Pride and Prejudice on a Centralized Academic Labor Market

Philippe Caillou and Michele Sebag

Abstract The Academic Labor Market in France can be viewed as a constrained Stable Marriage problem, pairing universities and candidates according to their (elitist) preferences. A Multi-Agent based model, calibrated after the empirical evidence, is used to investigate how universities can recruit the best candidates with high confidence. Extensive simulations suggest that universities can be divided in four categories: top and medium universities have no difficulty in attracting the candidates they have selected, contrarily to good and bad universities. In this paper, a learning mechanism is presented: universities are allowed to tune their expectations depending on whether they did succeed to attract candidates in the previous recruitment rounds. The impact of over/under estimations is analyzed with respect to the hiring efficiency and quality.

1 Introduction

National academic labor markets (ALMs) are strongly influenced by the culture and history of the country [1]. The French system examined in this paper reflects an egalitarian tradition; the hiring process globally aggregates the preferences of universities and candidates using a Stable Marriage-like algorithm [2]. Due to administrative constraints (limited size of the short list), this centralized procedure might entail some hiring inefficiencies, where good universities might select top candidates, who will ultimately prefer better universities. Universities might therefore use less elitist and more secure recruitment strategies. The goal of this paper is to examine how strategies based on raising/lowering the university expectations might improve their hiring efficiency. In an earlier study of the French ALM [3], a Multi-Agent (MA) simulation framework has been proposed to assess the hiring rate and the quality of the recruitment process. In this paper, this framework is extended, al-

LRI, Universite Paris Sud, F-91405 Orsay France
caillou;sebag@lri.fr

lowing universities to adjust their selectivity depending on their hiring success in the previous steps. The efficiency of the considered strategies is discussed with respect to the university position (relatively to the set of universities).

The paper is organized as follows. Section 2 briefly reviews and discusses relevant work. The French academic labor market (ALM) is presented in section 3. Section 4 describes the MA-based model. Section 5 discusses the lessons learned from extensive simulations conducted with this model, and the paper concludes with some perspectives for further research.

2 Related work

Centralized (labor) markets are based on the preferences of sellers (here, the candidates) and buyers (the universities). The combinatorial optimization problem of building an optimal pairing, referred to as *Stable Marriage* problem, has been extensively investigated since Gale & Shapley pioneering work [2]. The French ministry actually uses a variant of the Stable Marriage algorithm (akin [4]) to compute an optimal assignment of candidates to universities. Importantly, the procedure is shown to be *truthful*, in the sense that no agent could improve its outcome by lying about its preferences [5]. The optimality and truthfulness properties however only hold in an idealized setting (rational agents, unbounded shortlists).

The French academic labor market has more specifically been studied by [1] in a sociological perspective. This work focuses on organizational, societal and cognitive aspects; the professional efficiency, the university organization and department cohesiveness, and the quality assessment are related to the hiring process. The Local Hiring phenomenon has been investigated in a specific area by [6], empirically measuring how the proximity between the PhD jury of a candidate and the jury of highly competitive national examinations, is correlated with the probability of success of the candidate.

On the computational side, a variety of social and economical problems have been investigated using multi-agent systems (MAs) [7, 8]. MAs have demonstrated their ability to both represent (cognitive) agents and constrained interaction rules, and provide insights into the dynamics of the system. More generally, MAs are increasingly being considered as a flexible and versatile modelling framework, enabling positive and normative investigations of phenomena out of reach of analytical studies, and supported by efficient programming environments (e.g., ModulEco [9] and RePast [10]). In an earlier work [3], a MAS has been proposed to model the French ALM and study the local hiring problem; interestingly, local hiring (see below) was shown to be an efficient hiring strategy (as opposed to, a "bad university habit") in some settings.

3 The French Academic Labor Market

This section describes the French academic hiring process and the available ground truth.

3.1 The hiring process

• The list of all open positions in all universities is published by the State department. Every PhD[1] is free to apply to any such position; the number of applications is not restricted.
• In each university, for each position, a jury is designated, selects candidates and interviews the selected candidates.
• Every selected candidate goes to every interview (except for conflicting schedules or if he has been formerly top-listed in a University he prefers).
• For each position, the jury publishes a shortlist of at most five names, selected among the interviewed candidates.
• Each candidate is informed of the positions he has been shortlisted for, together with his rank; he symmetrically ranks all positions (no length constraint) according to his preferences.
• All university shortlists and candidate ranking lists are sent to the ministry; a stable marriage like algorithm is used to compute the actual matching.

3.2 Empirical Evidence

In 2007, the number of open positions, the number of candidates and the number of applications in every discipline were published by the State department (Table 1). Three main categories were distinguished. In the first category, including Law, Economics & Management (L&M) disciplines, each candidate applies on average to 50% of the opened positions. In the other categories, including Science on the one hand and Literature and Humanities on the other hand (H&S), the number of candidates per position (pressure) is significantly higher and the number of applications per candidate is significantly lesser. Globally, the hiring process is efficient in the sense that the recruitment rate is 98%. The local hiring rate, that is the percentage of universities recruiting candidates who passed their PhD in this same university, is circa 28% (37% in L&M, 24% in Humanities and 28% in Science).

[1] A pre-filter referred to as "qualification" is used to reject PhDs with no teaching experience. This step is left out of the study for the sake of simplicity.

Table 1 The French academic labor market in 2007.

2007	Total	L&M	Hum.	Science
Positions	2110	324	695	1000
Candidates	9318	555	3135	5540
Applications/candidate	8,3	26,8	8,8	6,0
HiringRate	98%	94%	99%	99%
LocalHiring rate	28%	37%	24%	28%
nbSections	57	6	25	23
Avg. nbJobs / sect.	37,0	54,0	27,8	43,5

4 Academic Labor Market Modelling

This section describes the Multi-Agent model proposed for the French academic labor market, together with the assumptions made regarding university and candidate preferences.

4.1 Agent Preferences

It might be safely assumed that universities aim at recruiting the best candidates while candidates aim at being recruited in the best universities. Usually each agent has however different quality criteria; and, would it exist, the “true quality” ordering is unknown.

The proposed modelling will thus proceed in a backward manner, assuming that there exists such a “true ordering”, of which each agent preference ordering is a perturbed variant. Only two types of perturbations are considered at the moment, respectively based on locality preferences and on random noise (see below).

Formally, letting U denote the number of universities, it is assumed with no loss of generality that the set of universities $\{u_1,\ldots,u_U\}$ is ordered according to the “true ordering”. Symmetrically, C denotes the number of candidates and $\{c_1,\ldots c_C\}$ the set of candidates ordered after the “true ordering”.

4.2 Multi-Agent based Model

The MAS involves two types of agents, candidates $\{c_1,\ldots,c_C\}$ and universities $\{u_1,\ldots,u_U\}$, where an agent index stands for its rank after the (unknown) “true ordering”. Furthermore the model is spatialized, that is, to each agent are associated 2D coordinates (in $[0,1]$). The home University of an candidate is the nearest one after the Euclidean distance.

4.2.1 Candidates

Candidate c_i is characterized from five parameters. The first two parameters (in $[0,1]$) govern his preference ordering: i) elitism e_i stands for his bias toward the best universities; ii) locality ℓ_i stands for his bias toward the nearest universities. A random perturbation, modelled as $(1-e_i-\ell_i)V$ with V uniformly drawn in $[0,1]$, accounts for his subjective preferences. Overall, the quality $q(i,t)$ of university u_t for candidate c_i is computed as (the lower the better):

$$q(i,t) = \mathrm{e}_i \times \frac{t}{U} + \ell_i \times d(c_i,u_t) + (1-e_i-\ell_i) \times V$$

Three more parameters are used to model the application strategy of candidate c_i. A risk-propensity parameter r_i determines whether he rather applies to the top-ranked universities (according to the preference ordering $q(i,\cdot)$), or to the universities best matching his own rank. Precisely, the strategic ordering of c_i is defined as (the lower the better):

$$s(i,t) = r_i q(i,t) + (1-r_i)\frac{|i-t|}{C}$$

His application strategy is finally defined from the number N_i of positions he will apply to; c_i applies to the top N_i universities after the ordering $s(i,t)$. Independently, c_i applies to his home University with probability h_i (empirically, candidates always apply to their home university).

4.2.2 Universities

Likewise, university u_t is characterized from four parameters. The first two parameters (in $[0,1]$) govern his preference ordering: i) elitism e_t stands for its bias toward the best candidates; ii) locality ℓ_t stands for its bias toward local candidates. Lastly, a random perturbation modeled as $(1-e_t)V$ with V uniformly drawn in $[0,1]$, accounts for the "subjective" preferences of university u_t. Overall, the quality $r(i,t)$ of candidate c_i for university u_t is:

$$r(i,t) = (\mathrm{e}_t \times \frac{i}{C} + (1-e_t)V)(1-\ell_t.\delta_{i,t})$$

where $\delta_{i,t}$ is 1 iff c_i is local to u_t and 0 otherwise.

University u_t selects the candidates to be interviewed after its risk propensity r_t and a *SelfAssessment* parameter o_t, where o_t is positive (respectively negative) if university u_t tends to consider itself less attractive (respectively more attractive) than it is after the "true" university ordering. More precisely, its strategic ordering is defined as:

$$s'(i,t) = r_t \times r(i,t) + (1-r_t) \times \frac{|i-(t+o_t)|}{C}$$

Two settings, referred to as *NoLearning* and *Learning*, are distinguished in the following. In the *NoLearning* setting, o_t is set to 0. In the *Learning* setting, o_t is adjusted after each “move” (yearly recruitment). In the case the university did not recruit its candidate, it lowers its expectation and o_t is incremented. Otherwise o_t is decremented with probability α, where α is computed after the empirical hiring rate[2]). s

4.2.3 Interaction rules

Every candidate c_i applies for the top N_i positions after ordering $s(i,\cdot)$, where N_i is uniformly selected in $[1, Max.Application]$, and he applies to his home university with probability h_i.

Every university u_t produces a shortlist of 5 names, the top 5 candidates after ordering $s'(\cdot,t)$[3].

Every candidate c_i thereafter ranks the universities having shortlisted him after the $q(i,\cdot)$ ordering. Eventually, the candidates and universities preferences are aggregated by a variant of Stable Marriage algorithm [4], an optimal matching is derived, and the recruitment decisions are made accordingly.

5 Simulation Results

5.1 Methodology and Experimental settings

The main two efficiency indicators of an ALM are the *HiringRate* (fraction of positions fulfilled) and the *LocalHiring* rate (fraction of positions fulfilled by local candidates). We further consider the *FameLoss* of each university, defined as the difference between the rank of the recruited candidate and its own rank (not fulfilled positions are not considered).

The key parameters of the MA-based model (number of positions, of candidates and maximal number of applications per candidate, size of the shortlist) are calibrated after the empirical evidence presented for the H&S disciplines[4] (section 4). The behavioral parameters (elitism, localism, risk-propensity) are set to the so-

---

[2] Formally, α is such that the average hiring rate converges toward the empiric rate ω. At the equilibrium, the expected increase equals the expected decrease:

$$1-\omega = \omega \times \alpha \Rightarrow \alpha = \frac{1-\omega}{\omega}$$

[3] For the sake of simplicity, the impact of the live interviews is not accounted for in the model.

[4] The L&M setting corresponds to a saturated market, where almost 50% of the candidates apply to every job; in this situation a high locality bias is needed to enforce a reasonable hiring rate, as shown in [3].

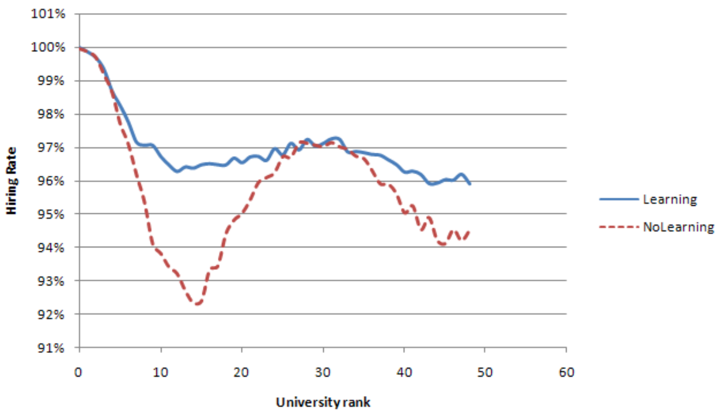

Fig. 1 *HiringRate* versus University Rank in the *NoLearning* and *Learning* cases

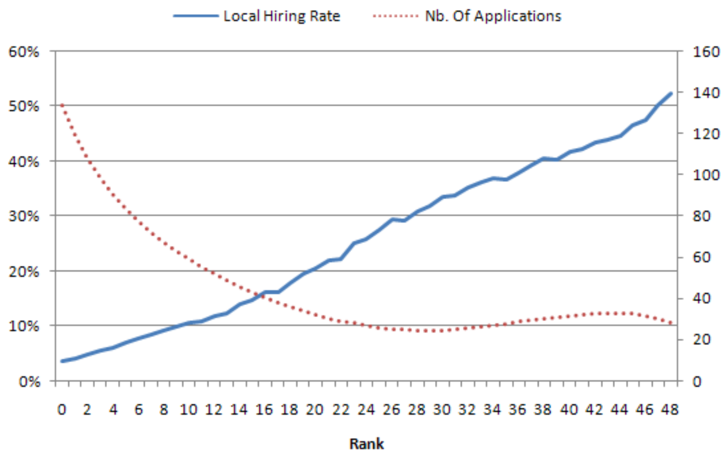

Fig. 2 Local Hiring Rate and Nb. of applications vs. university rank in the No-Learning case

called elitist settings (Table 2), studied and validated in [3]. The simulations were performed using the RePast Framework [10]. All reported results are averaged over 1,000 independent simulations with same parameter setting.

Table 2 Parameter Values in the Simulations

General				Candidates				Universities			
Positions	Candidates	*Max.Application*	ω	e_i	ℓ_i	r_i	h_i	e_t	ℓ_t	r_t	o_t
50	200	20	.97	.7	.1	U[.1; 1]	1	.7	U[0, .2]	U[.1;1]	0

5.2 No learning setting

In the *NoLearning* case ($o_t = 0$), the results closely match the available ground truth: *HiringRate* rate is 97% (vs 94% in empirical data) and Local Hiring rate is 28% (vs 28%). Interestingly, the *HiringRate* and the average number of received applications do not vary linearly w.r.t. the university rank. More precisely, four categories of universities can be distinguished (Fig. 1):

- The *Best* (Top 8 universities) have a high *HiringRate* and receive many applications. The *Best* universities choose the best candidates, who come. The *LocalHiring* rate is the lowest one (Fig. 2).
- The *Good* (rank between 9 and 21) have a low *HiringRate* although they receive a high number of applications. These universities, in the shadow of the *Best* ones, particularly suffer from the limited size of the short list to select the best candidates. While many good candidates apply to the *Good* universities, if they are selected they seldom come; they go to the *Best* universities.
- The *Medium* (between 22 and 39) have a high *HiringRate*, despite the fact that they receive few applications. They interview good risk-adverse candidates and local candidates. The short-listed candidates (including local candidates) come, as the *Medium* universities is the best they can pretend to. Both *HiringRate* and *LocalHiring* rate are high.
- The *Bad* (between 40 and 50) also receive few applications; they have a low *HiringRate* and a very high *LocalHiring* rate. Like *Medium* universities, they interview good risk-adverse and local candidates; however their top-listed candidates are more likely to defect if they can, and the *HiringRate* therefore decreases.

With respect to the *FameLoss* criteria, Fig 3 shows two groups of universities with significantly different behaviors: *Best* and *Good* universities recruit the best candidates they can attract whereas *Medium* and *Bad* universities recruit candidates

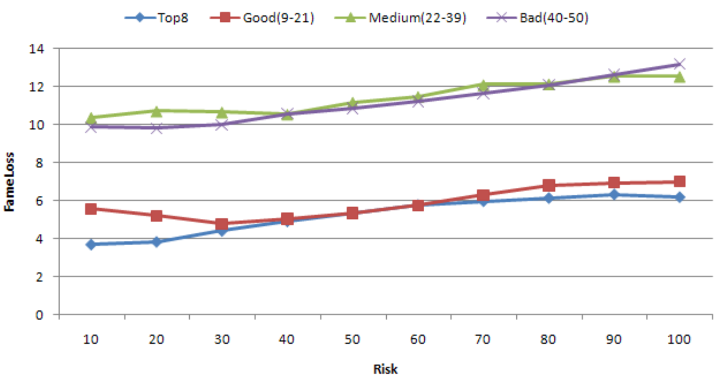

Fig. 3 Impact of risk-propensity on *FameLoss*, for Top, Good, Medium and Bad universities in the NoLearning case

with a disappointingly low rank. Furthermore, the *FameLoss* increases with the risk propensity (Fig. 3). This unexpected phenomena is blamed on the “subjectivity” effect involved in the preference $r(\cdot,t)$. The more risk-taker the university, the more it follows its own preference ranking, possibly selecting candidates with low rank due to subjective or local preferences.

5.3 Learning Universities Setting

In the *Learning* case, universities are allowed to increase/decrease their *SelfAssessment* depending on the success of the past hiring rounds. The results (averaged over 1000 independent runs) are measured after stabilization (1000 time steps). As could have been expected, learning makes universities more efficient: the *HiringRate* increases compared to the *NoLearning* case (Fig. 1). The *SelfAssessment* curve (Fig. 4) displays contrasted situations, mirroring the *HiringRate* curve (Fig. 1 - *NoLearning* case). Roughly speaking, the *Best* universities tend to overestimate themselves, the *Good* ones, to depreciate themselves (in order to anticipate the defection of their candidates), the *Medium* ones seemingly have no bias, and the *Bad* ones underestimate themselves.

The impact of the risk propensity is analyzed wrt the *FameLoss*, as the *HiringRate* does not discriminate among good and bad universities in the *Learning* case. Fig. 5 suggests that the risk propensity has no impact on the *FameLoss* except for the *Best* universities, that should rather have a conservative strategy (low risk propensity).

In the meanwhile, the *SelfAssessment* parameter features a high impact on the *FameLoss* (Fig. 6). If *Best* universities overestimate themselves, the weight of their subjective preferences increases, which results in recruiting lower-ranked candidates

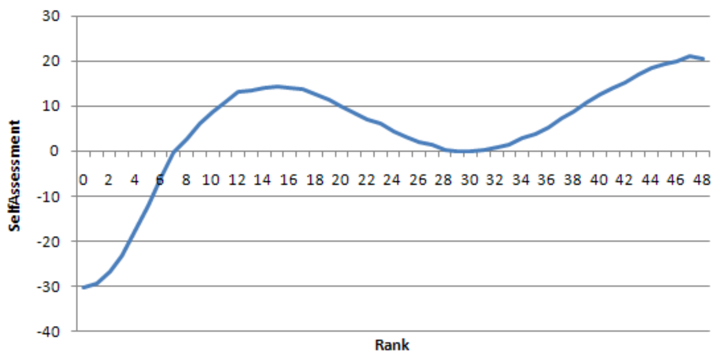

Fig. 4 Over and Under assessment of Universities vs Rank (learning setting). Top universities tend to over-estimate themselves; good and bad universities tend to under-estimate themselves; medium universities show no bias.

everything else being equal. Inversely, *Good* universities should not deprecate themselves in order to minimize the Fame loss. Quite the contrary, the *Medium* and *Bad* universities optimize their *FameLoss* by underestimating themselves.

6 Conclusion

This paper, resuming an earlier work devoted to the inefficiencies of Centralized Academic Labor Market [3], investigates how universities can increase their hiring rate. The proposed mechanism, relying on the self evaluation of the universities expectations, duly addresses the market inefficiencies regarding the Hiring rate. Extensive empirical investigations however suggest that this way of increasing the hiring rate can entail some undesirable Fame Loss. Specifically, Best and Good universities should not underestimate (respectively overestimate) themselves in order to recruit best or good candidates. Quite the contrary, Medium and Bad universities should deliberately underestimate themselves to secure the recruitment of acceptable candidates.

Further research will consider more comprehensive learning/optimizing setting for universities and candidates, allowing them to fine tune their behavioral parameters in order to maximize their consolidated Fame for universities, and their job quality for candidates.

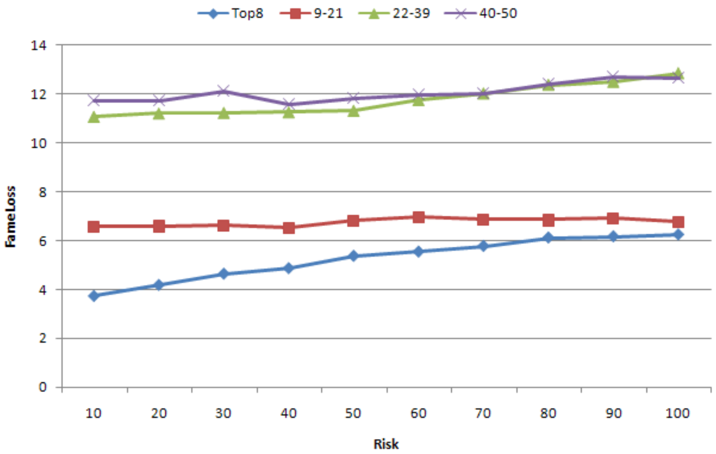

Fig. 5 Impact of risk-propensity on *FameLoss*, for Top, Good, Medium and Bad universities in the Learning setting

References

1. Musselin, C.: European academic labor markets in transition. Higher Education **49**(1-2) (2005) 135–154
2. Gale, D., Shapley, L.S.: College admissions and the stability of marriage. American Mathematical Monthly **69** (1962) 9–14
3. Caillou, P., Sebag, M.: Modelling a centralized academic labour market: Efficiency and fairness. In: ECCS 2008. (2008)
4. Baiou, M., Balinski, M.: Student admissions and faculty recruitment. Theor. Comput. Sci. **322**(2) (2004) 245–265
5. Ito, T., Parkes, D.: Instantiating the contingent bids model of truthful interdependent value auctions. In: 5th Int. Joint Conf. on Autonomous Agents and Multiagent Systems. (2006) 1151–1158
6. Combes, P., Linnemer, L., Visser, M.: Publish or peer-rich? the role of skills and networks in hiring economics professors. Labour Economics **15**(3) (2008) 423–441
7. Axelrod, R.: Advancing the art of simulation in the social sciences. Advances in Complex Systems **7**(1) (2004) 77–92
8. Tesfatsion, L.S.: A constructive approach to economic theory. In: Handbook of Computational Economics. Volume 2 Agent-Based Computational Economics of Handbooks in Economic Series. North-Holland (2006)
9. Phan, D.: From agent-based computational economics towards cognitive economics. In: Cognitive Economics. Handbook of Computational Economics. Springer Verlag (2004) 371–398
10. North, M., Collier, N., Vos, J.: Experiences creating three implementations of the repast agent modeling toolkit. ACM Transactions on Modeling and Computer Simulation **16**(1) (2006) 1–25

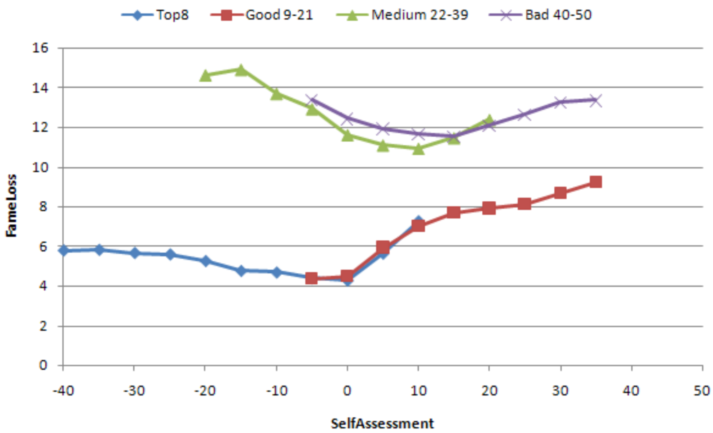

Fig. 6 *FameLoss* vs. university *SelfAssessment* for Top, Good, Medium and Bad universities in the Learning setting