Answers to Referee 1

1 Analysis

1. There is a missing λ in eq. (12), I agree with the referee. However I remark that the matching process is solved given λ which is treated as a constant. The correction does not affect the results.

2. The expected value of $a(\lambda, \theta)$ should be evaluated as $\int_{\theta^*}^{\bar{\theta}} a(\lambda, \theta)\gamma(\theta)d\theta$, as suggested by the referee. This correction should be applied also to the expected value of $\alpha(\lambda, \theta)$. I remark that these corrections do not affect the main results of the paper since when going from eq. 30 to eq. 31 (equation 31 contains the crucial results of the paper) only the last term of the second and of the third line in eq. 31 changes. The main results are confirmed after applying the correction.

3. and 4. In a dynamic setting, the urn-ball process applies as far as one condition holds (Blanchard and Diamond, 1994; ReStud): vacancies stay opened for a discrete period of time (see eqs. (1) and (2) in their paper and the discussion of these authors presented on p. 420). This is almost the same trick used by Moen (1999) in order to have a continuous time matching function identical to that of the instantaneous process. I explicitly refer to Blanchard and Diamond when going from the static form to the continuous time form of the urn-ball process. Probably, I should be more precise on this, linking this result to my framework. However, this result has already been set out in the literature.

Concerning the ability distribution among unemployed workers it is true that high-ability types have higher probability of being employed. Notwithstanding, the job destruction rate b affects all job positions homogeneously. This implies that under some conditions, in the steady-state (where inflows and outflows are the same) unemployed have the same ability distribution of the population $\Gamma(.)$ since the flow of workers going into unemployment is given, on average, by those workers who have been "preferred" when hiring took place. I agree this point should be absolutely better discussed in the paper and formalized.

5. The Bellman equations I reported are correct. Individual ability θ is observed when a match is realized. This implies that, if I define V_g^F as the value of a vacancy filled with a $\theta - type$ worker, I can then define the value of an unfilled vacancy by considering the probability that the vacant position has a match with a $\theta - type$ worker ($\alpha(\lambda, \theta)$). Of course, I can also define the expected value of an unfilled vacancy as indicated by the referee. However, in this case I would obtain an expression only for the expected wage. Since individuals take their decisions by looking at the wage they will gain in that specific sector (ability is observed when a match is realized) I need to define the wage conditional to θ hence I need to work with the value functions I have defined.

The expected value of an unfilled vacancy should be evaluated when the entry-decision is taken. In this case, the Bellman equation indicated by the referee is considered. This is exactly the meaning of the expression $E[V_g^V|\theta]$. This point needs to be clarified following the referee's suggestion.

Concerning the proof of proposition 1, in this proof I have clearly remarked that γ is the probability that an individual chooses the g sector and *ex-post* it is a density. Then, I have also proved that $\gamma = 1 - \Gamma(\theta^*)$, i.e., it is a density. This is the standard way in which BNE is discussed (see Fudenberg and Tirole Game Theory textbook pp. 211-212.). However, it is probably better to use a different symbol instead of γ .

2 Potential

The referee questions the main intuition of the paper, i.e., the existence of equilibria where, conditional on the selectivity of the higher education sector, firms find optimal not to add additional screening and choose randomly amongst the applications they received. The referee points are stated as follows:

a) Firms do not collude. Then, if firms do collude, they would do this not only in terms of ranking but they can collude in terms of number of vacancies.

Firstly, I need to remark that the interpretation of *subgame perfect (bayesian) nash equilibrium* solution concept as tacit collusion is only *one* amongst all the possible interpretations. Indeed, tacit collusion is what we observe *ex post* when we think of an interaction process where agents do not play the strategy maximizing their one-shot game payoff. Indeed, before saying that firms' behavior is the result of collusion, we should interpret the resulting equilibrium as a stable intersection between firms' best response functions in an infinitely repeated interaction process. As far as the agents care about future payoffs, we cannot exclude these equilibria. This being said, the question is how do agents arrive to these equilibria. This answer is not an easy one. There are tons of literature about agents' beliefs, the formation of common beliefs (evolution, learning, etc.) and the reasonability of common beliefs. However, as far as we accept the SPBNE as a solution concept to evaluate economic phenomena, we cannot exclude that these equilibria are actually achieved. These equilibria represent intersections of best responses and, consequently, perfectly rational behavior.

Secondly, I need to reply on why collusion is not modeled in terms of number of vacancies to be posted. It is true that the number of vacancies that are posted in the market generates externalities for other agents hence there is room for collusion in this respect too. However, at the outset, Hosios (1990) proved this is not necessarily the case, i.e., it is possible to have matching frictions in the presence of social efficiency so that collusion in terms of vacancies does not make sense. In addition, since vacancies can be created at no costs (free-entry) the value of a vacancy should be zero in the steady state. Hence, collusion among existing firms cannot arise whenever they keep the value of an unfilled vacancy greater than zero. Put differently, collusion among firms in terms of vacancies cannot arise if we assume free entry condition. Instead, the idea of the paper is perfectly consistent with the free entry condition. The intuition is the following. Assume that ranking applies. In this case, in the steady state we would have that the value of an unfilled vacancy is zero. Now consider the possibility that ranking is not optimal. If firms switch from ranking to no-ranking, they would attract more workers. The presence of more workers induces a rise of the value of a filled vacancy since having more workers implies lower wages (nash bargaining is assumed). Furthermore, the value of unfilled vacancies increases too (more chances of filling a vacancy) then more firms enter the market until the new steady state is achieved. The free entry condition holds in this case too. Whatever is the adopted ranking behavior, this is compatible with a value of unfilled vacancy equal to zero in the steady state, hence, with the free entry condition.

b) Why firms do not pay a bonus or higher wages to attract more workers instead of using ranking?

Answer: the paper assumes that ranking-decision is not costly, while wages are (of course) costly, then ranking is much a better choice to raise profit. Using wage in order to raise the share of graduates would not be a good idea.

I remark that ranking activities are not costly in my paper because I do not want to rely on an *ad hoc* assumption to exclude ranking. Because

ranking is not costly there are no death-weight losses related to the ranking activities in the ug sector.

c) No sense of cooperation in matching models.

I have already explained in answer to point a) the interpretation of SPBNE. Here, I remark that it is well possible to be in the presence of firms that are heterogeneous in terms of their exogenous characteristics (heterogeneity is at the root of matching models) that find optimal to apply the same ranking process. As regards the free entry condition, it holds since I remarked that the model follows the standard lines set out by Diamond-Mortensen-Pissarides. However I should be more explicit on this point.

3 Other Comments

A simultaneous move game of incomplete information with a finite number of heterogeneous agents, a set of types for each player, probability distribution over types, a payoff function for each player and a set of actions (pure strategy) for each player is defined as a Bayesian game (Fudenberg and Tirole, pp. 213-205).

It is true that Charlot and Decreuse (2005) use a perfectly segmented twosector model. However they do not use an urn-ball matching process. Instead Moen (1999) and Gravel (1999) do it. It is not true that Gravel (1999) does not use high-skill and low-skill sectors. Interestingly, referee 3 suggested the use of "skilled and unskilled" rather than "graduate and undergraduate" in order to "fitting in better with the literature on skilled/unskilled labor (and with Gavrel's paper)".

Blanchard and Diamond address ranking according to unemployment duration because their aim was to link directly the theory to the data concerning unemployment duration. However the reason why firms rank applicants according to unemployment duration is because they expect individuals with long unemployment histories to be less productive due to human capital deterioration. This implies ranking according to productivity. Quoting Blanchard and Diamond (1994) "An alternative assumption [......] would be that there is an arbitrarily small deterioration of skills with unemployment duration, so that, while workers are all acceptable, the firm marginally prefers those who have been unemployed the least time" (p. 422).

Notation can be modified, and it is probably better to use x for individual ability (as suggested by referre 3).