: Evidence from Danish Register Data," working paper,

[18] Lindenlaub, Ilse (2016). “Sorting Multidimensional Types: Theory and Application,” *Review of Economic Studies*, 84 (2), pp. 718-789.

[19] Lindenlaub, Ilse and Fabian Postel-Vinay (2017). “Multidimensional Sorting under Random Search,” working paper.

[20] Lise, Jeremy, Costas Meghir and Jean-Marc Robin (2016). “Matching, Sorting and Wages,” *Review of Economic Dynamics*, 19 (1), pp. 63-87.

[21] Lise, Jeremy and Fabian Postel-Vinay (2018). “Multidimensional Skills, Sorting, and Human Capital Accumulation,” working paper.

[22] Lopes del Melo, Rafael (2018). "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," *Journal of Political Economy (forthcoming).*

[23] McAfee, R. Preston (1993) “Mechanism Design by Competing Sellers,” *Econometrica*, Vol 61(6), pages 1281-1312.

[24] Mortensen, Dale (2003), Wage Dispersion: Why Are Similar Workers Paid Differently?, MIT Press.

[25] Peters, Michael, (1984) "Bertrand Equilibrium with Capacity Constraints and Restricted Mobility," *Econometrica*, 1117-27.

[26] Pissarides, Christopher A, (1994). "Search Unemployment with On-the-Job Search," *Review of Economic Studies*, 61(3),457-75.

[27] Postel-Vinay, Fabian and Jean-Marc Robin (2002). "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, vol. 70(6), pages 2295-2350.

[28] Shi, Shouyong (2001) "Frictional Assignment. I. Efficiency," *Journal of Economic Theory*, Elsevier, vol. 98(2), pages 232-260.

[29] Shimer, Robert (2005) "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," *Journal of Political Economy*, vol. 113(5), pages 996-1025.

[30] Shimer, Robert and Lones Smith (2000) "Assortative Matching and Search," *Econometrica*, vol. 68(2), pp. 343-69.

[31] Topkis, Donald M. (1998). “Supermodularity and Complementarity,” Princeton University Press.

[32] Woodcock, Simon D. (2015). "Match Effects," *Research in Economics*, vol. 69, pp. 100-121.

[33] Wright, Randall, Philipp Kircher, Benoît Julien, and Veronica Guerrieri (2017). "Directed Search: A Guided Tour", working paper.

A Theory Appendix

Derivation of equation (4)

We need to derive the conditional density of the second best offer, k_2, given a best offer, k_1.

$$g\left(k_2|k_1,\underline{k},h\right)=\frac{g\left(k_1,k_2|\underline{k},h\right)}{g^1\left(k_1|\underline{k},h\right)}$$

Therefore, we begin by deriving the joint density of the best and second best offer. Suppose a worker receives n job offers. The probability that j of these are below k_2 and that $n-j$ offers are above than k_1 where $\underline{k}\leq k_2\leq k_1\leq k^*$ is given by

$$\begin{pmatrix} n \\ j \end{pmatrix}\left(\frac{\phi\left(\underline{k}\right)-\phi\left(k_2|\underline{k},h\right)}{\phi\left(\underline{k}\right)}\right)^j\left(\frac{\phi\left(k_1|\underline{k},h\right)}{\phi\left(\underline{k}\right)}\right)^{n-j}$$

where $\begin{pmatrix} n \\ j \end{pmatrix}=\frac{n!}{(n-j)!j!}$. Taking the negative cross-derivative of this delivers the joint density of the j'th and $j+1$'th order statistics

$$\begin{aligned} g_{j,j+1}\left(k_1,k_2|n,\underline{k},h\right) &= \left[\begin{pmatrix} n \\ j \end{pmatrix} j\left(\frac{\phi\left(\underline{k}\right)-\phi\left(k_2|\underline{k},h\right)}{\phi\left(\underline{k}\right)}\right)^{j-1}(n-j)\left(\frac{\phi\left(k_1|\underline{k},h\right)}{\phi\left(\underline{k}\right)}\right)^{n-j-1}\frac{\phi'\left(k_1|\underline{k},h\right)}{\phi\left(\underline{k}\right)}\frac{\phi'\left(k_2|\underline{k},h\right)}{\phi\left(\underline{k}\right)}\right] \\ &= \frac{n!\left(\frac{\phi(\underline{k})-\phi(k_2|\underline{k},h)}{\phi(\underline{k})}\right)^{j-1}\left(\frac{\phi(k_1|\underline{k},h)}{\phi(\underline{k})}\right)^{n-j-1}\frac{\phi'(k_1|\underline{k},h)}{\phi(\underline{k})}\frac{\phi'(k_2|\underline{k},h)}{\phi(\underline{k})}}{(n-j-1)!\,(j-1)!} \end{aligned}$$

We are only interested in the best and second best offers, so we set $j=n-1$ and $j+1=n$. This gives us

$$g\left(k_1,k_2|n,\underline{k},h\right)=\frac{n!\left(\frac{\phi(\underline{k})-\phi(k_2|\underline{k},h)}{\phi(\underline{k})}\right)^{n-2}\frac{\phi'(k_1|\underline{k},h)}{\phi(\underline{k})}\frac{\phi'(k_2|\underline{k},h)}{\phi(\underline{k})}}{(n-2)!}$$

Summing over all possible number of job offers we obtain

$$
\begin{aligned}
g\left(k_1,k_2|\underline{k},h\right) &= \sum_{n=2}^{\infty} \frac{e^{-\phi(\underline{k})}\phi\left(\underline{k}\right)^n}{n!} \frac{n!\left(\frac{\phi(\underline{k})-\phi(k_2|\underline{k},h)}{\phi(\underline{k})}\right)^{n-2} \frac{\phi'(k_1|\underline{k},h)}{\phi(\underline{k})}\frac{\phi'(k_2|\underline{k},h)}{\phi(\underline{k})}}{(n-2)!} \\
&= \frac{\phi'\left(k_1\right)}{\phi\left(\underline{k}\right)}\frac{\phi'\left(k_2|\underline{k},h\right)}{\phi\left(\underline{k}\right)}\phi\left(\underline{k}\right)^2 e^{-\phi(\underline{k})}\sum_{n=2}^{\infty}\frac{\left(\phi\left(\underline{k}\right)-\phi\left(k_2|\underline{k},h\right)\right)^{n-2}}{(n-2)!} \\
&= \phi'\left(k_1|\underline{k},h\right)\phi'\left(k_2|\underline{k},h\right)e^{-\phi(\underline{k})}\sum_{n=2}^{\infty}\frac{\left(\phi\left(\underline{k}\right)-\phi\left(k_2|\underline{k},h\right)\right)^{n-2}}{(n-2)!} \\
&= \phi'\left(k_1|\underline{k},h\right)\phi'\left(k_2|\underline{k},h\right)e^{-\phi(k_2|\underline{k},h)}e^{-(\phi(\underline{k})-\phi(k_2|\underline{k},h))}\sum_{n=2}^{\infty}\frac{\left(\phi\left(\underline{k}\right)-\phi\left(k_2|\underline{k},h\right)\right)^{n-2}}{(n-2)!} \\
&= \phi'\left(k_1|\underline{k},h\right)\phi'\left(k_2|\underline{k},h\right)e^{-\phi(k_2|\underline{k},h)}\sum_{n=0}^{\infty}\frac{e^{-(\phi(\underline{k})-\phi(k_2|\underline{k},h))}\left(\phi\left(\underline{k}\right)-\phi\left(k_2|\underline{k},h\right)\right)^{n}}{n!} \\
&= \phi'\left(k_1|\underline{k},h\right)\phi'\left(k_2|\underline{k},h\right)e^{-\phi(k_2|\underline{k},h)}
\end{aligned}
$$

Using this, we can write

$$
\begin{aligned}
g\left(k_2|k_1,\underline{k},h\right) &= \frac{g\left(k_1,k_2|\underline{k},h\right)}{g^1\left(k_1|\underline{k},h\right)} \\
&= \frac{\phi'\left(k_1|\underline{k},h\right)\phi'\left(k_2|\underline{k},h\right)e^{-\phi(k_2|\underline{k},h)}}{-\phi'\left(k_1|\underline{k},h\right)e^{-\phi(k_1|\underline{k},h)}} \\
&= \frac{-\phi'\left(k_2|\underline{k},h\right)e^{-\phi(k_2|\underline{k},h)}}{e^{-\phi(k_1|\underline{k},h)}} \\
&= \frac{g^1\left(k_2|\underline{k},h\right)}{G^1\left(k_1|\underline{k},h\right)}
\end{aligned}
$$

A firm approaching a worker of type h will hire the worker with probability $G^1\left(k|\underline{k},h\right)$ and have the following expected profits

$$
\begin{aligned}
\pi\left(k|\underline{k},h\right) =& G^1\left(k|\underline{k},h\right)\int_{\underline{k}}^{k}\left(y\left(k,h\right)-y\left(z,h\right)\right)dG\left(z|k,\underline{k},h\right)-k \\
=& \int_{\underline{k}}^{k}\left(y\left(k,h\right)-y\left(z,h\right)\right)dG^1\left(z|\underline{k},h\right)-k
\end{aligned}
$$

Proof of proposition 1

The lowest job type in the worker's matching set is $\hat{k}(\underline{k}, h)$ and the productivity associated with this job is $y(\hat{k}(\underline{k}, h), h)$. Satisfaction of the free entry condition implies

$$e^{-\phi\left(\hat{k}(\underline{k},h),h\right)}\left(y\left(\hat{k}(\underline{k}, h), h\right) - y(\underline{k}, h)\right) = \hat{k}(\underline{k}, h)$$

Since the the lowest quality job type $\hat{k}(\underline{k}, h)$ for a worker of type h earns a positive return of $y(\hat{k}(\underline{k}, h), h)$ if and only if there is no other firm at the local market of this worker, $\phi(\hat{k}(\underline{k}, h), h)$ is also the measure of jobs greater than $\hat{k}(\underline{k}, h)$ for workers of type h. We then find $\hat{k}(\underline{k}, h)$ by solving for the maximum number of jobs consistent with free entry. Thus

$$\phi(\hat{k}(\underline{k}, h), h) = \arg\max\left\{\phi(k, h) \mid e^{-\phi(k,h)}(y(k, h) - y(\underline{k}, h)) = k\right\}$$

The total mass of jobs can also not exceed $\phi(\hat{k}(\underline{k}, h), h)$, because the argmax operator in equation (7) looks for the largest possible value of $\phi(k, h)$ that satisfies free entry of low type jobs. Likewise, the total mass of jobs cannot be less than $\phi(\hat{k}(\underline{k}, h), h)$ since this would imply positive profits for type $\hat{k}(\underline{k}, h)$ jobs.

The number of jobs above this threshold (and also the maximum job type) is derived as follows. It is useful to discretize the number of job types and take the limit as the number of job types gets large. The payoff of each firm type entering the matching market is a function of its productivity and the probability that it faces a competitor of type k. Given the distribution of firm types over a discrete set of firm types, we have a simple expression for the payoff of a type k firm. Thus the expected return of a type k_i firm who enters the type $n(\underline{k})$ submarket in the free entry equilibrium (when $\phi(k \mid \underline{k})$ is positive) is given by

$$k_i = \sum_{j=1}^{i}(y(k_i) - y(k_{j-1}))\exp(-\phi(k_i \mid \underline{k}))(1 - \exp(-(\phi(k_{j-1} \mid \underline{k}) - \phi(k_j \mid \underline{k}))))$$

where $k_1 = \hat{k}(\underline{k})$, $k_0 = \underline{k}$ and $\phi(\underline{k} \mid \underline{k}) = \infty$. Differencing the payoffs and opportunity costs of any pair of adjacent job types, we get the following difference equation

$$k_{j+1} - k_j = (y(k_{j+1}, h) - y(k_j, h))\, e^{-\phi(k_{j+1},h)} \tag{21}$$

which must be satisfied for all job types offered in equilibrium. Let $\triangle = k_j - k_{j\text{-}1}$ denote the interval of successive job types for which equation (21) holds. Our functional form for $\phi(k)$ follows by taking the limit of equation (21) as the interval $\triangle$ becomes small. We note that the function $\phi(k)$ gives a positive density of wages over the support of this job offer distribution if

$$\left[-\frac{y_{11}(k, h)}{(y_1(k))^2}\right][\log(y_1(k, h))] < 0.$$

Obviously, we require $y_{11}(k,h) < 0$.

Proof of proposition 2

It is straightforward to show that the higher bound of the matching set is increasing/ decreasing in h when the production function is supermodular/submodular. Instead, we consider the distribution of employment and the lower bound of the matching set.

Part 1: The distribution of employment.

The distribution of employed workers across firms of $k \geq \hat{k}(k,h)$ is given by

$$N(k \mid h) = \frac{G^1(k \mid h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)}{1 - G^1\left(\hat{k}(\underline{k},h) \mid h\right)}$$

Differentiating with respect to h gives us

$$\begin{aligned}
N_2(k \mid h) &= \frac{\begin{pmatrix} \left[G_2^1(k \mid h) - G_2^1\left(\hat{k}(\underline{k},h) \mid h\right)\right]\left[1 - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right] \\ + \left[G^1(k \mid h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right] G_2^1\left(\hat{k}(\underline{k},h) \mid h\right) \end{pmatrix}}{\left[1 - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right]^2} \\
&= \frac{G_2^1(k \mid h)\left[1 - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right] - G_2^1\left(\hat{k}(\underline{k},h) \mid h\right)\left[1 - G^1(k \mid h)\right]}{\left[1 - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right]^2} \\
&= \frac{\frac{y_{12}\left(\hat{k}(\underline{k},h),h\right)}{\left[y_1\left(\hat{k}(\underline{k},h),h\right)\right]^2}\left[1 - G^1(k \mid h)\right] - \frac{y_{12}(k,h)}{[y_1(k,h)]^2}\left[1 - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right]}{\left[1 - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right]^2}
\end{aligned}$$

Since we only consider $k \geq \hat{k}(\underline{k},h)$, we always have that $G^1(k \mid h) \geq G^1\left(\hat{k}(\underline{k},h) \mid h\right)$. If $\frac{y_{12}(k,h)}{y_1(k,h)}$ is increasing in k, we have that $\frac{y_{12}(k,h)}{[y_1(k,h)]^2} > \frac{y_{12}\left(\hat{k}(\underline{k},h),h\right)}{\left[y_1\left(\hat{k}(\underline{k},h),h\right)\right]^2}$. This will imply that $N_2(k \mid h) < 0$ such the distribution of employment for high type workers stochastically dominate the distribution of employment for lower worker types. When $\frac{y_{12}(k,h)}{y_1(k,h)} < 0$ is decreasing in k, $N_2(k \mid h) > 0$.

Part 2: The lower bound of the matching set.

The first-order condition of equation (7) is given by

$$y_1\left(\hat{k}(\underline{k},h),h\right)\hat{k}(\underline{k},h) = y\left(\hat{k}(\underline{k},h),h\right) - y(\underline{k},h)$$

Differentiate with respect to h

$$\left[y_{12}\left(\hat{k}\left(\underline{k},h\right),h\right)+y_{11}\left(\hat{k}\left(\underline{k},h\right),h\right)\hat{k}_2\left(\underline{k},h\right)\right]\hat{k}\left(\underline{k},h\right)+y_1\left(\hat{k}\left(\underline{k},h\right),h\right)\hat{k}_2\left(\underline{k},h\right)=$$

$$y_2\left(\hat{k}\left(\underline{k},h\right),h\right)-y_2\left(\underline{k},h\right)+y_1\left(\hat{k}\left(\underline{k},h\right),h\right)\hat{k}_2\left(\underline{k},h\right)\Leftrightarrow$$

$$\hat{k}_2\left(\underline{k},h\right)=\frac{1}{y_{11}\left(\hat{k}\left(\underline{k},h\right),h\right)}\left[\frac{y_2\left(\hat{k}\left(\underline{k},h\right),h\right)-y_2\left(\underline{k},h\right)}{\hat{k}\left(\underline{k},h\right)}-y_{12}\left(\hat{k}\left(\underline{k},h\right),h\right)\right]\Leftrightarrow$$

$$\hat{k}_2\left(\underline{k},h\right)=\frac{y_1\left(\hat{k}\left(\underline{k},h\right),h\right)}{y_{11}\left(\hat{k}\left(\underline{k},h\right),h\right)}\left[\frac{y_2\left(\hat{k}\left(\underline{k},h\right),h\right)-y_2\left(\underline{k},h\right)}{y\left(\hat{k}\left(\underline{k},h\right),h\right)-y\left(\underline{k},h\right)}-\frac{y_{12}\left(\hat{k}\left(\underline{k},h\right),h\right)}{y_1\left(\hat{k}\left(\underline{k},h\right),h\right)}\right]\Leftrightarrow \tag{22}$$

$$\hat{k}_2\left(\underline{k},h\right)=\frac{-y_1\left(\hat{k}\left(\underline{k},h\right),h\right)}{y_{11}\left(\hat{k}\left(\underline{k},h\right),h\right)}\left[\frac{y_{12}\left(\hat{k}\left(\underline{k},h\right),h\right)}{y_1\left(\hat{k}\left(\underline{k},h\right),h\right)}-\frac{\int_{\underline{k}}^{\hat{k}(\underline{k},h)}y_{12}\left(z,h\right)dz}{\int_{\underline{k}}^{\hat{k}(\underline{k},h)}y_1\left(x,h\right)dx}\right]$$

The first term is positive because $y_{11}\left(\hat{k}\left(\underline{k},h\right),h\right)<0$. Define a function $\Psi(k,h)\equiv\frac{y_{12}(k,h)}{y_1(k,h)}$. We can then write

$$\hat{k}_2\left(\underline{k},h\right)=\frac{-y_1\left(\hat{k}\left(\underline{k},h\right),h\right)}{y_{11}\left(\hat{k}\left(\underline{k},h\right),h\right)}\left[\frac{\Psi(\hat{k}\left(\underline{k},h\right),h)y_1\left(\hat{k}\left(\underline{k},h\right),h\right)}{y_1\left(\hat{k}\left(\underline{k},h\right),h\right)}-\frac{\int_{\underline{k}}^{\hat{k}(\underline{k},h)}\Psi(z,h)y_1\left(z,h\right)dz}{\int_{\underline{k}}^{\hat{k}(k,h)}y_1\left(x,h\right)dx}\right]\Leftrightarrow$$

$$\hat{k}_2\left(\underline{k},h\right)=\frac{-y_1\left(\hat{k}\left(\underline{k},h\right),h\right)}{y_{11}\left(\hat{k}\left(\underline{k},h\right),h\right)}\left[\Psi(\hat{k}\left(\underline{k},h\right),h)-\int_{\underline{k}}^{\hat{k}(\underline{k},h)}\Psi(z,h)\frac{y_1\left(z,h\right)dz}{\int_{\underline{k}}^{\hat{k}(k,h)}y_1\left(x,h\right)dx}\right]$$

When $\Psi(k,h)\equiv\frac{y_{12}(k,h)}{y_1(k,h)}$ is non-decreasing in k, the second term must be positive since the last term of the squared bracket is simply a weighted average of $\Psi(k,h)$ in between $\underline{k}$ and $\hat{k}\left(\underline{k},h\right)$. Hence, when $\Psi(k,h)$ is non-decreasing in k, we have that $\hat{k}_2\left(\underline{k},h\right)>0$, while $\hat{k}_2\left(\underline{k},h\right)<0$ if $\Psi(k,h)$ is non-increasing in k.

Furthermore, from equation 22, it is clear that setting $y\left(\underline{k},h\right)=0$ for all h implies

$$\hat{k}_2\left(\underline{k},h\right)=\frac{y_1\left(\hat{k}\left(\underline{k},h\right),h\right)}{y_{11}\left(\hat{k}\left(\underline{k},h\right),h\right)}\left[\frac{y_2\left(\hat{k}\left(\underline{k},h\right),h\right)}{y\left(\hat{k}\left(\underline{k},h\right),h\right)}-\frac{y_{12}\left(\hat{k}\left(\underline{k},h\right),h\right)}{y_1\left(\hat{k}\left(\underline{k},h\right),h\right)}\right]\Leftrightarrow$$

$$\hat{k}_2\left(\underline{k},h\right)=\frac{-1}{y\left(\hat{k}\left(\underline{k},h\right),h\right)y_{11}\left(\hat{k}\left(\underline{k},h\right),h\right)}\left[y\left(\hat{k}\left(\underline{k},h\right),h\right)y_{12}\left(\hat{k}\left(\underline{k},h\right),h\right)-y_1\left(\hat{k}\left(\underline{k},h\right),h\right)y_2\left(\hat{k}\left(\underline{k},h\right),h\right)\right]$$

which is positive/negative if the production function is log-supermodular/log-submodular.

Proof of constrained efficiency

To prove that the decentralized solution is constrained efficient, we need only to show that the solution to the planner's problem for a particular firm of type k approaching a type h worker employed in a type k_1 firm correspond to the free entry condition. We consider the planner's problem for the case of discrete firm distribution. Define$\zeta(k,h) = \phi(k,h) - \phi(k-1,h)$ such that $\phi(k,h) = \sum_{z=k}^{k^*(h)} \zeta(z,h)$. Then, we can write the planner's problem as

$$\Lambda(k_1,h)\, e^{-\phi\left(\hat{k}(k_1,h),h\right)} + \sum_{z=\hat{k}(k_1,h)}^{k^*(h)} \left[\Lambda(z,h)\, e^{-\phi(z+1,h)} \left[1 - e^{-\zeta(z,h)}\right] - \zeta(z,h)\, z\right]$$

$$\Lambda(k_1,h)\, e^{-\phi\left(\hat{k}(k_1,h),h\right)} + \sum_{z=\hat{k}(k_1,h)}^{k^*(h)} \left[\Lambda(z,h) \left(e^{-\phi(z+1,h)} - e^{-\phi(z,h)}\right) - \zeta(z,h)\, z\right]$$

Differentiate with respect to $\zeta(k,h)$

$$-\Lambda(k_1,h)\, e^{-\phi\left(\hat{k}(k_1,h),h\right)} - \sum_{z=\hat{k}(k_1,h)}^{k-1} \Lambda(z,h) \left(e^{-\phi(z+1,h)} - e^{-\phi(z,h)}\right) + \Lambda(k,h)\, e^{-\phi(k,h)} - k = 0 \Leftrightarrow$$

$$\left[\Lambda(z,h) - \Lambda(k_1,h)\right] e^{-\phi\left(\hat{k}(k_1,h),h\right)} - \sum_{z=\hat{k}(k_1,h)}^{k-1} \left[\Lambda(k,h) - \Lambda(z,h)\right] \left(e^{-\phi(z+1,h)} - e^{-\phi(z,h)}\right) - k = 0 \Leftrightarrow$$

In the limit where the distribution of firms is continuous, we have

$$\left[\Lambda(z,h) - \Lambda(k_1,h)\right] e^{-\phi\left(\hat{k}(k_1,h),h\right)} - \int_{\hat{k}(k_1,h)}^{k} \left[\Lambda(k,h) - \Lambda(z,h)\right] \left[-\phi_1(z,h)\, e^{-\phi(z,h)}\right] dz - k = 0 \Leftrightarrow$$

$$\left[\Lambda(k,h) - \Lambda(k_1,h)\right] G^1\left(\hat{k}(k_1,h)\,|k_1,h\right) - \int_{\hat{k}(k_1,h)}^{k} \left[\Lambda(k,h) - \Lambda(z,h)\right] dG^1(z|k_1,h) - k = 0$$

Hence, the planner's first order condition is identical to the firm's free-entry condition in equation (10).

Proof of lemma 2

The job value function for a worker of type h employed in a type $k \geq p^*$ firm so $G^1\left(\hat{k}(k,h) \mid h\right) = 1$ and with a second-best offer also being k is given by

$$\Lambda(k,h) = y(k,h) + \beta(1-\delta)\Lambda(k,h) + \beta\delta W(\underline{k},\underline{k},h)$$

Differentiating with respect to k and re-arranging gives

$$y_1(k,h) = (1 - \beta(1-\delta))\, \Lambda_1(k,h)$$

For $k < p^*$ consider first equation (11) for a worker employed in a type k firm and with a second-best offer also being k

$$W(k,k,h) = \Lambda(k)\, G^2\left(\hat{k}(k,h) \mid h\right) + \int_{\hat{k}(k,h)}^{k^*(h)} \Lambda(z,h)\, dG^2(z \mid h)$$

Differentiating this with respect to k gives

$$\begin{aligned} W_1(k,k,h) &= \Lambda_1(k,h)\, G^2\left(\hat{k}(k,h) \mid h\right) + \Lambda(k,h)\, G_1^2\left(\hat{k}(k,h) \mid h\right) \hat{k}_1(k,h) \\ &\quad - \Lambda\left(\hat{k}_1(k,h), h\right) G_1^2\left(\hat{k}(k,h) \mid h\right) \hat{k}_1(k,h) \\ &= \Lambda_1(k,h)\, G^2\left(\hat{k}(k,h) \mid h\right) + \left(\Lambda(k,h) - \Lambda\left(\hat{k}_1(k,h), h\right)\right) G_1^2\left(\hat{k}(k,h) \mid h\right) \hat{k}_1(k,h) \end{aligned}$$

Using this in the differentiated version of equation (12) for a worker employed in a type k firm and with a second-best offer also being k delivers

$$\begin{aligned} \mathit{\Lambda}_1(k,h) &= y_1(k,h) + \beta(1-\delta) \left[\begin{array}{c} \mathit{\Lambda}_1(k,h)\, G^2\left(\hat{k}(k,h) \mid h\right) \\ + \left(\Lambda(k,h) - \Lambda\left(\hat{k}(k,h), h\right)\right) G_1^2\left(\hat{k}(k,h) \mid h\right) \hat{k_1}(k,h) \end{array} \right] \\ &= y_1(k,h) + \beta(1-\delta) \left[\begin{array}{c} \mathit{\Lambda}_1(k,h)\, G^2\left(\hat{k}(k,h) \mid h\right) \\ -\mathit{\Lambda}_1\left(\hat{k}(k,h), h\right) G_1^2\left(\hat{k}(k,h) \mid h\right) \hat{k}(k,h)\, \hat{k_1}(k,h) \end{array} \right] \end{aligned} \tag{23}$$

where we have used that $\hat{k}_1(k,h) = \frac{\mathit{\Lambda}\left(\hat{k}(k,h),h\right) - \mathit{\Lambda}(k,h)}{\mathit{\Lambda}_1\left(\hat{k}(k,h),h\right)}$. Differentiating $\hat{k}_1(k,h)$ with respect to k delivers

$$\begin{aligned} \hat{k}_1(k,h)\, \Lambda_1\left(\hat{k}(k,h), h\right) + \hat{k}(k,h)\, \Lambda_{11}\left(\hat{k}(k,h), h\right) \hat{k}_1(k,h) &= \Lambda_1\left(\hat{k}(k,h), h\right) \hat{k}_1(k,h) - \Lambda_1(k,h) \Leftrightarrow \\ \hat{k}(k,h)\, \Lambda_{11}\left(\hat{k}(k,h), h\right) \hat{k}_1(k,h) &= -\Lambda_1(k,h) \Leftrightarrow \\ \hat{k}(k,h)\, \hat{k}_1(k_1,h) &= \frac{-\Lambda_1(k,h)}{\Lambda_{11}\left(\hat{k}(k,h), h\right)} \end{aligned} \tag{24}$$

Besides this, we also plug in $G^2\left(\hat{k}(k,h)\mid h\right)=\frac{1+\log\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)}{\Lambda_1\left(\hat{k}(k,h),h\right)}$ and $G_1^2\left(\hat{k}(k,h)\mid h\right)=-\frac{\Lambda_{11}\left(\hat{k}(k,h),h\right)\log\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)}{\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)^2}$ in equation (23) and re-arrange to obtain

$$\begin{aligned}
\Lambda_1(k,h) &= y_1(k,h)+\beta(1-\delta)\begin{bmatrix} \Lambda_1(k,h)\frac{1+\log\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)}{\Lambda_1\left(\hat{k}(k,h),h\right)} \\ -\Lambda_1\left(\hat{k}(k,h),h\right)\frac{-\Lambda_{11}\left(\hat{k}(k,h),h\right)\log\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)}{\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)^2}\frac{-\Lambda_1(k,h)}{\Lambda_{11}\left(\hat{k}(k,h),h\right)}\end{bmatrix}\Leftrightarrow \\
y_1(k,h) &= \Lambda_1(k,h)-\beta(1-\delta)\Lambda_1(k,h)\begin{bmatrix}\frac{1+\log\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)}{\Lambda_1\left(\hat{k}(k,h),h\right)} \\ -\left(\frac{\log\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)}{\Lambda_1\left(\hat{k}(k,h),h\right)}\right)\end{bmatrix}\Leftrightarrow \\
y_1(k,h) &= \left(1-\beta(1-\delta)\frac{1}{\Lambda_1\left(\hat{k}(k,h),h\right)}\right)\Lambda_1(k,h) \qquad (25)
\end{aligned}$$

Differentiating this with respect to k yields

$$y_{11}(k,h)=\left(1-\beta(1-\delta)\frac{1}{\Lambda_1\left(\hat{k}(k,h),h\right)}\right)\Lambda_{11}(k,h)+\beta(1-\delta)\frac{\Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_1(k,h)}{\Lambda_1\left(\hat{k},h\right)}\Lambda_1(k,h)$$

Then, using the equation for $\hat{k}_1(k,h)$ and (24), we obtain

$$\begin{aligned}
y_{11}(k,h) &= \left(1-\beta(1-\delta)\frac{1}{\Lambda_1\left(\hat{k}(k,h),h\right)}\right)\Lambda_{11}(k,h) \\
&\quad +\beta(1-\delta)\frac{\Lambda_{11}\left(\hat{k}(k,h),h\right)}{\Lambda_1\left(\hat{k}(k,h),h\right)\hat{k}(k,h)}\frac{-\Lambda_1(k,h)}{\Lambda_{11}\left(\hat{k}(k,h),h\right)}\Lambda_1(k,h) \\
&= \left(1-\beta(1-\delta)\frac{1}{\Lambda_1\left(\hat{k}(k,h),h\right)}\right)\Lambda_{11}(k,h)-\beta(1-\delta)\frac{[\Lambda_1(k,h)]^2}{\Lambda\left(\hat{k}(k,h),h\right)-\Lambda(k,h)}
\end{aligned}$$

where both terms on the right hand side are negative.

Proof of lemma 3.

The proof consists of four parts. In the first and second parts, we sign $\frac{\partial w(k_1,k_2,h)}{\partial k_1}$ and $\frac{\partial w(k_1,k_2,h)}{\partial k_2}$. Next, we show that the k_1 and k_2 are positively correlated. Finally, we show that for any worker type h, the expected wage in a k_1 firm is higher than the expected wage in a $k_1' < k_1$ firm.

Part 1: Signing $w_1(k_1,k_2,h)$

Differentiating equation (15) with respect to k_1 delivers

$$
\begin{aligned}
w_1(k_1,k_2,h) &= y_1(k_1,h) + \beta(1-\delta) G_1^1\left(\hat{k}(k_1,h) \mid h\right) \hat{k}_1(k_1,h)\left[\Lambda(k_1,h) - \Lambda(k_2,h)\right] \\
&\quad - \left[1 - \beta(1-\delta) G^1\left(\hat{k}(k_1,h) \mid h\right)\right] \Lambda_1(k_1,h)
\end{aligned}
$$

Using equation (25) to substitute $y_1(k_1,h)$, we can write

$$
w_1(k_1,k_2,h) = \beta(1-\delta) G_1^1\left(\hat{k}(k_1,h) \mid h\right) \hat{k}_1(k_1,h)\left[\Lambda(k_1,h) - \Lambda(k_2,h)\right] \tag{26}
$$

In order to show that this is positive, we need to show that $\hat{k}_1(k_1,h) > 0$. In the derivation for equation (25) we derived that $\hat{k}(k,h)\hat{k}_1(k_1,h) = \frac{-\Lambda_1(k,h)}{\Lambda_{11}\left(\hat{k}(k,h),h\right)}$, so $\hat{k}_1(k_1,h) > 0$ since $\Lambda_{11}\left(\hat{k}(k,h),h\right) < 0$. Obviously, $G_1^1\left(\hat{k}(k_1,h) \mid h\right) > 0$ so we always have that $w_1(k_1,k_2,h) > 0$.

Part 2: Signing $w_2(k_1,k_2,h)$.

Differentiate equation (15) with respect to k_2

$$
w_2(k_1,k_2,h) = \left[1 - \beta(1-\delta) G^1\left(\hat{k}(k_1,h) \mid h\right)\right] \Lambda_1(k_2,h) \tag{27}
$$

which is always positive.

Part 3: Showing that k_1 and k_2 are positively correlated.

Consider an unemployed worker searching for a job. Conditional on becoming employed in a firm at k_1 or below, we want to examine the relationship between the first and second best offers, k_1 and k_2. To begin with, we examine the first period of employment after being unemployed. Below, we also consider the relationship between k_1 and k_2 after search while employed. We derive the conditional distribution function of k_2 given k_1 to show that the conditional distribution of k_2 for a given k_1 stochastically dominates the conditional distribution of k_2 for a $k_1' < k_1$.

We want to calculate the probability of leaving unemployment to a job with best and second best offers (k_1,k_2), where $k_1 \geq k_2$. When leaving unemployment, we know that the best offer is at $\hat{k}(\underline{k},h)$ or above, whereas the second best offer only needs to be at $\underline{k}$ or above. The probability is given by

$$
G(k_1,k_2 \mid h) - G\left(\hat{k}(\underline{k}), \hat{k}(\underline{k}) \mid h\right) = \left(1 + \phi(k_2,h) - \phi(k_1,h)\right)\exp\left(-\phi(k_2,h)\right) - \exp\left(-\phi\left(\hat{k}(\underline{k},h),h\right)\right)
$$

The conditional distribution of k_2 given k_1 and $k_1 \geq \hat{k}(\underline{k})$ is given by

$$
\begin{aligned}
G\left(k_2 \mid k_1, k_1 \geq \hat{k}(\underline{k}), h\right) &= \frac{G(k_1, k_2 \mid h) - G\left(\hat{k}(\underline{k}, h), \hat{k}(\underline{k}, h) \mid h\right)}{G(k_1, k_1 \mid h) - G\left(\hat{k}(\underline{k}, h), \hat{k}(\underline{k}, h) \mid h\right)} \\
&= \frac{(1 + \phi(k_2, h) - \phi(k_1, h)) \exp(-\phi(k_2, h)) - \exp\left(-\phi\left(\hat{k}(\underline{k}, h), h\right)\right)}{\exp(-\phi(k_1, h)) - \exp\left(-\phi\left(\hat{k}(\underline{k}, h), h\right)\right)}
\end{aligned}
$$

Since $\phi(k, h) = \log(\Lambda_1(k, h))$, we have that

$$
G\left(k_2 \mid k_1, k_1 \geq \hat{k}(\underline{k}), h\right) = \frac{(1 + \log(\Lambda_1(k_2, h)) - \log(\Lambda_1(k_1, h))) \frac{1}{\Lambda_1(k_2,h)} - \frac{1}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}}{\frac{1}{\Lambda_1(k_1,h)} - \frac{1}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}}
$$

Differentiating with respect to k_1 delivers

$$
\begin{aligned}
& G_2\left(k_2 \mid k_1 \geq \hat{k}(\underline{k}, h), h\right) \\
&= \frac{-\frac{\Lambda_{11}(k_1,h)}{\Lambda_1(k_1,h)} \frac{1}{\Lambda_1(k_2,h)} \left[\frac{1}{\Lambda_1(k_1,h)} - \frac{1}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}\right] + \frac{\Lambda_{11}(k_1,h)}{\Lambda_1(k_1,h)} \left[(1 + \log(\Lambda_1(k_2, h)) - \log(\Lambda_1(k_1, h))) \frac{1}{\Lambda_1(k_2,h)} - \frac{1}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}\right]}{\left[\frac{1}{\Lambda_1(k_1,h)} - \frac{1}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}\right]^2} \\
&= \frac{\Lambda_{11}(k_1, h)}{\Lambda_1(k_1, h)} \left(\frac{-\frac{1}{\Lambda_1(k_2,h)} \left[\frac{1}{\Lambda_1(k_1,h)} - \frac{1}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}\right] + \left[(1 + \log(\Lambda_1(k_2, h)) - \log(\Lambda_1(k_1, h))) \frac{1}{\Lambda_1(k_2,h)} - \frac{1}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}\right]}{\left[\frac{1}{\Lambda_1(k_1,h)} - \frac{1}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}\right]^2}\right) \\
&= \frac{\Lambda_{11}(k_1, h)}{\Lambda_1(k_1, h)} \frac{1}{\left[\frac{1}{\Lambda_1(k_1,h)} - \frac{1}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}\right]^2} \left(\left[\frac{1}{\Lambda_1(k_1, h)} - \frac{1}{\Lambda_1\left(\hat{k}(\underline{k}, h), h\right)}\right] \left(1 - \frac{1}{\Lambda_1(k_2, h)}\right) + \frac{\log(\Lambda_1(k_2, h)) - \log(\Lambda_1(k_1, h))}{\Lambda_1(k_2, h)}\right) \\
&= \frac{\Lambda_{11}(k_1, h)}{\Lambda_1(k_1, h) \Lambda_1(k_2, h)} \frac{1}{\left[\frac{1}{\Lambda_1(k_1,h)} - \frac{1}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}\right]^2} \left(\left[\frac{1}{\Lambda_1(k_1, h)} - \frac{1}{\Lambda_1\left(\hat{k}(\underline{k}, h), h\right)}\right] (\Lambda_1(k_2, h) - 1) + \log\left(\frac{\Lambda_1(k_2, h)}{\Lambda_1(k_1, h)}\right)\right)
\end{aligned}
$$

which is negative since all terms on the r.h.s. are positive besides the first term, $\frac{\Lambda_{11}(k_1,h)}{\Lambda_1(k_1,h)\Lambda_1(k_2,h)} < 0$. This shows that the distribution of k_2 given k_1 stochastically dominates the distribution of k_2 given k_1' for $k_1' < k_1$.

We want to extend this result of the distribution of k_2 given k_1 following any period of employment. There are three cases to consider. Let the workers' current best offer be given by k. First, with the event that the workers receive no offers above $\hat{k}(k, h)$, nothing happens and the stochastic dominance result from the unemployed search still holds. Second, when a worker receives two or more offers above $\hat{k}(k, h)$, we can simply replace $\underline{k}$ with k in the proof above and, hence, the stochastic dominance result holds. Third, when the worker only receives one offer above $\hat{k}(k, h)$, the new offer becomes k_1 and $k_2 = k$. This is straightforward to show since the conditional distribution of k_1 given k_2 (i.e. the previous period's best offer) when only getting a single new offer at $\hat{k}(k, h)$ or above is

$$
F(k_1 \mid k_2, h) = \frac{G^1(k_1 \mid h) - G^1\left(\hat{k}(k_2, h) \mid h\right)}{1 - G^1\left(\hat{k}(k_2, h) \mid h\right)}
$$

Differentiating with respect to k_2 delivers

$$\begin{aligned}
&\frac{\partial F\left(k_1 \mid k_2, h\right)}{\partial k_2} \\
&= \frac{-G_1^1\left(\hat{k}\left(k_2,h\right) \mid h\right)\hat{k}_1\left(k_2,h\right)\left[1-G^1\left(\hat{k}\left(k_2,h\right) \mid h\right)\right]+G_1^1\left(\hat{k}\left(k_2,h\right) \mid h\right)\hat{k}_1\left(k_2,h\right)\left[G^1\left(k_1 \mid h\right)-G^1\left(\hat{k}\left(k_2,h\right) \mid h\right)\right]}{\left[1-G^1\left(\hat{k}\left(k_2,h\right) \mid h\right)\right]^2} \\
&= \frac{G_1^1\left(\hat{k}\left(k_2,h\right) \mid h\right)\hat{k}_1\left(k_2,h\right)\left[-1+G^1\left(\hat{k}\left(k_2,h\right) \mid h\right)+G^1\left(k_1 \mid h\right)-G^1\left(\hat{k}\left(k_2,h\right) \mid h\right)\right]}{\left[1-G^1\left(\hat{k}\left(k_2,h\right) \mid h\right)\right]^2} \\
&= -\frac{G_1^1\left(\hat{k}\left(k_2,h\right) \mid h\right)\hat{k}_1\left(k_2,h\right)\left[1-G^1\left(k_1 \mid h\right)\right]}{\left[1-G^1\left(\hat{k}\left(k_2,h\right) \mid h\right)\right]^2} < 0
\end{aligned}$$

Part 4: With the results established in part 1-3 of this proof, we can now examine whether the expected wage in a k_1 firm is higher than in a $k_1' < k_1$ firm (conditional on worker type, h). Denoting the conditional density of second best offers by $f\left(k_2 \mid k_1, h\right)$, we can write this as

$$\begin{aligned}
& \int_{\underline{k}}^{k_1} w\left(k_1,z,h\right) f\left(z \mid k_1,h\right) dz - \int_{\underline{k}}^{k_1'} w\left(k_1',z,h\right) f\left(z \mid k_1',h\right) dz \\
= & \int_{\underline{k}}^{k_1} w\left(k_1,z,h\right) f\left(z \mid k_1,h\right) dz - \int_{\underline{k}}^{k_1} w\left(k_1',z,h\right) f\left(z \mid k_1',h\right) dz \\
= & \int_{\underline{k}}^{k_1} w\left(k_1,z,h\right) f\left(z \mid k_1,h\right) dz - \int_{\underline{k}}^{k_1} w\left(k_1',z,h\right) f\left(z \mid k_1',h\right) dz \\
& + \int_{\underline{k}}^{k_1} w\left(k_1',z,h\right) f\left(z \mid k_1,h\right) dz - \int_{\underline{k}}^{k_1} w\left(k_1',z,h\right) f\left(z \mid k_1,h\right) dz \\
= & \int_{\underline{k}}^{k_1} \left[w\left(k_1,z,h\right) - w\left(k_1',z,h\right)\right] f\left(z \mid k_1,h\right) dz + \int_{\underline{k}}^{k_1} w\left(k_1',z,h\right)\left[f\left(z \mid k_1,h\right) - f\left(z \mid k_1',h\right)\right] dz
\end{aligned}$$

The first term on the r.h.s. is positive because when $k_1 > k_1'$ we have that $w\left(k_1,k_2,h\right) > w\left(k_1',k_2,h\right)$. The second term is positive since $w\left(k_1',k_2,h\right)$ is increasing in k_2 and since the distribution of second best offers given k_1 stochastically dominates the distribution of second best offers given $k_1' < k_1$, i.e. $F\left(z \mid k_1,h\right) \leq F\left(z \mid k_1',h\right)$.

Proof of lemma 4.

The proof is divided into the case of $\Lambda_{12}\left(k,h\right) > 0$ and the case of $\Lambda_{12}\left(k,h\right) \leq 0$.

1) Assume that $\Lambda_{12}(k,h) > 0$ for all (k,h):

$$
\begin{aligned}
W(k,k,h) &= \Lambda(k,h)\,G^2\left(\hat{k}(k,h)\,|h\right) + \int_{\hat{k}(k,h)}^{k^*(h)} \Lambda(z,h)\,dG^2(z|h) \\
&= \Lambda(k,h)\,G^2\left(\hat{k}(k,h)\,|h\right) + \left[\Lambda(z,h)\,G^2(z|h)\right]_{\hat{k}(k,h)}^{k^*(h)} - \int_{\hat{k}(k,h)}^{k^*(h)} \Lambda_1(z,h)\,G^2(z|h)\,dz \\
&= \Lambda(k,h)\,G^2\left(\hat{k}(k,h)\,|h\right) + \Lambda(k^*(h),h) - \Lambda\left(\hat{k}(k,h),h\right)G^2\left(\hat{k}(k,h)\,|h\right) - \int_{\hat{k}(k,h)}^{k^*(h)} \Lambda_1(z,h)\,G^2(z|h)\,dz \\
&= \Lambda(k,h)\,G^2\left(\hat{k}(k,h)\,|h\right) + \Lambda\left(\hat{k}(k,h),h\right)\left[1 - G^2\left(\hat{k}(k,h)\,|h\right)\right] + \left[\Lambda(k^*(h),h) - \Lambda\left(\hat{k}(k,h),h\right)\right] \\
&\quad - \int_{\hat{k}(k,h)}^{k^*(h)} \Lambda_1(z,h)\,G^2(z|h)\,dz \\
&= \Lambda(k,h)\,G^2\left(\hat{k}(k,h)\,|h\right) + \Lambda\left(\hat{k}(k,h),h\right)\left[1 - G^2\left(\hat{k}(k,h)\,|h\right)\right] \\
&\quad + \int_{\hat{k}(k,h)}^{k^*(h)} \Lambda_1(z,h)\left[1 - G^2(z|h)\right]dz
\end{aligned}
$$

Differentiate with respect to h

$$
\begin{aligned}
W_3(k,k,h) &= \Lambda_2(k,h)\,G^2\left(\hat{k}(k,h)\,|h\right) + \Lambda(k,h)\,\frac{\partial G^2\left(\hat{k}(k,h)\,|h\right)}{\partial h} + \Lambda_1\left(\hat{k}(k,h),h\right)\left[1 - G^2\left(\hat{k}(k,h)\,|h\right)\right]\hat{k}_2(k,h) \\
&\quad + \Lambda_2\left(\hat{k}(k,h),h\right)\left[1 - G^2\left(\hat{k}(k,h)\,|h\right)\right] - \Lambda\left(\hat{k}(k,h),h\right)\frac{\partial G^2\left(\hat{k}(k,h)\,|h\right)}{\partial h} \\
&\quad + \Lambda_1(k^*(h),h)\left[1 - G^2(k^*(h)\,|h)\right]k_1^*(h) - \Lambda_1\left(\hat{k}(k,h),h\right)\left[1 - G^2\left(\hat{k}(k,h)\,|h\right)\right]\hat{k}_2(k,h) \\
&\quad + \int_{\hat{k}(k,h)}^{k^*(h)} \Lambda_{12}(z,h)\left[1 - G^2(z|h)\right]dz - \int_{\hat{k}(k,h)}^{k^*(h)} \Lambda_1(z,h)\,\frac{\partial G^2(z|h)}{\partial h}dz \\
&= \Lambda_2(k,h)\,G^2\left(\hat{k}(k,h)\,|h\right) + \left[\Lambda(k,h) - \Lambda\left(\hat{k}(k,h),h\right)\right]\frac{\partial G^2\left(\hat{k}(k,h)\,|h\right)}{\partial h} \\
&\quad + \Lambda_2\left(\hat{k}(k,h),h\right)\left[1 - G^2\left(\hat{k}(k,h)\,|h\right)\right] + \int_{\hat{k}(k,h)}^{k^*(h)} \Lambda_{12}(z,h)\left[1 - G^2(z|h)\right]dz \\
&\quad - \int_{\hat{k}(k,h)}^{k^*(h)} \Lambda_1(z,h)\,\frac{\partial G^2(z|h)}{\partial h}dz
\end{aligned}
$$

The first and third terms are positive since $\Lambda_2(k,h) > 0$. The second term is positive since $\left[\Lambda(k,h) - \Lambda\left(\hat{k}(k,h),h\right)\right] < 0$ and $\frac{\partial G^2\left(\hat{k}(k,h)|h\right)}{\partial h} < 0$. The fourth term is positive as $\Lambda_{12}(k,h) > 0$, while the fifth term is positive because $\frac{\partial G^2(z|h)}{\partial h} < 0$. Hence, $W_3(k,k,h) > 0$ when $\Lambda_{12}(k,h) > 0$ for all (k,h) (as well as $\Lambda_1(k,h) > 0$, and $\Lambda_2(k,h) > 0$ for all (k,h)).

2) Assume that $\Lambda_{12}(k,h) \leq 0$ for all (k,h):

$W(k,k,h)$ is given by

$$\begin{aligned}
W(k,k,h) &= \Lambda(k,h)\,G^2\left(\hat{k}(k,h)\,|h\right) + \int_{\hat{k}(k,h)}^{k^*(h)} \Lambda(z,h)\,dG^2(z|h) \\
&= \Lambda(k,h)\,G^2\left(\hat{k}(k,h)\,|h\right) + \left[\Lambda(z,h)\,G^2(z|h)\right]_{\hat{k}(k,h)}^{k^*(h)} - \int_{\hat{k}(k,h)}^{k^*(h)} \Lambda_1(z,h)\,G^2(z|h)\,dz \\
&= \Lambda(k,h)\,G^2\left(\hat{k}(k,h)\,|h\right) + \left[\Lambda(z,h)\,G^2(z|h)\right]_{\hat{k}(k,h)}^{k^*(h)} - \int_{\hat{k}(k,h)}^{k^*(h)} (1+\log(\Lambda_1(z,h)))\,dz \\
&= \Lambda(k^*(h),h) + \left[\Lambda(k,h) - \Lambda\left(\hat{k}(k,h),h\right)\right] G^2\left(\hat{k}(k,h)\right) - \int_{\hat{k}(k,h)}^{k^*(h)} (1+\log(\Lambda_1(z,h)))\,dz \\
&= \Lambda(k^*(h),h) + \left[\Lambda(k,h) - \Lambda\left(\hat{k}(k,h),h\right)\right] \frac{1+\log\left(\Lambda_1\left(\hat{k}(k,h)\right)\right)}{\Lambda_1\left(\hat{k}(k,h),h\right)} - \int_{\hat{k}(k,h)}^{k^*(h)} (1+\log(\Lambda_1(z,h)))\,dz \\
&= \Lambda(k^*(h),h) - \hat{k}(k,h)\left(1+\log\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)\right) - \int_{\hat{k}(k,h)}^{k^*(h)} (1+\log(\Lambda_1(z,h)))\,dz
\end{aligned}$$

where we have used that $\hat{k}(k,h) = \frac{\Lambda\left(\hat{k}(k,h),h\right)-\Lambda(k,h)}{\Lambda_1\left(\hat{k}(k,h),h\right)}$ and $G^2(k|h) = \frac{1+\log(\Lambda_1(k,h))}{\Lambda_1(k,h)}$. Then, differentiating $W(k,k,h)$ with respect to h delivers

$$\begin{aligned}
W_3(k,k,h) &= \Lambda_1(k^*(h),h)\,k_1^*(h) + \Lambda_2(k^*(h),h) - \hat{k}_2(k,h)\left(1+\log\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)\right) \\
&- \hat{k}(k,h)\,\frac{\Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) + \Lambda_{12}\left(\hat{k}(k,h),h\right)}{\Lambda_1\left(\hat{k}(k,h),h\right)} \\
&- (1+\log(\Lambda_1(k^*(h),h)))\,k_1^*(h) + \left(1+\log\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)\right)\hat{k}_2(k,h) - \int_{\hat{k}(k,h)}^{k^*(h)} \frac{\Lambda_{12}(z,h)}{\Lambda_1(z,h)}dz \\
&= k_1^*(h) + \Lambda_2(k^*(h),h) - \hat{k}(k,h)\,\frac{\Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) + \Lambda_{12}\left(\hat{k}(k,h),h\right)}{\Lambda_1\left(\hat{k}(k,h),h\right)} \\
&- k_1^*(h) - \int_{\hat{k}(k,h)}^{k^*(h)} \frac{\Lambda_{12}(z,h)}{\Lambda_1(z,h)}dz \\
&= \Lambda_2(k^*(h),h) - \hat{k}(k,h)\,\frac{\Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) + \Lambda_{12}\left(\hat{k}(k,h),h\right)}{\Lambda_1\left(\hat{k}(k,h),h\right)} - \int_{\hat{k}(k,h)}^{k^*(h)} \frac{\Lambda_{12}(z,h)}{\Lambda_1(z,h)}dz
\end{aligned}$$

where we have used that $\Lambda_1(k^*(h),h) = 1$. The first term is positive and the third term is positive when $\Lambda_{12}(k,h) < 0$ for all (k,h). The sign of the second term depends on the sign of $\Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) + \Lambda_{12}\left(\hat{k}(k,h),h\right)$. When the sign of this term is negative, $W_3(k,k,h) > 0$. To show that it is negative, differentiate $\hat{k}(k,h)\,\Lambda_1\left(\hat{k}(k,h),h\right) = \Lambda\left(\hat{k}(k,h),h\right) - \Lambda(k,h)$ with respect to h. This delivers

$$\begin{aligned}
\begin{pmatrix} \hat{k}_2(k,h)\,\Lambda_1\left(\hat{k}(k,h),h\right) + \hat{k}(k,h)\,\Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) \\ +\hat{k}(k,h)\,\Lambda_{12}\left(\hat{k}(k,h),h\right) \end{pmatrix} &= \begin{pmatrix} \Lambda_2\left(\hat{k}(k,h),h\right) \\ +\Lambda_1\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) - \Lambda_2(k,h) \end{pmatrix} \Leftrightarrow \\
\hat{k}(k,h)\left[\Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) + \Lambda_{12}\left(\hat{k}(k,h),h\right)\right] &= \Lambda_2\left(\hat{k}(k,h),h\right) - \Lambda_2(k,h) \Leftrightarrow \\
\Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) + \Lambda_{12}\left(\hat{k}(k,h),h\right) &= \frac{\int_k^{\hat{k}(k,h)} \Lambda_{12}(z,h)\,dz}{\hat{k}(k,h)}
\end{aligned}$$

where the r.h.s. is negative when $\Lambda_{12}(k,h) < 0$ for all (k,h).

Proof of lemma 5.

Evaluating equation (15) in $k_1 = k^*(h)$, we obtain

$$\begin{aligned} \Lambda(k_2,h) &= w(k^*(h),k_2,h) + \beta(1-\delta)\Lambda(k_2,h) + \beta\delta W(\underline{k},\underline{k},h) \Leftrightarrow \\ w(k^*(h),k_2,h) &= [1-\beta(1-\delta)]\Lambda(k_2,h) - \beta\delta W(\underline{k},\underline{k},h) \end{aligned}$$

Next, integrate from $\underline{k}$ to $k^*(h)$ over $G^2(k \mid h)$

$$\begin{aligned} \int_{\underline{k}}^{k^*(h)} w(k^*(h),k_2,h)\,dG^2(k_2 \mid h) &= [1-\beta(1-\delta)]\int_{\underline{k}}^{k^*(h)} \Lambda(k_2,h)\,dG^2(k_2 \mid h) - \beta\delta W(\underline{k},\underline{k},h) \Leftrightarrow \\ \int_{\underline{k}}^{k^*(h)} w(k^*(h),k_2,h)\,dG^2(k_2 \mid h) &= (1-\beta)W(\underline{k},\underline{k},h) \end{aligned}$$

Derivation of equations (16) and (17)

Setting the inflow into unemployment equal to the flow out of unemployment delivers

$$\begin{aligned} (1-u)\delta G^1\left(\hat{k}(\underline{k},h) \mid \underline{k},h\right) &= u(h)\left[1-G^1\left(\hat{k}(\underline{k},h) \mid \underline{k},h\right)\right] \Leftrightarrow \\ \delta G^1\left(\hat{k}(\underline{k},h) \mid \underline{k},h\right) &= u(h)\left[1-G^1\left(\hat{k}(\underline{k},h) \mid \underline{k},h\right) + \delta G^1\left(\hat{k}(\underline{k},h) \mid \underline{k},h\right)\right] \Leftrightarrow \\ u(h) &= \frac{\delta G^1\left(\hat{k}(\underline{k},h) \mid \underline{k},h\right)}{1-(1-\delta)G^1\left(\hat{k}(\underline{k},h) \mid \underline{k},h\right)} \end{aligned}$$

Setting inflow into the mass $N(k \mid h)$ equal to outflow from the same mass, we have that

$$\begin{aligned} &u(h)\left(G^1(k,h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) \\ &+(1-u(h))(1-N(k \mid h))\delta\left(G^1(k \mid h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) = (1-u(h))N(k \mid h)\delta\left[1-\left(G^1(k \mid h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right)\right] \\ &\qquad + (1-\delta)(1-u(h))\textstyle\int_{\hat{k}(\underline{k},h)}^{k} n(z \mid h)\left[1-G^1\left(\max\left(\hat{k}(z,h),k\right) \mid h\right)\right]dz \Leftrightarrow \\ &\tfrac{u(h)}{1-u(h)}\left(G^1(k,h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) \\ &+(1-N(k \mid h))\delta\left(G^1(k \mid h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) = N(k \mid h)\delta\left[1-\left(G^1(k \mid h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right)\right] \\ &\qquad + (1-\delta)\textstyle\int_{\hat{k}(\underline{k},h)}^{k} n(z \mid h)\left[1-G^1\left(\max\left(\hat{k}(z,h),k\right) \mid h\right)\right]dz \Leftrightarrow \\ &\tfrac{u(h)}{1-u(h)}\left(G^1(k,h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) \\ &+\delta\left(G^1(k \mid h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) \\ &-N(k \mid h)\delta\left(G^1(k \mid h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) = N(k \mid h)\delta\left[1-\left(G^1(k \mid h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right)\right] \\ &\qquad + (1-\delta)\textstyle\int_{\hat{k}(\underline{k},h)}^{k} n(z \mid h)\left[1-G^1\left(\max\left(\hat{k}(z,h),k\right) \mid h\right)\right]dz \Leftrightarrow \\ &\left[\delta + \tfrac{u(h)}{1-u(h)}\right]\left(G^1(k,h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) = N(k \mid h)\delta + (1-\delta)\textstyle\int_{\hat{k}(\underline{k},h)}^{k} n(z \mid h)\left[1-G^1\left(\max\left(\hat{k}(z,h),k\right) \mid h\right)\right]dz \\ &N(k \mid h) = \left[\delta + \tfrac{u(h)}{1-u(h)}\right]\left(G^1(k,h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) \\ &\qquad + (1-\delta)\textstyle\int_{\hat{k}(\underline{k},h)}^{k} n(z \mid h)G^1\left(\max\left(\hat{k}(z,h),k\right) \mid h\right)dz \end{aligned}$$

Proof of proposition 3

Differentiate equation (17) with respect to h

$$
\begin{aligned}
N_2(k \mid h) \quad = \quad & \frac{\partial\left(\frac{u(h)}{1-u(h)}\right)}{\partial h}\left(G^1(k,h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right) + \left[\delta + \frac{u(h)}{1-u(h)}\right]\left(\frac{\partial\left(G^1(k,h) - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right)}{\partial h}\right) \\
& - (1-\delta)\,\hat{k}_2(\underline{k},h)\, n\left(\hat{k}(\underline{k},h) \mid h\right) g^1\left(\max\left(\hat{k}\left(\hat{k}(\underline{k},h), h\right), k\right) \mid h\right) \\
& + (1-\delta)\int_{\hat{k}(\underline{k},h)}^{k} n_2(z \mid h)\, G^1\left(\max\left(\hat{k}(z,h), k\right) \mid h\right) dz \\
& + (1-\delta)\int_{\hat{k}(\underline{k},h)}^{k} n(z \mid h)\, \frac{\partial G^1\left(\max\left(\hat{k}(z,h),k\right) \mid h\right)}{\partial h} dz
\end{aligned} \tag{28}
$$

When $\Lambda_{12}(k,h) > 0$, $\Lambda_1(k,h) > 0$, $\Lambda_{11}(k,h) < 0$ and $\Lambda_{112}(k,h) \geq 0$, then the first three terms and the fifth term of equation (28) are negative. When $\Lambda_{12}(k,h) < 0$, $\Lambda_1(k,h) > 0$, $\Lambda_{11}(k,h) < 0$ and $\Lambda_{112}(k,h) \leq 0$, then the first three terms and the fifth term are positive. We will consider each of the terms in turn.

Consider the first term of equation (28). Since $k \geq \hat{k}(\underline{k},h)$ the first term has the same sign as $\frac{\partial\left(\frac{u(h)}{1-u(h)}\right)}{\partial h}$. We can write $\frac{u(h)}{1-u(h)} = \frac{\delta G^1\left(\hat{k}(\underline{k},h)|h\right)}{1-G^1\left(\hat{k}(\underline{k},h)|h\right)}$. Differentiating gives us

$$
\begin{aligned}
\frac{\partial\left(\frac{u(h)}{1-u(h)}\right)}{\partial h} &= \frac{\delta \frac{\partial G^1\left(\hat{k}(\underline{k},h)|h\right)}{\partial h}\left[1 - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right] + \frac{\partial G^1\left(\hat{k}(\underline{k},h)|h\right)}{\partial h}\delta G^1\left(\hat{k}(\underline{k},h) \mid h\right)}{\left[1 - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right]^2} \\
&= \frac{\delta \frac{\partial G^1\left(\hat{k}(\underline{k},h)|h\right)}{\partial h}}{\left[1 - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right]^2} \\
&= \frac{-\delta \frac{\Lambda_{11}\left(\hat{k}(\underline{k},h),h\right)\hat{k}_2(\underline{k},h) + \Lambda_{12}\left(\hat{k}(\underline{k},h),h\right)}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}}{\left[1 - G^1\left(\hat{k}(\underline{k},h) \mid h\right)\right]^2}
\end{aligned}
$$

To sign $\Lambda_{11}\left(\hat{k}(\underline{k},h), h\right)\hat{k}_2(\underline{k},h) + \Lambda_{12}\left(\hat{k}(\underline{k},h), h\right)$, we differentiate $\hat{k}(\underline{k},h)$ with respect to h

$$
\begin{aligned}
& \hat{k}_2(k,h)\,\Lambda_1\left(\hat{k}(k,h),h\right) + \hat{k}(k,h)\,\Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) + \hat{k}(k,h)\,\Lambda_{12}\left(\hat{k}(k,h),h\right) \\
= \quad & \Lambda_2\left(\hat{k}(k,h),h\right) + \Lambda_1\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) - \Lambda_2(k,h) \Leftrightarrow \\
& \hat{k}_2(k,h)\,\Lambda_1\left(\hat{k}(k,h),h\right) + \hat{k}(k,h)\left[\Lambda_{12}\left(\hat{k}(k,h),h\right) + \Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h)\right] \\
= \quad & \Lambda_2\left(\hat{k}(k,h),h\right) - \Lambda_2(k,h) + \Lambda_1\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) \Leftrightarrow
\end{aligned}
$$

$$\begin{aligned}
\hat{k}(k,h)\left[\Lambda_{12}\left(\hat{k}(k,h),h\right)+\Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h)\right] &= \Lambda_2\left(\hat{k}(k,h),h\right)-\Lambda_2(k,h)\Leftrightarrow \\
\Lambda_{12}\left(\hat{k}(k,h),h\right)+\Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) &= \frac{\Lambda_2\left(\hat{k}(k,h),h\right)-\Lambda_2(k,h)}{\hat{k}(k,h)}\Leftrightarrow \\
\Lambda_{12}\left(\hat{k}(k,h),h\right)+\Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h) &= \frac{\int_{z=k}^{\hat{k}}\Lambda_{12}(z,h)\,dz}{\Lambda\left(\hat{k}(k,h),h\right)-\Lambda(k,h)}\Lambda_1\left(\hat{k}(k,h),h\right)
\end{aligned} \tag{29}$$

where the right hand side is positive when $\Lambda_{12}(k,h)>0$ and negative when $\Lambda_{12}(k,h)<0$. Therefore, the first term of the r.h.s. of equation (28) is negative when $\Lambda_{12}(k,h)>0$ and positive when $\Lambda_{12}(k,h)<0$.

To sign the second term of equation (28), we need to sign

$$\begin{aligned}
\frac{\partial\left(G^1(k,h)-G^1\left(\hat{k}(\underline{k},h)\mid h\right)\right)}{\partial h} &= \frac{-\Lambda_{12}(k,h)}{\Lambda_1(k,h)}-\frac{-\left(\Lambda_{11}\left(\hat{k}(\underline{k},h),h\right)\hat{k}_2(\underline{k},h)+\Lambda_{12}\left(\hat{k}(\underline{k},h),h\right)\right)}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)} \\
&= -\left(\frac{\Lambda_{12}(k,h)}{\Lambda_1(k,h)}-\frac{\Lambda_{12}\left(\hat{k}(\underline{k},h),h\right)}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}\right)+\frac{\Lambda_{11}\left(\hat{k}(\underline{k},h),h\right)\hat{k}_2(\underline{k},h)}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}
\end{aligned}$$

Under the assumptions that $\Lambda_{12}(k,h)>0$, $\Lambda_1(k,h)>0$, $\Lambda_{11}(k,h)<0$ and $\Lambda_{112}(k,h)\geq 0$, $\frac{\Lambda_{12}(k,h)}{\Lambda_1(k,h)}\geq\frac{\Lambda_{12}\left(\hat{k}(\underline{k},h),h\right)}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}$ since $k\geq\hat{k}(\underline{k},h)$ and hence the first term is negative. The second term is also negative since $\Lambda_{11}\left(\hat{k}(\underline{k},h),h\right)<0$ and $\hat{k}_2(\underline{k},h)>0$ under the maintained assumptions (see proposition 2). Under the assumptions that $\Lambda_{12}(k,h)<0$, $\Lambda_1(k,h)>0$, $\Lambda_{11}(k,h)<0$ and $\Lambda_{112}(k,h)\leq 0$, $\frac{\Lambda_{12}(k,h)}{\Lambda_1(k,h)}\leq\frac{\Lambda_{12}\left(\hat{k}(\underline{k},h),h\right)}{\Lambda_1\left(\hat{k}(\underline{k},h),h\right)}$ since $k\geq\hat{k}(\underline{k},h)$ and hence the first term is positive. The second term is also positive since $\Lambda_{11}\left(\hat{k}(\underline{k},h),h\right)<0$ and $\hat{k}_2(\underline{k},h)<0$.

The third term of equation (28) is negative when $\Lambda_{12}(k,h)>0$, $\Lambda_1(k,h)>0$, $\Lambda_{11}(k,h)<0$ and $\Lambda_{112}(k,h)\geq 0$ due to proposition 2. Furthermore, the third term of equation (28) is positive when $\Lambda_{12}(k,h)<0$, $\Lambda_1(k,h)>0$, $\Lambda_{11}(k,h)<0$ and $\Lambda_{112}(k,h)\leq 0$.

The fifth term of equation (28) is negative when $\Lambda_{12}(k,h)>0$ following the argument for the first term of equation (28). The fifth term is positive when $\Lambda_{12}(k,h)<0$.

Finally, we turn to the fourth term of equation (28). Suppose that $\Lambda_{12}(k,h)>0$, $\Lambda_1(k,h)>0$, $\Lambda_{11}(k,h)<0$ and $\Lambda_{112}(k,h)\geq 0$ such that the first three terms and the fifth term on the r.h.s. of equation (28) are negative. The remaining question is whether the fourth term could be positive and imply that the right hand side becomes positive as well. Therefore, suppose that $(1-\delta)\int_{\hat{k}(\underline{k},h)}^{k}n_2(z\mid h)\,G^1\left(\max\left(\hat{k}(z,h),k\right)\mid h\right)dz>0$, then we must have that $N_2(k\mid h)>(1-\delta)\int_{\hat{k}(\underline{k},h)}^{k}n_2(z\mid h)\,G^1\left(\max\left(\hat{k}(z,h),k\right)\mid h\right)dz$, which is not possible as the remaining terms on the right hand side of the equation are negative, i.e. a contradiction. A similar argument can be made for the case where $\Lambda_{12}(k,h) < 0$, $\Lambda_1(k,h) > 0$, $\Lambda_{11}(k,h) < 0$ and $\Lambda_{112}(k,h) \leq 0$.

Derivation of equation (18)

The values of y_1/Λ_1, $\Lambda_1/\hat{\Lambda}_1$ and $\left[\hat{\Lambda} - \Lambda\right]$ are all positive. The terms Λ_{12} and $\hat{\Lambda}_2 - \Lambda_2 = \int_{z=k}^{\hat{k}} \Lambda_{12}(z,h)\, dz$ are positive if and only if $\Lambda(k,h)$ is super-modular. The result is straightforward to derive for $k \geq p^*$, so here we only consider the case where $k < p^*$. Differentiate equation (25) with respect to h to obtain

$$\begin{aligned} y_{12}(k,h) &= \left(1 - \beta(1-\delta)\frac{1}{\Lambda_1\left(\hat{k}(k,h),h\right)}\right)\Lambda_{12}(k,h) \\ &\quad + \beta(1-\delta)\frac{\Lambda_1(k,h)\left[\Lambda_{12}\left(\hat{k}(k,h),h\right) + \Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h)\right]}{\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)^2} \\ &= \left(1 - \beta(1-\delta)\frac{1}{\Lambda_1\left(\hat{k}(k,h),h\right)}\right)\Lambda_{12}(k,h) \\ &\quad + \beta(1-\delta)\frac{\Lambda_1(k,h)\left[\Lambda_{12}\left(\hat{k}(k,h),h\right) + \Lambda_{11}\left(\hat{k}(k,h),h\right)\hat{k}_2(k,h)\right]}{\left(\Lambda_1\left(\hat{k}(k,h),h\right)\right)^2} \end{aligned}$$

Using equations (25) and (29), we can write

$$\begin{aligned} y_{12}(k,h) = \left[\tfrac{y_1(k,h)}{\Lambda_1(k,h)}\right]\Lambda_{12}(k,h) + \beta(1-\delta)\tfrac{\Lambda_1(k,h)}{\Lambda_1\left(\hat{k}(k,h),h\right)}\tfrac{\Lambda_2\left(\hat{k}(k,h),h\right) - \Lambda_2(k,h)}{\Lambda\left(\hat{k}(k,h),h\right) - \Lambda(k,h)} \Leftrightarrow \\ \tfrac{y_{12}(k,h)}{y_1(k,h)} = \tfrac{\Lambda_{12}(k,h)}{\Lambda_1(k,h)} + \tfrac{\beta(1-\delta)\Lambda_1(k,h)}{y_1(k,h)\Lambda_1\left(\hat{k}(k,h),h\right)}\tfrac{\int_k^{\hat{k}(k,h)}\Lambda_{12}(z,h)dz}{\int_k^{\hat{k}(k,h)}\Lambda_1(z,h)dz} \Leftrightarrow \\ \tfrac{y_{12}(k,h)}{y_1(k,h)} = \tfrac{\Lambda_{12}(k,h)}{\Lambda_1(k,h)} + \tfrac{\beta(1-\delta)}{\Lambda_1\left(\hat{k}(k,h),h\right) - \beta(1-\delta)}\tfrac{\int_k^{\hat{k}(k,h)}\Lambda_{12}(z,h)dz}{\int_k^{\hat{k}(k,h)}\Lambda_1(z,h)dz} \end{aligned}$$

Signing $w_3(k_1, k_2, h)$.

Differentiate equation (15) respect to h

$$\begin{aligned} w_3(k_1,k_2,h) &= y_2(k_1,h) + \beta(1-\delta)\frac{\partial G^1\left(\hat{k}(k,h) \mid h\right)}{\partial h}\left[\Lambda(k_1,h) - \Lambda(k_2,h)\right] \\ &\quad + \left[1 - \beta(1-\delta)G^1\left(\hat{k}(k,h) \mid h\right)\right]\left[\Lambda_2(k_1,h) - \Lambda_2(k_2,h)\right] \\ &= y_2(k_1,h) + \beta(1-\delta)\frac{\partial G^1\left(\hat{k}(k,h) \mid h\right)}{\partial h}\left[\Lambda(k_1,h) - \Lambda(k_2,h)\right] \\ &\quad - \left[1 - \beta(1-\delta)G^1\left(\hat{k}(k,h) \mid h\right)\right]\int_{k_2}^{k_1}\Lambda_{12}(z,h)\, dz \end{aligned} \tag{30}$$

The first term of equation (30) is always positive, whereas we know from the proof of proposition 2 that $\frac{\partial G^1\left(\hat{k}(k,h)|h\right)}{\partial h} < 0$ if $\Lambda_{12}\left(\hat{k}(k,h),h\right) > 0$ and $\frac{\partial G^1\left(\hat{k}(k,h)|h\right)}{\partial h} > 0$ if $\Lambda_{12}\left(\hat{k}(k,h),h\right) < 0$. The implication is that the second and third terms are negative/ positive if PAM/ NAM. The right hand side is positive whenever the job value function is super-modular and negative when the job value function is positive. This implies that $\frac{\partial G^1\left(\hat{k}(k,h)|h\right)}{\partial h} < 0$ when the job value function is super-modular whereas $\partial G^1\left(\hat{k}(k,h) \mid h\right) > 0$ when the job value function is sub-modular Therefore, super-modularity of the job value function implies that the second and third terms of equation (30) are negative. This is reversed if the job value function is sub-modular.

B Additional Results of Global Absolute Advantage Test

We conduct the following robustness checks for the global absolute advantage test:

- In Figure A1 we use poaching index to rank firms and workers;
- In Figure A2 we conduct the global absolute advantage test for each of the five geographical regions in Denmark: Hovedstaden, Midtjylland, Nordjylland, Sjælland and Syddanmark;
- In Figure A3 we conduct the global absolute advantage test for each four-year time period from 1995 to 2010;
- In Figure A4 we conduct the global absolute advantage test for male and female workers separately;
- In Figure A5 we conduct the global absolute advantage test for different age groups.

In all of these tests, we find that the lowest firm types have negative coefficients, suggesting that global absolute advantage fail to hold within regions, time periods, and gender and age groups.

In Figure A6 we test global absolute advantage for each industry sector. Most sectors do not have global absolute advantage, but there are two exceptions: finance and knowledge services. These are high-skilled sectors and the results suggest that for high-skilled jobs we do have global absolute advantage, that is, workers who are employed at better firms are also more productive in high-skilled jobs.

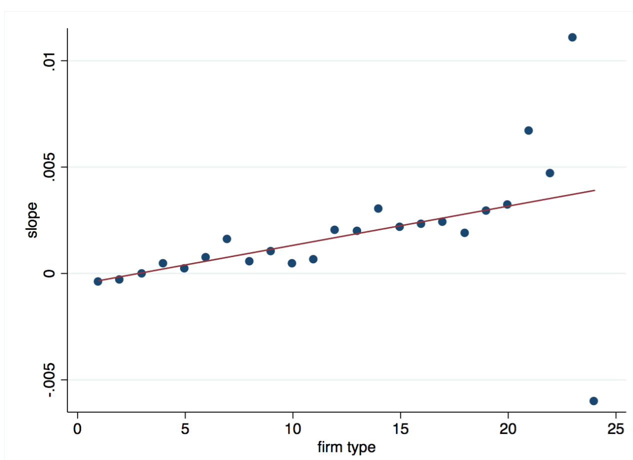
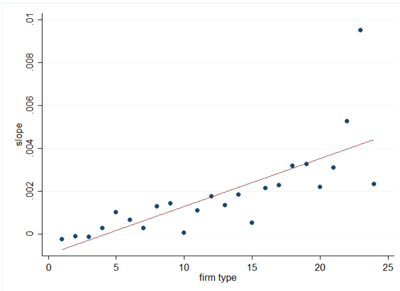
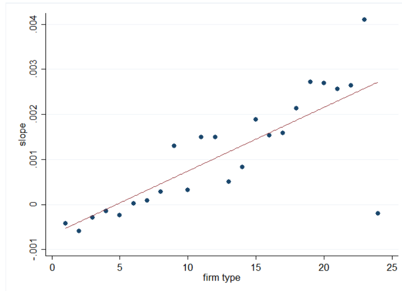


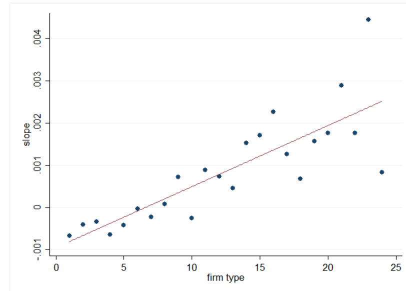
Figure A1: Global absolute advantage test: rank firms using poaching index



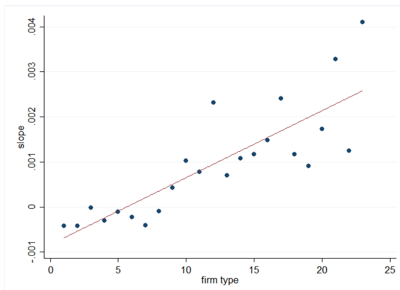
(a) Region 1



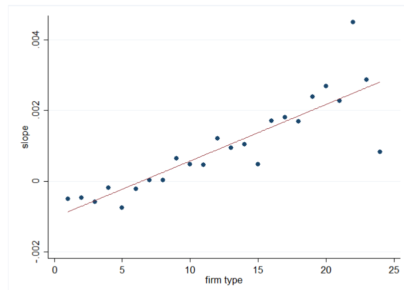
(b) Region 2



(c) Region 3

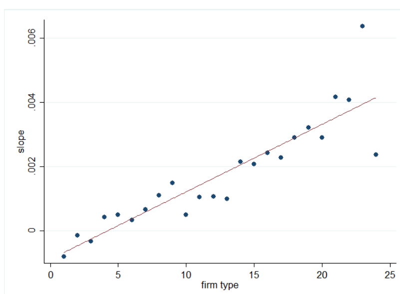


(d) Region 4

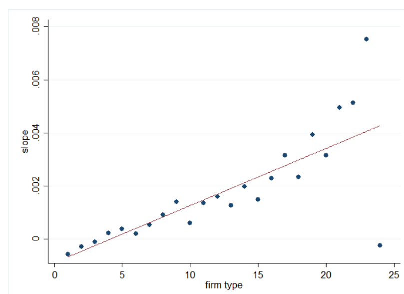


(e) Region 5

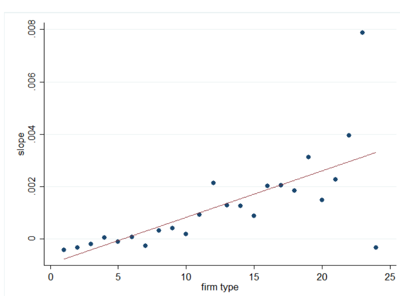
Figure A2: Testing Global Absolute Advantage by Regions



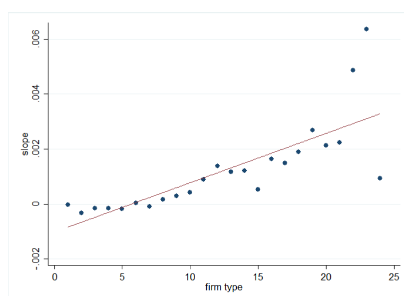
(a) 1995-1998



(b) 1999-2002

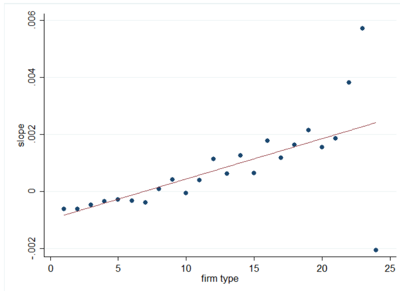


(c) 2003-2006

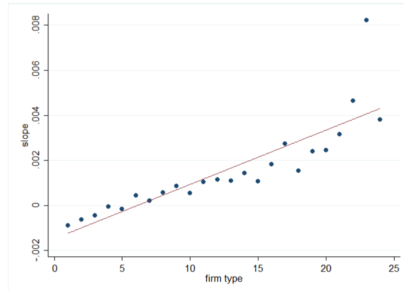


(d) 2007-2010

Figure A3: Testing Global Absolute Advantage by Year

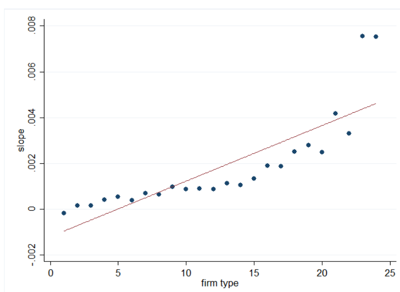


(a) Male

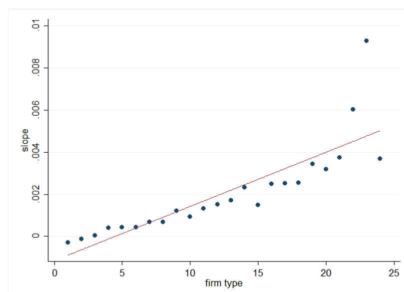


(b) Female

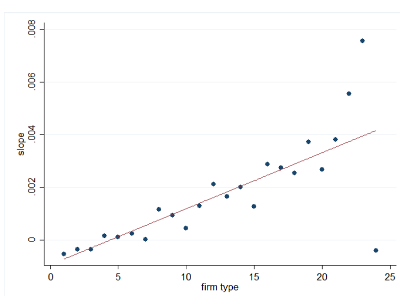
Figure A4: Testing Global Absolute Advantage by Gender



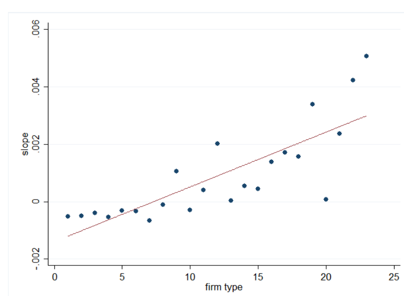
(a) Age 25-29



(b) Age 30-39

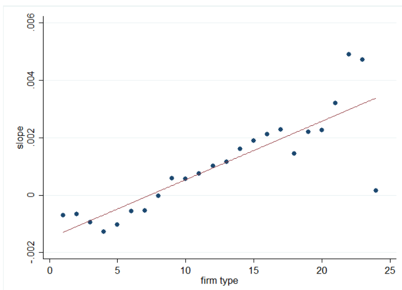


(c) Age 40-49

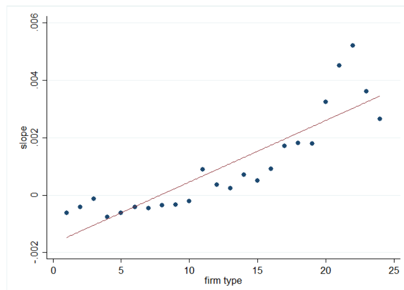


(d) Age 50-59

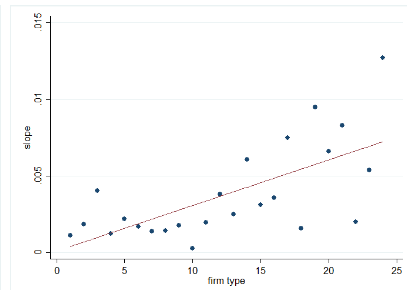
Figure A5: Testing Global Absolute Advantage by Age



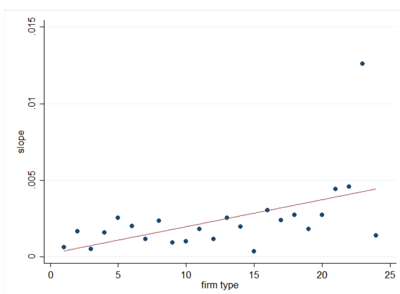
(a) Manufacturing



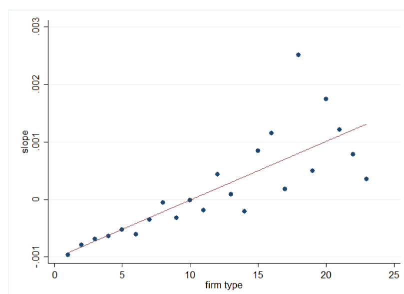
(b) Retail and Wholesale



(c) Finance



(d) Knowledge Services



(e) Construction

Figure A6: Testing Global Absolute Advantage by Industries

Economics Working Papers

2017-11: Erik Strøjer Madsen: Branding and Performance in the Global Beer Market

2017-12: Yao Amber Li, Valerie Smeets and Frederic Warzynski: Processing Trade, Productivity and Prices: Evidence from a Chinese Production Survey

2017-13: Jesper Bagger, Espen R. Moen and Rune M. Vejlin: Optimal Taxation with On-the-Job Search

2018-01: Eva Rye Johansen, Helena Skyt Nielsen and Mette Verner: Long-term Consequences of Early Parenthood

2018-02: Ritwik Banerjee, Nabanita Datta Gupta and Marie Claire Villeval: Self Confidence Spillovers and Motivated Beliefs

2018-03: Emmanuele Bobbio and Henning Bunzel: The Danish Matched Employer-Employee Data

2018-04: Martin Paldam: The strategies of economic research - An empirical study

2018-05: Ingo Geishecker, Philipp J.H. Schröder, and Allan Sørensen: One-off Export Events

2018-06: Jesper Bagger, Mads Hejlesen, Kazuhiko Sumiya and Rune Vejlin: Income Taxation and the Equilibrium Allocation of Labor

2018-07: Tom Engsted: Frekvensbaserede versus bayesianske metoder i empirisk økonomi

2018-08: John Kennes, Daniel le Maire and Sebastian Roelsgaard: Equivalence of Canonical Matching Models

2018-09: Rune V. Lesner, Anna Piil Damm, Preben Bertelsen and Mads Uffe Pedersen: Life Skills Development of Teenagers through Spare-Time Jobs

2018-10: Alex Xi He, John Kennes and Daniel le Maire: Complementarity and Advantage in the Competing Auctions of Skills