

école —————
normale —————
supérieure —————
paris – saclay —————

2016-2017

Mémoire d'initiation à la recherche

On the performative effects of university
rankings: evidence of increasing social
polarization among French universities

Author:

Romain AVOUAC

Supervisor:

Hugo HARARI-KERMADEC

Abstract

In recent years, university rankings have been increasingly used to evaluate and compare performances of higher education institutions. Many have argued that the performative effects of rankings could eventually generate two tier education systems: on the one hand, a few number of "world-class universities" and on the other, a large share of second-class universities. As a result, a Matthew effect is likely to be observed: best endowed universities will keep getting additional resources while the situation of others will stagnate or even decline. The purpose of this master's thesis is to investigate whether such an effect is observed in the case of France, using data from the exhaustive database of French universities students registrations, from 2007 to 2015. We choose to test this assertion relatively to social composition of universities, as we believe that social capital is a resource that universities actually compete for. First, we run a static analysis which highlights an already very unequal university landscape as regards to social composition of their students. Then, we combine these results and the list of French universities that appear in main university rankings to show that they actually measure social realities and not merely academic performance. Finally, a dynamic analysis reveals a trend of increasing social polarization – irrespective of the tool used to measure it – with a stabilization in more recent years. These results support the hypothesis of an ongoing Matthew effect among French universities.

Contents

1	Introduction	1
2	A typology of French universities	4
2.1	Data	4
2.2	Method	4
2.2.1	Principal component analysis	4
2.2.2	Implementation of the PCA	5
2.2.3	Ascending hierarchical clustering	5
2.3	Results	6
2.3.1	A very unequal French university landscape	6
2.3.2	A typology of French universities	8
3	What do university rankings really measure ?	11
4	Is polarization increasing among French universities ?	15
4.1	Method	15
4.1.1	Using PCA coordinates as a university social composition index	15
4.1.2	Presentation and justification of the ER polarization measure	15
4.2	Results	18
5	Dicussion	22
6	Conclusion	23
	Appendices	27
	Appendix A	27
	Appendix B	28
	Appendix C	30

1 Introduction

Over the last decade, there has been a rising interest in quantitative performance indicators to evaluate higher education institutions. This trend is epitomized by the increasing number of international university rankings: the ARWU (Shanghai ranking), the Times Higher Education ranking, the QS World University Ranking, etc. This highlights a significant challenge posed by the massification of education systems and the ever increasing international mobility of students: how can one make informed choices about the best places to study at ? Given the considerable variety in size and purpose of higher education institutions, performance indicators and all the more so rankings appear as a straightforward and convenient way for students and parents to find their bearing between potential orientation choices.

Quantitative indicators are praised for enabling the comparison of universities performances in a consistent and transparent way by making productivity directly assessable. As a result, they have been increasingly used as benchmarks, i.e. performance metrics. For instance, through its Education and Training 2010 (ET 2010) framework, the European Commission has established several quantitative indicators for European education systems. Because these are non-coercive in nature, these policies rely on a form of peer pressure to further homogenize the education systems, since no member state wants to display the worst results. Thus, "it may be argued that benchmarking prescribes the necessity of competitiveness by embedding the managerial rationality into co-operation among member states." (BRUNO, 2006). Similarly, in France, the legislation governing public finance (LOLF) implemented in 2001 and the law on the autonomy of universities (LRU) of 2008 have also introduced the New Public Management (NPM) principles of market competition to the field of higher education, under the form of quantitative indicators (employment rates, success and attendance rates at exams, etc.) and "best practices".

Two types of quantification should be distinguished: cardinal – like the aforementioned quantitative indicators used for evaluation of higher education institutions – and ordinal, i.e. rankings. Although both play an undeniable role in the observed transformations of education systems, the substantial media coverage which surrounds every year the publication of the main university rankings give them an unrivaled influence on the strategies of actors in education systems. French "excellence initiatives" (IDEX) – mergers of French universities based on state funded projects – are a case in point of this idea, as they are often justified by the opportunity to improve their rank as a whole in international rankings (DOCAMPO, EGRET & CRAM, 2015).

As a consequence of this generalized use of rankings to guide education policies, university managers have become more inclined to use them. Since these benchmarks are promoted as targets to reach at country level, complying to them as a university likely means it will get additional funding. Besides, good ranks can also be used as an advertising tool in order to attract better students – especially foreigners, for whom additional resources are often offered – and well-known researchers. These, in turn, can serve as justification to charge higher tuition fees – on the model of Paris-Dauphine university or more recently some Master's degree of Paris-Saclay university.

The increasing use of rankings to assess and compare performances of higher education institutions has raised several concerns. The first concern is a general caveat associated with the use of quantitative indicators in policy making. As stated before, such tools are generally seen as objective measures of performance because of their numerical nature (PORTER, 1996). However, quantification is in fact a two-step process: first, it requires to agree upon what is to be measured, and then to measure it (DESROSIERES, 2008). Thus, quantifying a phenomenon is not neutral since the resulting figure is strongly dependent upon the implicit convention that was made in the first step. University rankings are a good illustration of this underlying problem of quantitative indicators. The most influential one, the Shanghai ARWU, is presented by its

creators as a ranking of the "top 500 world universities". This title conveys the idea that the ARWU is an accurate reflection of the relative performances of ranked universities. But the resulting order cannot be analyzed independently from the criteria used to build it – number of alumni winning Nobel Prizes, number of articles published in journals of Nature and Science, etc. – i.e. the convention made to assess university performance. However, such criteria appear to be of little relevance for students searching for the best universities to study at, whose primary concerns might more likely be quality of teaching, job openings and tuition fees.

Another critic relates to methodological weaknesses of the main rankings. After analyzing the Shanghai ranking methodology in the light of multiple criteria decision making (MCDM), some authors conclude "that the aggregation methodology is plagued by a number of major problems and that the whole exercise suffers from an insufficient attention paid to fundamental structuring issues" (BILLAUT, BOUYSSOU & VINCKE, 2010). Although often presented as objective and sound thanks to rigorous methodologies, university rankings actually prove inconsistent – in some situations, a university can be downgraded in the ranking even if it has improved as regards all the criteria used to build it – and their building actually involves many undocumented micro-decisions. We thus end up in a highly paradoxical situation in which the shortcomings of the main rankings are widely acknowledged, but those are still taken as a reference, particularly to guide higher education policies. For instance, the former French minister of higher education and research Valérie Pécresse stated in 2007 that "there is indeed scope for criticism of the Shanghai ranking but [...] as we can't escape from its influence, we must move up in it."¹

A third concern raised by critics of the use of rankings in higher education is that it might lead to a generalization of two tier education systems: on the one hand, a few number of "world-class universities" and on the other, a large share of second-class universities (VAN PARIJS, 2009). Every year since the Shanghai ranking creation, Harvard has gotten the maximum score of 100 and every other ranked university is thus evaluated based on how much it resembles it in terms of research standards. Therefore, this indicator generates strong incentives for countries and universities to transform their research models in order to meet the criteria used to compute the index: high number of publications per researcher – potentially at the expense of actual quality – a preference for hard science fields in order to get published in Nature or Science, and more generally a decrease of research activity that can't be assessed based on these criteria. According to proponents of NPM, rankings could thus spur universities to compete on a standardized basis, which should eventually entail an homogenization of productivity levels. However, for most countries, it is not possible nor desirable to have more than several world-class universities, since "many nations simply cannot afford to support such institutions" (ALTBACH, 2004). As a result, a Matthew effect (MERTON, 1968) is likely to be observed: a small number of elite universities will get the largest share of resources at the expense of others.

The main purpose of this master's thesis is to investigate the legitimacy of this last claim in the case of France, i.e. whether the French higher education system exhibits a Matthew effect. However, this requires to determine according to which resource we are going to test the presence of such an effect. As we are interested in finding out the underlying social reality of university rankings, we choose to assess it based on social composition of university students. It is a well-established sociological fact that individuals from privileged social backgrounds are more likely to go to university than their underprivileged counterparts. Although there has been an opening of higher education in France – between 1984 and 2009, the share of working class children who hold a university degree has increased from 6% to 24% (PEUGNY, 2013) – this figure alone is not sufficient to conclude that there has been a real democratization. Inequality in higher education cannot be reduced to its unequal access, it is also the product of unequally distributed resources in social and cultural capital. Students from high cultural capital backgrounds have interiorized

¹http://www.senat.fr/cra/s20070711/s20070711_8.html Our translation from French.

a culture which is valued by the education system, and thus possess an inner knowledge of what is necessary to succeed in it (BOURDIEU & PASSERON, 1964). Consequently, privileged students constitute a resource which is worth competing for from a university perspective.

If there is a Matthew effect going on among the French higher education system, we shall thus observe that some universities are becoming increasingly endowed over the years in terms of social composition of their public, and that the situation of other universities is worsening or at least stagnating. This hypothesis will be tested using data from the exhaustive database of French universities students registrations (SISE database), available from 2007 to 2015.

The first step will consist in building a typology of French universities in order to get an overview of the level of heterogeneity they already exhibit statically. Since the variables we use to assess the social composition of universities are quantitative, principal component analysis (PCA) with universities as individuals is a convenient tool to do so as it both enables one to get a good idea of the structure of correlations between the variables at hand, and to check if there already exists clearly defined groups of universities in terms of social composition of their public. This approach has been used previously in the case of French universities (BRUSADELLI & LEBARON, 2012) and indicates a high level of heterogeneity. My results strongly corroborate that conclusion. Furthermore, an ascending hierarchical classification (AHC) confirms the relevance of the distinct groups of universities highlighted by the PCA.

Then, using the results from the PCA, we will try to determine what is the actual reality which is measured by the main university rankings as regards French universities. BILLAUT, BOUYSSOU and VINCKE have shown that the Shanghai ranking, instead of merely measuring academic performance, assess a series of criteria which have nothing to do with research quality. Applying a similar analysis to French universities which appear in the Shanghai ranking, I find that it actually measures social composition of universities, scientificity level of their main field of teaching and their size.

The final step will consist in a dynamic analysis of social composition of universities in order to see whether the aforementioned Matthew effect is observed. Many different methods can be used to do so. BRUSADELLI and LEBARON monitor the evolution of a "democratization index of universities", and find a stagnation of democratization (Ibid.). However, the time span of their study is small (2007 to 2009) because of a lack of available data. At a worldwide level, HALFFMAN and LEYDESDORFF use the GINI coefficient to conclude that there is an increasing homogenization of top 500 universities of the Shanghai ranking (HALFFMAN & LEYDESDORFF, 2010). However, they compute their index based on the ARWU score of universities, and thus knowingly measure another reality (the evolution of inequalities in research output) that the one we are interested in. Besides, although the GINI coefficient can be useful to appreciate the evolution of inequalities in universities social composition, it doesn't enable one to see where the changes occur in the distribution. More importantly, in case of a clustering of two opposed groups of universities around their local means, an inequality measure would register an unambiguous decline whereas it would actually be a clear sign of a Matthew effect. To deal with this limit, I will also employ polarization measures, which are robust to this type of phenomenon and thus complement inequality measures conveniently (ESTEBAN & RAY, 1994). Irrespective of the tool used to measure it, I find an unambiguous trend of increasing social polarization with a stabilization in recent years. These results confirm the hypothesis of a Matthew effect as regards social composition of universities.

2 A typology of French universities

We want to determine whether French universities exhibit growing social polarization, i.e. if some universities are getting increasingly endowed in terms of social composition of their public while the situation of others is deteriorating. To assess that, we must first run a static analysis to check if there exists clearly identified groups of universities relatively to this criterion. Using a principal component analysis (PCA) followed by an ascending hierarchical clustering (AHC), we build a typology of French universities which confirms the existence of distinct groups of universities. This typology highlights a very unequal distribution of the social composition of students between these groups.

2.1 Data

The data used throughout this master's thesis are the exhaustive databases of French universities students registrations ("SISE registrations database"), available from 2007 to 2015. They were obtained following an access request filled on the Quetelet Network (French data archives for social sciences). Each year, the SISE survey lists all the students registered in various French higher education institutions. For each student, many variables describe the institution attended, the socio-demographic status of the student, his prior education and information about his current school level (prepared diploma, field of study, etc.). In 2015, the database included 1,579,145 students attending 85 different institutions (78 public universities, and several specific institutions such as University Paris-Dauphine or Sciences Po Paris).

Engineering and business schools are not included in our analysis. Similarly, French "grandes écoles" are absent although they are generally pointed out as an important determinant of the lack of democratization exhibited by the French higher education system, because of the social inequality which remains particularly high regarding their access (ALBOUY & WANECQ, 2003). Nonetheless, the absence of these different types of institutions is beneficial in our case since it enables us to study a population – French universities – which possess identical missions. We decided to keep University Paris-Dauphine and Sciences Po Paris in our sample as comparison purpose, since they can be seen as intermediary between French *Grandes Écoles* and public universities. Nonetheless, it should be noted that these institutions have specific characteristics – in particular, a selection process and higher tuition fees.

2.2 Method

2.2.1 Principal component analysis

Principal component analysis is a statistical method used to reduce the dimension of a given set of observations while maximally retaining the initial information, in order to facilitate data visualization. A set containing n individuals and p variables is located in a p -dimensional space and thus can't be represented if $p > 3$. PCA reduces the original set of observations into an inferior number of "dimensions" (principal components) q – generally 2 or 3 – which can then be plotted. Each of these components is built as a linear combination of the variables included in the PCA:

$$\begin{aligned} C_1 &= a_1^1 X_1 + \dots + a_1^p X_p \\ &\dots \\ C_q &= a_q^1 X_1 + \dots + a_q^p X_p \end{aligned}$$

where C_i is the i th principal component (PC), X_1, \dots, X_p are the variables and a_i^j is the weight of the j th variable in the i th PC. The weights a_1^1, \dots, a_1^p in the first PC C_1 are determined by maximizing its variance, so that C_1 explains the largest possible amount of information contained in the original data. The same is true for the second PC C_2 , with the additional constraint that C_1 and C_2 must be uncorrelated. The information explained by C_2 is thus additional, but also less important than the amount explained by C_1 . And so on for the next components. For technical reasons, we must also impose the constraints: $\forall i \in \{1, \dots, q\}, \sum_{j=1}^p (a_i^j)^2 = 1$.

In a PCA, all the variables must be quantitative. They should also be correlated: the more the original variables are correlated, the fewer components required for the retained information to be substantial. In practice, the variables are always centered for convenience and because this doesn't affect their dispersion. If the variables are heterogeneous in nature (different unit of measure, different scale, etc.), they should also be reduced prior to the PCA.

2.2.2 Implementation of the PCA

From the initial 41 variables available in the databases, 2 were relevant to characterize the social composition of a university: the socio-professional category of the students' parents and the type of *baccalauréat* obtained by the student – which in France is an important marker of social status (DURU-BELLAT & KIEFFER, 2008). After having identified the relevant modalities of these variables relatively to our object, i.e. the ones that would constitute potential sources of polarization, we calculated for each modality the percentage of students exhibiting it for each university. We thus obtained a new table with universities as individuals, and the proportions of students displaying each modality as variables. For instance, the row associated to Bordeaux 1 university shows that 77,14% of its students hold a scientific *baccalauréat*, 7,79% are children of workers, etc.

These 7 new variables were included as active variables in the PCA, i.e. they served to build the principal components as presented in the method section. Other variables such as the percentages of males/females and the percentages of undergraduate/graduate/PhD students were included as supplementary variables: they did not participate in the calculations of the PCs, but they were projected on the PCA resulting graphs so that their correlations with the other variables could still be analyzed. A table detailing active and supplementary variables, their definition as well as basic descriptive statistics is presented in appendix A (Table 5).

2.2.3 Ascending hierarchical clustering

Ascending hierarchical clustering (AHC) is an unsupervised statistical learning method which seeks to split the n individuals from the sample into a given number of clusters, based on an iterative process. Initially, there are n clusters: every individual constitutes a cluster. At each of its iteration, the AHC algorithm regroups the 2 closest clusters, i.e. between which a certain dissimilarity measure is minimized. The most commonly used metric is Ward's distance, which aims at maximizing the inter-cluster inertia. The algorithm stops when there is only one remaining cluster, which then contains all the individuals from the sample.

The main result of AHC is a dendrogram, i.e. a tree which represents all the mergers operated by the algorithm. The stake is then to determine where to cut the tree, i.e. the appropriate number of clusters to keep. Many empirical criteria have been put forward in the literature in order to do so (ROUSSEEUW, 1987 ; TIBSHIRANI, WALTHER & HASTIE, 2001). However these criteria always remain somewhat arbitrary. The appropriate number of clusters is generally to be found based on how interpretable the chosen clusters are in the light of knowledge of the studied field and various descriptive statistics of the clusters.

In our case, we used hierarchical clustering on principal components (HCPC): AHC was not performed directly on the individuals of the sample (French universities), but on their coordinates on the first 2 principal components generated by the PCA. PCA and clustering methods are indeed complementary tools to explore data. PCA enables to keep only the first principal components that carry the most information, and thus get rid of the noise in the data. As a result, AHC is more robust and results are more easily interpretable.

2.3 Results

2.3.1 A very unequal French university landscape

The PCA encompasses 7 active variables so that the original scatter plot of individuals belongs to a 7-dimensional space. On average, each dimension should summarize about $1/8 = 14\%$ of the original inertia. In 2007, the first two dimensions (PC) produced by the MCA summarize 54,16% and 25% of the initial information, i.e. respectively the average information of 4 and 2 dimensions (see Appendix B, Table 6). The reduction of the original cloud into 2 dimensions is thus of very good quality as 80% of its information is retained. These results enable us to unreservedly interpret the graphical representations obtained in the first factorial plan, i.e. in the coordinate system made of the first two principal components. The remaining 6 PCs each summarize less than the average information of a dimension and so can be considered as containing mainly noise. Consequently, the results on these dimensions didn't partake in the interpretation of the PCA.

Since we are going to analyze the results of the PCA only on the first factorial plan, the first step is to give a meaning to its axes. This is done by looking at the correlations between the initial variables and the first two PCs. These correlations are graphically represented on the correlation circle (Figure 1). Variables which are represented in black are active variables, whereas those represented in blue are supplementary.

The first PC (horizontal) is negatively correlated to the variable *pcs_plus* and positively correlated to the variables *ouvriers*, *bac techno* and *bac pro*. Besides, all these correlations are greater than 0.8 in absolute terms and can thus be safely interpreted. These results suggest that the first dimension is an indicator of social composition of universities. On the right of the first axis, we find universities with high proportion of workers children and students who hold either a technological or a professional *baccalauréat*. Inversely, universities with low proportion of such students and high shares of students from economically and culturally privileged backgrounds are located on the left.

The second PC (vertical) is mainly correlated with variables indicating the proportions of students based on the type of *baccalauréat* they hold: positively with scientific *baccalauréat* and negatively with literary and economic/social *baccalauréat*. It thus seem that the second dimension can be interpreted as an indicator of the main field of study offered by each university: predominantly scientific universities will be located at the top of the first factorial plan whereas universities dominated by social sciences and literature will be located at the bottom.

The fact that the variable *bac_S* is negatively and significantly correlated to the first component whereas the opposite is true for variables *bac_ES* and *bac_L* – albeit to a lesser extent – confirms the socially differentiated values of *baccalauréats* of different types. This implies that universities which exhibit a large share of students holding a scientific *baccalauréat* also have higher chances of displaying a great proportion of students coming from privileged backgrounds. However, the interpretation made of the second dimension holds true. Indeed, the orthogonality of principal components implies that the interpretation of the second axis should be made for a given social composition. This ensures that the opposition highlighted by the second dimension is effectively the result of different fields of study of universities.

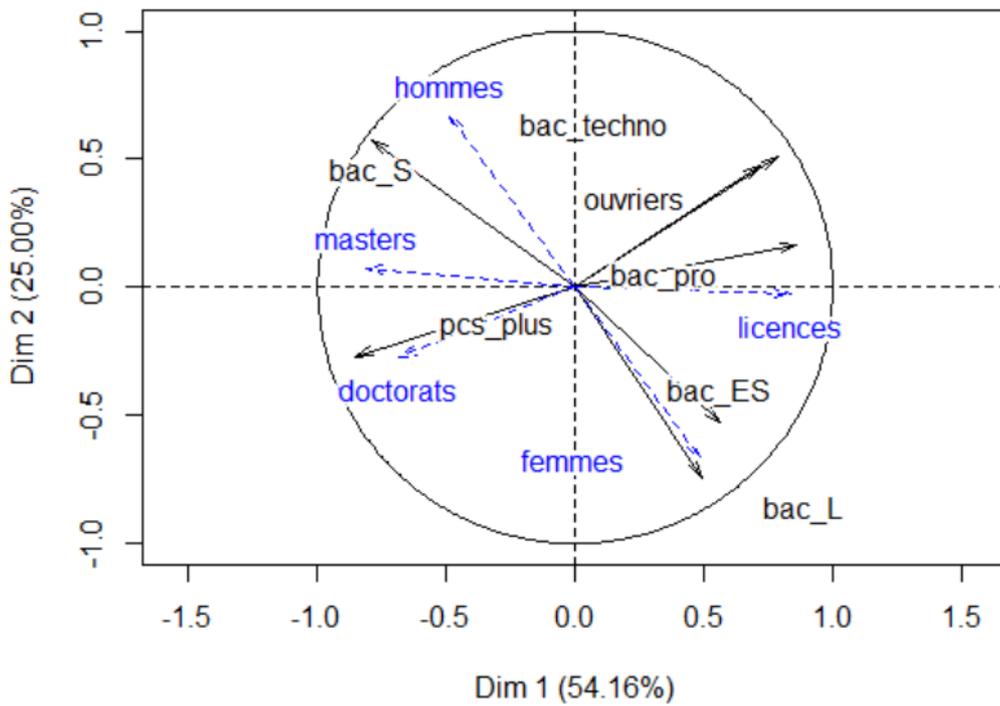


Figure 1: Graph of 2007 PCA active (black) and supplementary (blue) variables on dimensions 1 (horizontal) and 2 (vertical)

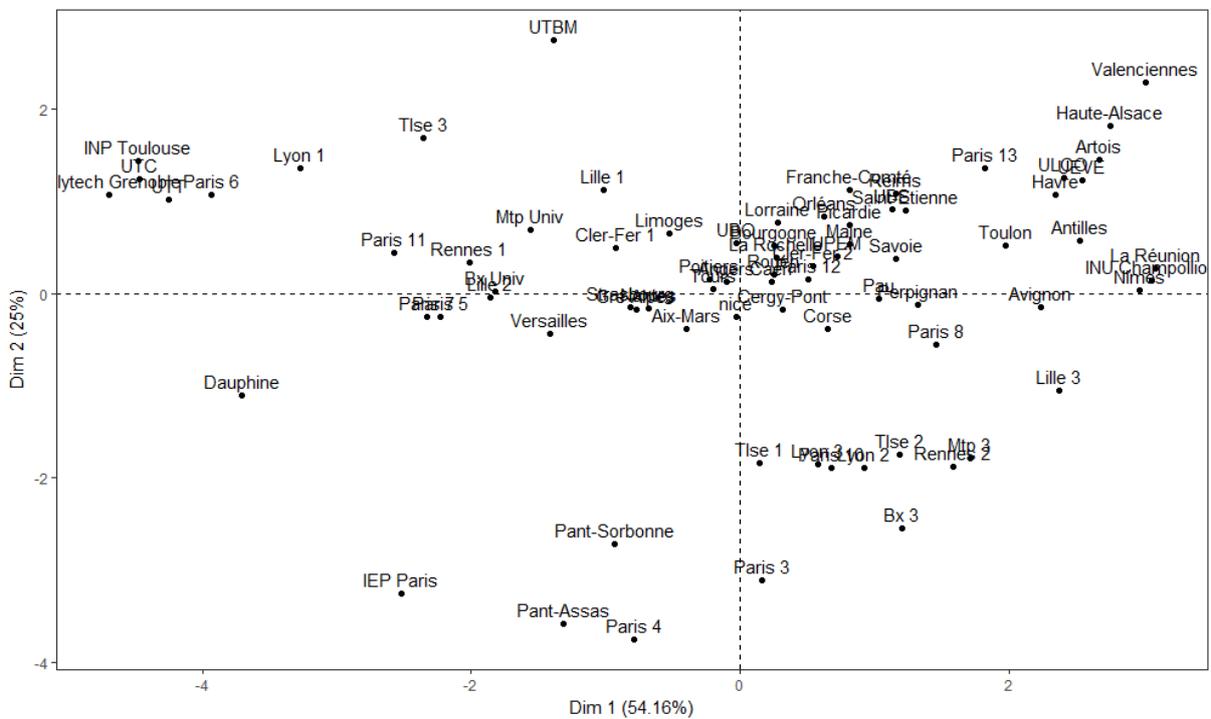


Figure 2: Graph of 2007 PCA individuals on dimensions 1 (horizontal) and 2 (vertical)

The correlations between the 2 first PCs and the supplementary variables – represented in blue on Figure 1 – also provide pertinent information. These variables didn't contribute to building the axes, they were simply projected on the first factorial plan. The fact that these variables exhibit strong correlations with the PCs thus appear all the more significant.

The supplementary variables *hommes* and *femmes* are highly correlated to the second PC, and respectively very close from the active variables *bac_S* and *bac_L / bac_ES*. This result confirms that although female students tend to be more successful in their studies whatever the level and the field of study, they are still over-represented in literary and social sciences fields as opposed to their male counterparts (ROSENWALD, 2006).

The first PC is positively correlated to the supplementary variable *licences* whereas it is negatively correlated to the variable *masters*. This indicates that universities which exhibit a large share of students from privileged social backgrounds also tend to have an important proportion of students at a master level. Conversely, universities in which underprivileged students are numerous generally exhibit a large share of undergraduate students. These findings corroborate the assessment that not only children from underprivileged background are much less likely to access higher education, but those who do access it have also less chances to hold a master's degree (PEUGNY, 2013). The quality of the representation of the variable *doctorats* (proportion of PhD students) is poor – the sums of squared cosinus that represents its proximity with the first PCs is equal to 0.24 whereas a good representation should exceed 0.7 – which doesn't enable us to interpret it.

Now that the two first PCs have a clear interpretation, we can analyze the graph of the individuals in 2007 (Figure 2). The first dimension offers the most interesting results as regards to the object of this master's thesis. Indeed, the extent to which universities are spread over this dimension shows how much unequal the French university landscape already is when we look at it in a static manner. On the left side, we can observe an over-representation of Parisian higher education institutions (Dauphine, Sciences Po Paris, Panthéon-Sorbonne, Panthéon-Assas, etc.) and more generally universities from major French cities. Conversely, we find on the right of the graph universities from relatively small provincial cities as well as universities from French overseas departments. These results corroborate findings according to which a large part of inequalities between universities is a reflection of unequal social compositions of different French regions (NICOURD, SAMUEL & VILTER, 2011).

2.3.2 A typology of French universities

Our purpose is to find out whether polarization is increasing, which requires to monitor the evolution of clearly defined groups of universities. The dispersion of individual on the first factorial plan suggests that a typology, i.e. distinct groups of universities, can be built based on the PCA results. To do so, we ran an ascending hierarchical clustering using the coordinates of universities on the first 2 PCs from the PCA.

Based on the empirical criterion of minimization of the relative loss of inertia, the algorithm suggests a 3-clusters (Figure 3) configuration (3CC). It distinguishes a group of rather scientific universities with high social capital on the left, a group of rather literary universities with average social composition on the middle, and a group of rather scientific universities with high shares of students coming from underprivileged backgrounds. However, the privileged and underprivileged groups appear coarsely separated in the middle. The 4-clusters (Figure 4) configuration (4CC) appears more satisfactory in this regard since it enables to distinguish a supplementary group of scientific universities exhibiting average social composition. The 5-clusters configuration (5CC) merely distinguishes among privileged universities a small group of universities exhibiting very high social capital.

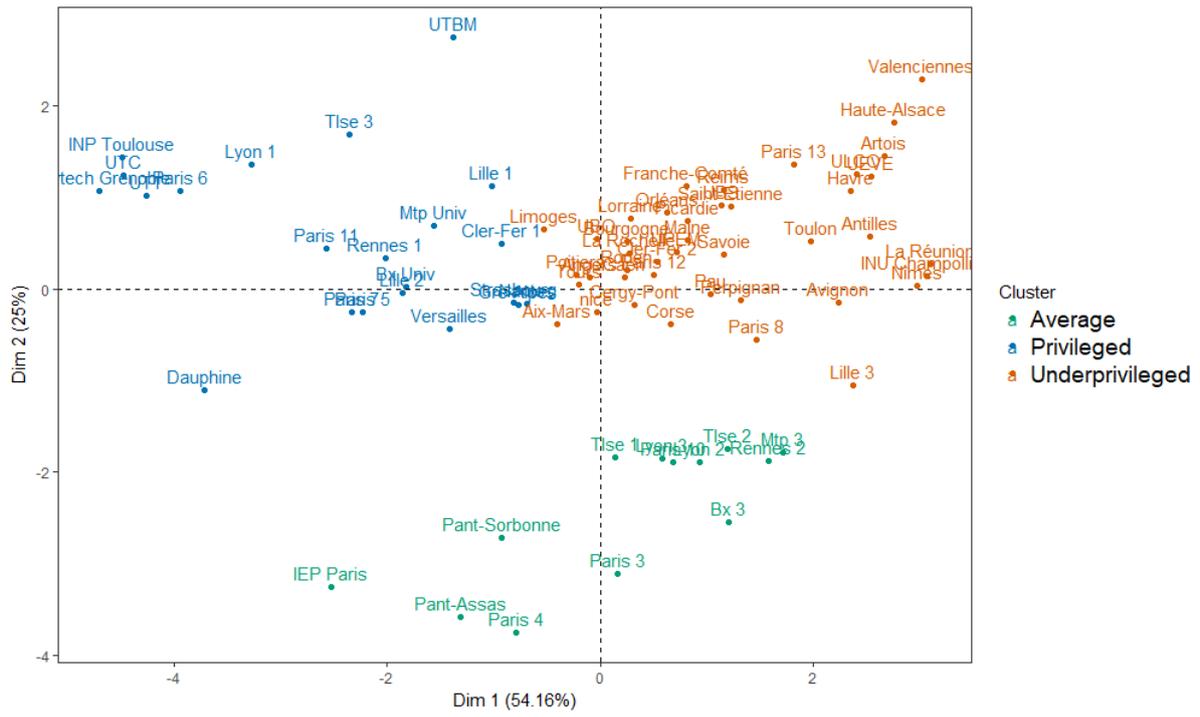


Figure 3: HCPC results with a 3 clusters configuration (3CC)

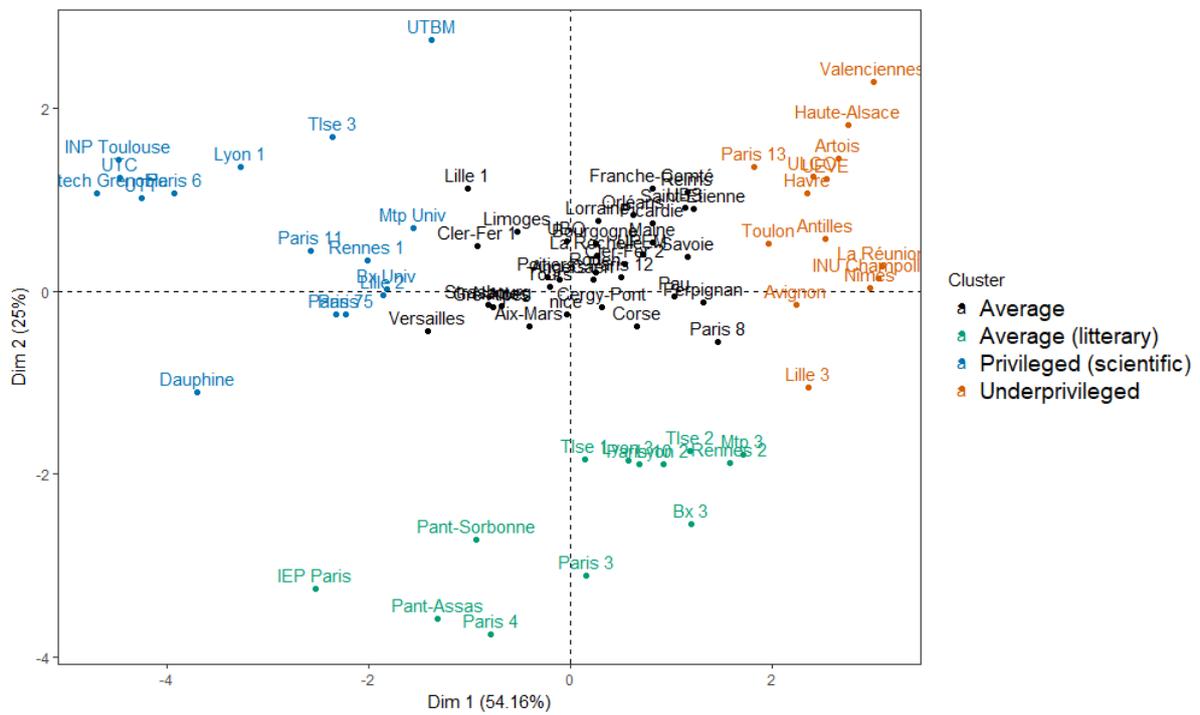


Figure 4: HCPC results with a 4 clusters configuration (4CC)

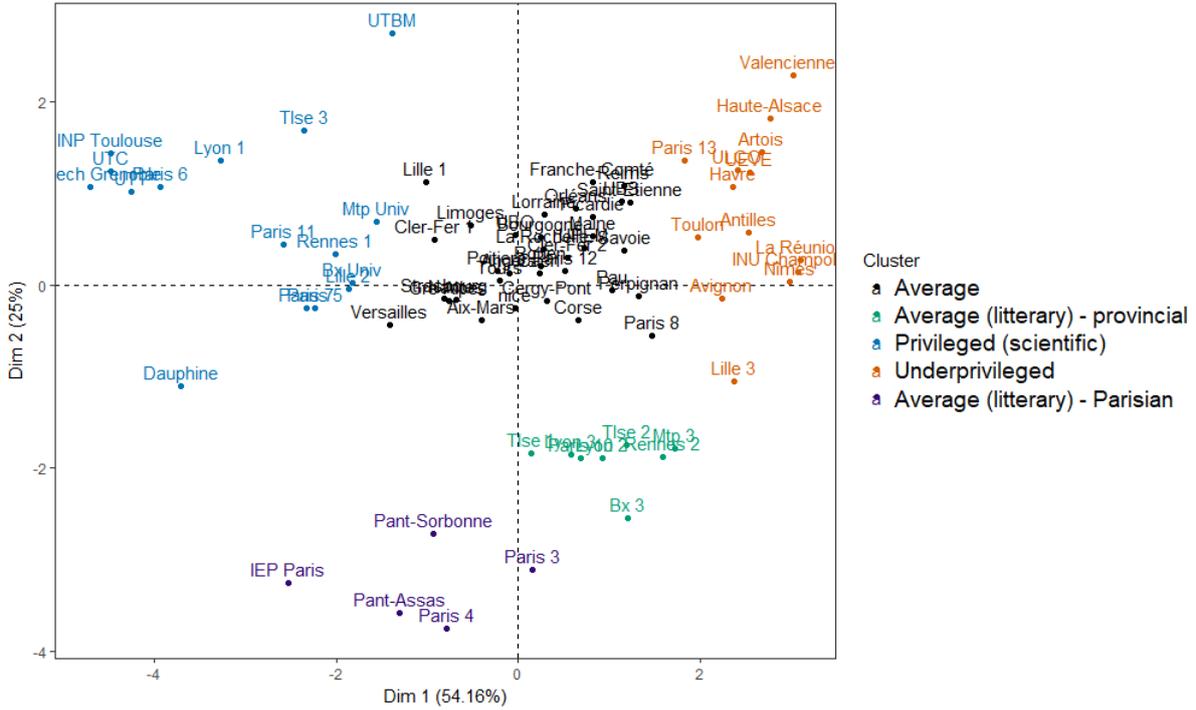


Figure 5: HPCP results with a (modified) 5 clusters configuration (M5CC)

We must now decide which configuration(s) will be used for the following analysis. The subgroup generated by the 5CC actually appears highly heterogeneous : it aggregates Paris-Dauphine – a university which is mainly specialized in economics, finance and management – with universities specialized in fundamental science and public engineering schools. Thus, this configuration is not retained. The gain of within-clusters inertia associated to the 4CC relatively to the 3CC isn't particularly high, meaning that it doesn't increase by much the homogeneity of the different clusters in terms of inertia. But more relevant is the fact that the 4CC results in more homogenous groups in terms of social composition, which is important as we seek to measure polarization based on this criterion. However, it can be argued that the 4CC "average (litterary)" cluster (in blue on (Figure 4)) appears insufficiently homogenous regarding social composition. We will thus also consider a modified 5-clusters (Figure 5) configuration (M5CC) in which we distinguish among this group a subgroup of Parisian universities and Sciences Po Paris (on the left, i.e. better endowed in terms of social composition) and provincial universities (on the right). We retain 4CC as the reference for the rest of this paper, but we will also perform the analysis on 3CC and M5CC since the three of them exhibit pros.

We can now further characterize the four clusters of 4CC. On the top left corner, we find scientific universities with large share of students coming from socially favored background. At the bottom, we have a group of universities predominantly dealing with social sciences and literacy, heterogenous in terms of social composition. On the middle of the graph, we observe a large cluster of universities, average as regards both dimensions. Finally, the right of the graph exhibits a group of universities, rather scientific, with a majority of students coming from underprivileged backgrounds.

A table showing the mean of each active and supplementary variables of the 2007 PCA according to the different clusters of the 4CC is presented in appendix B (Table 7). These statistics confirm the relevance of the clusters chosen since they highlight a strong inter-clusters heterogeneity as regards the variables used in the PCA.

3 What do university rankings really measure ?

University rankings all aim at answering the same question: what are the best universities in the world ? The declared purpose is to "provide young people and their families with trusted guidance to help with this [university] decision" ².

Many criticisms have been made concerning the inadequacies of the most influential rankings to prove relevant as tools to assist such a decision. First, they mainly aim at assessing research performance whereas students concerns might more likely be quality of teaching, job openings, tuition fees, etc. Furthermore, even if academic performance alone was relevant to gauge the quality of a university, the main rankings actually appear to poorly assess it because of a flawed methodology (BILLAUT, BOUYSSOU & VINCKE, 2010). As a result, rankings generally end up measuring another reality that the one they were designed for. For instance, BILLAUT, BOUYSSOU & VINCKE derive from the criteria used to build the Shanghai ranking its implicit definition of a "good university", which "roughly consists in an institution:

1. that is old and has kept its name (preferably a simple English one without diacritical signs) throughout its history,
2. coming from a country that has not experienced major political and social changes and that is peaceful democracy,
3. coming from a country in which the organization of the higher education system is simple (no dual system, no research centers around),
4. coming from an English speaking country,
5. having much freedom with respect to its governance,
6. having much freedom in hiring and firing staff and deciding on salaries,
7. being well funded,

and then point out that "it more or less corresponds to the definition of the Ivy League (plus "Oxbridge") and conclude that "unsurprisingly, these institutions are quite well represented at the top of the Shanghai ranking".

It appears clear that this definition doesn't quite reflect the nature of French universities – many were split and regrouped throughout their history and they have little freedom over their governance and staff management ; researchers often have a dual affiliation (university and a research unit at CNRS) ; there is a substantial parallel system with *classes préparatoires* and *Grandes Écoles*, etc. Thus, we may wonder what are the characteristics of French universities which do appear in the main rankings. In particular, we can check whether they fit the alternative ranking we built using PCA in order to investigate the underlying social reality of these rankings.

The three most influential international rankings are the ARWU (the "Shanghai ranking"), the QS World University Rankings (QS) and the Times Higher Education World University Rankings (THE). The purpose of the researchers behind the ARWU is unambiguous: "institutions are ranked according to their academic or research performance". By contrast, the QS and the THE rankings creators claim to encompass some qualitative criteria such as academic

²<https://www.timeshighereducation.com/student/advice/world-university-rankings-explained>
Other rankings exhibit similar goals in their presentation.

reputation (QS), employer reputation (QS), teaching reputation (THE), etc. However, in both of these rankings, the weights assigned to each criterion are such that more than 60% of the final score depends on academic performance. As a result, these rankings are very similar to the ARWU. In particular, the French universities which are ranked are almost the same for the three rankings each year. We will thus use the ARWU for our analysis as it remains the most influential overall, but the results would very likely be identical if we performed it using the THE or the QS rankings.

In Figure 6, we plotted the 2015 data on French universities as supplementary individuals on the 2007 PCA, and marked the universities which were ranked in the 2015 ARWU. Those appear overly represented in the left part of the graph, and all the more so in the top left corner. This result suggests that the French universities which are ranked in the ARWU tend to exhibit large shares of students coming from privileged backgrounds who hold a scientific *baccalauréat*.

To confirm this conjecture, we performed a logistic regression in order to model the probability for a French university to appear in the Shanghai ranking depending on its position on the PCA graph, i.e. on its respective coordinates on the first two principal components. The number of students (in thousands) per university was also included to control for institutions size. The results are presented in Table 1. The advantage of this approach is that since the PCs are synthetic variables, it avoids the multicollinearity issues that would likely have occurred if we performed the regression using directly the initial variables. Indeed, the PCA showed that several variables were highly correlated (see Figure 1).

The variable *Dim.1* exhibits a statistically significant effect, at the one percent significance level. Since universities are in reverse order on the first PC – universities which exhibit privileged social composition have negative coordinates and conversely – the negative sign of the coefficient of *Dim.1* implies that the higher the social composition of a university, the better its chances of appearing in the ARWU all other things being equal.

Contrary to the conjecture we made from Figure 6, the coordinate of a given university on the second PC doesn't seem to exert a statistically significant effect ($p=0.25$) on its probability of being ranked. However, it is interesting to note that among the universities which are ranked in the ARWU and located close to the center of the graph, we find a substantial group of universities – Strasbourg, Aix-Marseille, Lorraine, Auvergne, Bordeaux and Grenoble-Alpes – which are the result of recent mergers of several previous institutions. Their relatively average positions on the second PC is due to the fact that they were built by aggregating heterogeneous universities in terms of predominant field of study – generally, a mainly scientific university and one or several social sciences/literary universities. The effect of mergers could thus mitigate the effect of the field of study on the probability of being ranked.

To confirm this, we ran the same analysis on 2012 data, i.e. before the mergers of the universities of Grenoble-Alpes, Bordeaux and Montpellier. In 2012, their scientific component – respectively Grenoble 1, Bordeaux 1 and Montpellier 2 – were already ranked in the ARWU, and were logically located at the top of the graph. And the logistic regression results then showed a statistically significant ($p<0.05$) positive effect of the variable *Dim.2*, in accordance with our initial conjecture. This averaging effect of mergers isn't limited to the three universities we just mentioned, nor to the second PC : mergers also aggregate heterogeneous institutions in terms of social composition. As a result, it is possible that the magnitude of the effect of social composition (*Dim.1*) on the probability of being ranked in the ARWU is mitigated as well.

Finally, the variable *nb_students* exerts a statistically significant effect on the probability of a university to appear in the ARWU, at the one percent significance level. For a given position on the first factorial plan of the PCA, a greater size – approximated from the number of registered students – also increases the chances for a university to be ranked. This result corroborates the conclusions of DOCAMPO, EGRET and CRAM (op. cit., 2015) according to which "the ARWU

ranking methodology favors universities that are large and excellent over those that are small and excellent".

However, our results also indicate that, as far as French universities are concerned, the ARWU doesn't (only) measure "excellence". Similarly to what BILLAUT, BOUYSSOU & VINCKE did in a general manner, we can characterize the implicit definition of a "good French university" according to the ARWU, which consists in an institution:

1. that exhibits a high share of students coming from socially privileged backgrounds
2. of great size, possibly following a merger of previous institutions.
3. in which fundamental science tends to be the predominant field of study

Against that background, it seems very likely that a Matthew effect will eventually be observed among French universities. Rankings, like other quantitative indicators, don't just describe the reality at a specific point in time, they also have a performative effect: agents adapt their behavior accordingly. For instance, DOCAMPO, EGRET and CRAM report that France exhibits "many examples where merger proposals have been coupled politically to the claim that ranking or "visibility" would be enhanced." And since these mergers are generally accompanied by substantial public funds, it is likely that these "world universities" in progress will keep expanding and thus improve their ranking positions. And this in turn will enable them to attract better students – especially foreigners – as well as justify higher tuition fees, which will likely result in even higher levels of polarization in terms of social composition.

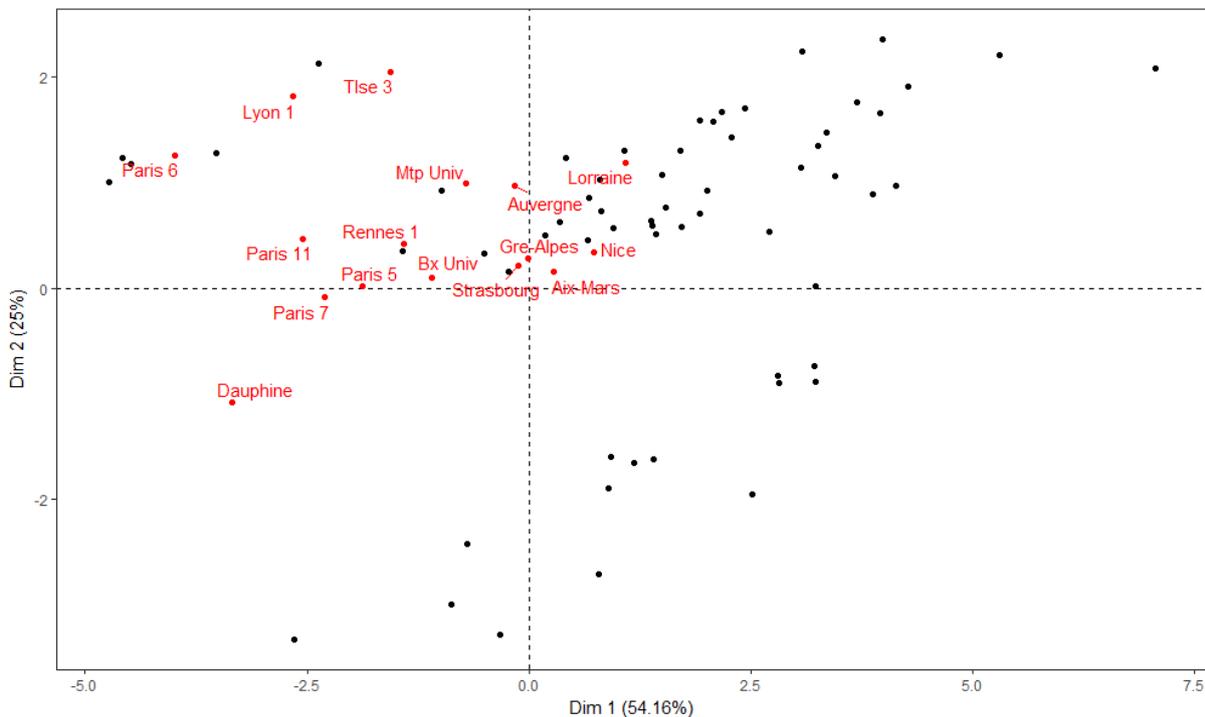


Figure 6: French universities ranked in the 2015 ARWU

Table 1: Logit regression modelling the probability for a university to appear in the 2015 ARWU

<i>Dependent variable:</i>	
	shanghai
Dim.1	-0.936*** (0.301)
Dim.2	0.553 (0.480)
nb_students	0.187*** (0.051)
Constant	-6.502*** (1.604)
Observations	77
Log Likelihood	-15.239

Note: *p<0.1; **p<0.05; ***p<0.01

4 Is polarization increasing among French universities ?

The typology we built shows that the French university landscape already appears highly unequal in terms of social composition of universities when we look at it on a static level. We now seek to find out whether this situation is evolving according to a Matthew effect, i.e. if universities which are already the most endowed in terms of social composition are getting even "richer", while the situation of the others is stagnating or worsening. First, we will show how we can derive from the PCA a university social composition index which can be used to monitor its evolution over time. Then, we will use this index to compute the GINI coefficient and the ER polarization measure and track their evolution.

4.1 Method

4.1.1 Using PCA coordinates as a university social composition index

An index that immediately comes to mind is the variable constituted by the coordinates of universities on the first principal component of the PCA, as it represents a continuum of universities ranging from those with poor social composition to those with high social composition. The idea of deriving an uni-dimensional measure of a multi-dimensional phenomenon using PCA has been used previously in the literature, particularly to compute socio-economic status (SES) indices (GWATKIN et al., 2000 ; FILMER & PRITCHETT, 2001 ; MCKENZIE, 2005 ; VYAS & KUMARANAYAKE, 2006). Albeit not exactly similar, the purpose of all these studies appears analogous to ours. Since accurate individual income data are very costly to collect – especially in developing countries – these authors use PCA coordinates to derive an uni-dimensional approximation of individual SES using available data (household ownership of durable assets, housing characteristics, etc.). Then, they use the obtained measure for comparison (between countries, over time, etc.). Similarly, we use PCA to derive from various data on social characteristics of each university students a social composition index of universities, in order to use it for comparison over time.

For this dynamic analysis to be sound, we must ensure the comparability of the index across time. Although the structure of correlations between the variables used in the PCA remains very stable over time – for comparison purposes, graphs of 2007 and 2015 PCA are presented in appendix C (Figure 9 and 10) – the first two principal components still vary somewhat. As a result, we can't directly compare the coordinates of universities on the first PC of subsequent PCA across time. A satisfactory solution to this issue is to set the axes once and for all and then to monitor the evolution of coordinates relatively to these axes. To do so, we simply project the observations of the 2008 to 2015 data sets on the axes built in 2007 – they thus constitute supplementary individuals, similarly to the supplementary variables we presented earlier and which didn't partake in building the axes. Then, we obtain a social composition index of universities for each year which we can safely use to compute and monitor inequality and polarization measures over time.

4.1.2 Presentation and justification of the ER polarization measure

Our purpose is to find out whether we observe the aforementioned Matthew effect among French universities using the index we derived from the PCA. To achieve such a goal, the economic literature traditionally resorts to inequality measures such as the GINI coefficient in order to assess whether the concentration of the statistical distribution increases. However, a shortcoming

of such measures is that they fail to adequately distinguish between "convergence" to the global mean and "clustering" around "local means". For instance, if we observe that a group of poor social capital universities converge to their common mean while the same is true for a high social capital group, inequality measure would register an unambiguous decline whereas it would actually be a clear expression of a Matthew effect. A way to deal with this issue is to employ polarization measures, which are robust to this type of phenomenon.

Polarization measure has been first introduced by ESTEBAN and RAY (op. cit., 1994). Since then, many authors have worked on refining this measure and applying it to various populations (WANG & TSUI, 2000 ; DUCLOS, ESTEBAN & RAY, 2004 ; ANDERSON, 2011). The starting point of this literature is to define a population as "polarized" when it can be "grouped into significantly-sized clusters, such that each cluster is very "similar" in terms of the attributes of its members, but different clusters have members with very "dissimilar" attributes". According to the authors, many social phenomena that are sources of social tensions (labor market segmentation, racial issues, nationalist conflicts, etc.) have more to do with the opposition of well-defined clusters of attributes – e.g. revenue – than with the inequality of their distribution over the population.

ESTEBAN and RAY then define an axiomatic of polarization, i.e. conditions under which polarization increases in a given population. These axioms impose three conditions that must be met by the distribution of attributes: there must be a high degree of homogeneity within each group, a high degree of heterogeneity across groups, and there must be a small number of significantly sized groups. Under these conditions, the authors provide a mathematical framework of polarization that we summarize below.

Let $(\pi, y) \equiv (\pi_1, \dots, \pi_n, y_1, \dots, y_n)$ be a distribution where π_i is the weight attributed to the cluster i and y_i the mean log income of the cluster i . Let $I(p)$ be an "identification" function, which accounts for the fact that an individual feels a sense of identification with other individuals who have the same income as him. Let $a(\delta(y_i, y_j))$ be an "alienation" function, which expresses the alienation felt by an individual from others that are "far away" from him in the income distribution. Finally, let $T(I, a)$ be an "effective antagonism" function, which accounts for the fact that an individual's feelings of identification may influence the expression of his alienation.

The total polarization P in the society is then postulated to be the sum of all the effective antagonisms:

$$P(\pi, y) = \sum_{i=1}^n \sum_{j=1}^n \pi_i \pi_j T [I(\pi_i), a(\delta(y_i, y_j))]$$

And the authors further demonstrate that, for this measure to meet the aforementioned axioms of polarization, it must be of the form:

$$P^*(\pi, y) = K \sum_{i=1}^n \sum_{j=1}^n \pi_i^{1+\alpha} \pi_j |y_i - y_j|$$

with $K > 0$ a constant and $\alpha \in [0; 1.6]$. This expression is the ER polarization measure.

This form of the polarization measure is directly computable by choosing values of K and α . The constant K can be used for population normalization, but has no influence whatsoever on the order. The parameter α can be interpreted as a "sensitivity to polarization" of the population, which depends on the degree of homogeneity of the clusters. If $\alpha = 0$, the groups are supposed to be perfectly homogenous and polarization is assumed to be maximal. However, the "sensitivity to polarization" of the population is difficult to assess *a priori*. A common practice in the literature while using the ER polarization measure is thus to compute it for different values of α in order to test the robustness of the results to changes in this parameter.

The formula of the ER measure bears resemblance to the GINI coefficient formula. In fact, for $\alpha = 0$ and $K = \frac{1}{2n^2\bar{x}}$, P^* is the GINI coefficient, i.e. a measure of the inequality among clusters. As the authors note, "it is precisely the fact that population weights are raised to a power exceeding unity in the formula above that gives rise to distinctly different behavior of a polarization measure".

Several elements suggest that the polarization framework is appropriate to our particular case. First, the necessary hypothesis appear to be met: the AHC has produced a small number (4) of homogenous groups, which are fairly dissimilar from one another. The inter-clusters heterogeneity can be appreciated on the table showing the mean of the variables included in the PCA according to the different clusters (Appendix B, Table 7). Another condition is that the clusters should be "significantly sized". In view of the vagueness of this hypothesis, it is difficult to know if the cluster of privileged literary universities is significant enough, as it is only composed of 5 institutions. To deal with this issue, we will also compute the polarization measure using the 3-clusters configuration, in which each cluster represents about 1/3 of the universities sample.

Second, the identification/alienation framework fits the social reality of French universities. For illustration purposes, in 2015, 21 presidents of French universities collectively signed a public press article in which they denounced "the obsession for international rankings which prevails in the choices of public funding of universities"³. Most of these universities happen to be located in the "underprivileged" cluster of our 4-clusters configuration of the AHC, some in the "average" cluster while none of them belong in the two privileged clusters. Symmetrically, a recent article was published by 7 presidents of various French institutions to plead for a "gathering of the best research components of the higher education institutions, within several research universities which possess the potential of being competitive at a worldwide level"⁴. Of these seven institutions, one appear in our sample – University Paris-Dauphine, which is located on the far left of the first PC – but others are Grandes Écoles and were thus not present in our initial database. However, it is very likely that they would also be located on the far left of the graph since these institutions usually exhibit very high shares of students coming from privileged backgrounds (ALBOUY & WANECQ, 2003). These examples highlight the identification existing between universities which exhibit similar social compositions as much as the alienation they feel towards universities located at the other end of the social spectrum.

Finally, the polarization measure can be computed based on the social composition index of universities we derived from the PCA. In the original paper, ESTEBAN and RAY use log incomes of individuals to compute the ER measure, but also explain that "even if the attribute space is multidimensional, the analysis is undisturbed if the differences between attributes can be represented by a metric." This is precisely the case in our study, the multidimensional attribute being the social composition of a university and the metric its coordinates on the first principal component.

We should note that both the GINI coefficient and the ER polarization measure can only be computed on positive metrics – traditionally, income or log income. However, the coordinates of universities on the first PC range from negative to positive, and the order is reversed: best quality universities have the lowest (negative) values while worst quality universities have the highest coordinates. We must thus apply a normalization in order to be able to use these metrics to compute the inequality and polarization measure. First, for each year, we multiply the vector of coordinates by -1 in order to get the true ordering of universities based on their quality. Second,

³<https://blogs.mediapart.fr/edition/les-invites-de-mediapart/article/290515/quel-avenir-pour-l-enseignement-superieur-et-la-recherche-francais> Our translation from French.

⁴https://lemonde.fr/idees/article/2017/05/30/enseignement-superieur-rapprocher-les-meilleurs-organismes-de-recherche-des-meilleurs-etablisements_5135898_3232.html Our translation from French.

we compute the minimum coordinate of all the vectors associated to each year and subtract this value to each vector. As a result, we have for each year a vector of positive coordinates in the good order on which we can calculate the desired measures. This approach has been used previously in the literature (SAHN & STIFEL, 2010 ; BHORAT & VAN DER WESTHUIZEN, 2013).

The downside of this method is that it is rather ad-hoc: by adding the same constant to all values of each vector, we do not change the order of the coordinates and the distance between one another, but we modify the absolute total. As a result, we won't be able to interpret the values of the GINI coefficient or the polarization measure *per se*. But as we are interested in monitoring their evolution and not in interpreting their absolute value, we believe that this concern doesn't mitigate the validity of our results.

Finally, we will also employ a measure put forward by MCKENZIE, in line with the SES literature presented earlier (MCKENZIE, 2005). This measure was precisely built to assess inequality over groups and/or over time of index built in an analogous way to ours, i.e. by using individuals coordinates on the first PC of a PCA. In its original form, the MCKENZIE measure is:

$$I_c = \frac{\sigma_c}{\sqrt{\lambda}}$$

where σ_c is the standard deviation of coordinates of individuals from group c on the first PC, and λ is both the eigenvalue corresponding to the first principal component, and also the variance of coordinates on the first PC over the whole sample. To monitor the evolution of inequality within a whole population over time, we just have to take for λ the eigenvalue corresponding to the first PC of the reference year, and for σ_c the standard deviation of coordinates of individuals from year c on the first PC – obtained by projecting them as supplementary individuals on the PCA built in the reference year. Thus, in our case, the measure will necessary be equal to one in 2007 (our reference year). On subsequent years, it will be greater than one if inequalities have risen, and conversely.

The main benefit of this indicator is that MCKENZIE proves that it "satisfies the standard four axioms required of an inequality measure" (anonymity, scale independence, population independence and the PIGOU-DALTON transfer principle).

4.2 Results

As we are going to use changes in our index – i.e. coordinates on the first dimension of the PCA – over time, the first thing we can do is plot the evolution of these coordinates. Figure 7 represents the 2007 PCA with both the 2007 and 2015 universities population, the latter being projected as supplementary individuals. Labels of universities were removed to ensure legibility. Although it is hard to conclude with any certainty at this point, this graph already shows first signs of a Matthew effect within that period. Indeed, there seems to be a somewhat general shift to the right of universities coordinates between 2007 and 2015. However, this shift appears highly prominent for underprivileged universities, whereas it seems less and less pronounced as we move towards universities with high social capital. This graph thus suggests that social composition is declining for already underprivileged universities, while there isn't any sign of a potent social opening among privileged universities.

To support this first result, we plotted in Figure 8 the evolution over the year of the barycentres of the four clusters derived from the AHC. Similarly, we find that all the clusters have their middle point shifted to the right over the year. But the speed of the shift appears greater for the average and particularly for the underprivileged group of universities, than for the privileged cluster. Furthermore, it is interesting to note that the shift happens diagonally and thus seems to

follow the "pure social origin" axis constituted by the variables *boursiers*, *ouvriers*, *bac_techno* and *pcs_plus* (see Figure 1). This suggests that the shift to the right of universities is done for a constant share of the different types of *baccalauréat* per university. We can thus unambiguously interpret this shift to the right of the average and underprivileged groups of universities as a decrease of their social composition.

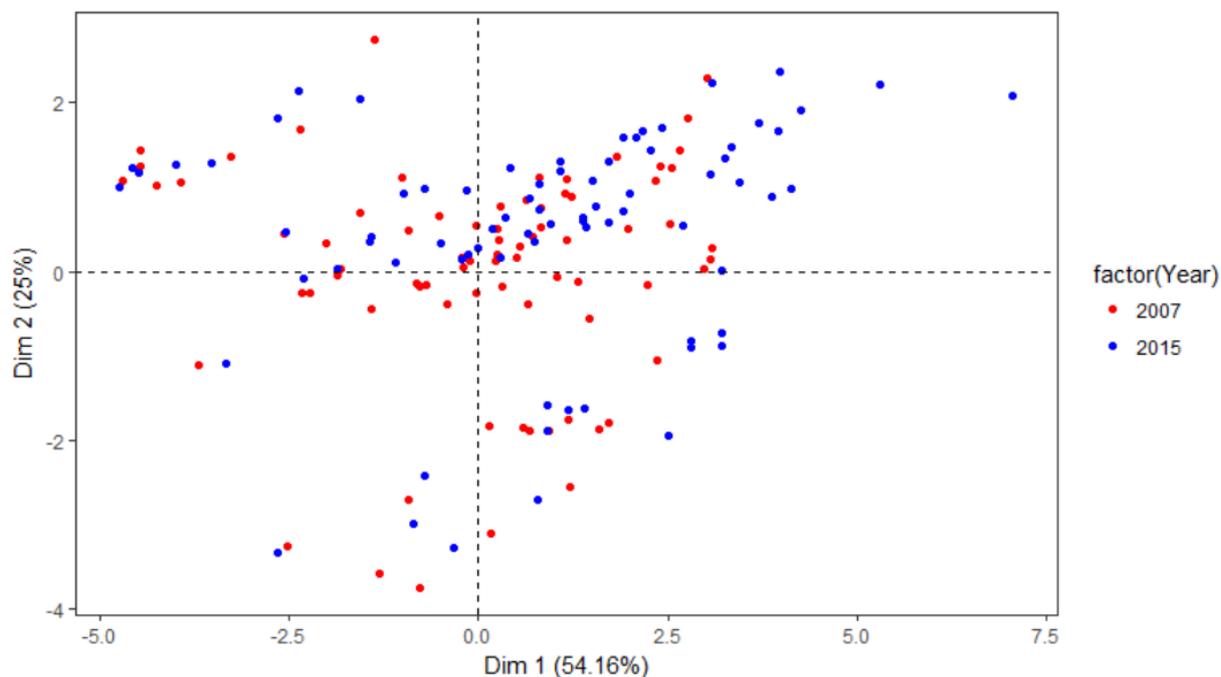


Figure 7: Projection of 2015 universities as supplementary individuals on the 2007 PCA

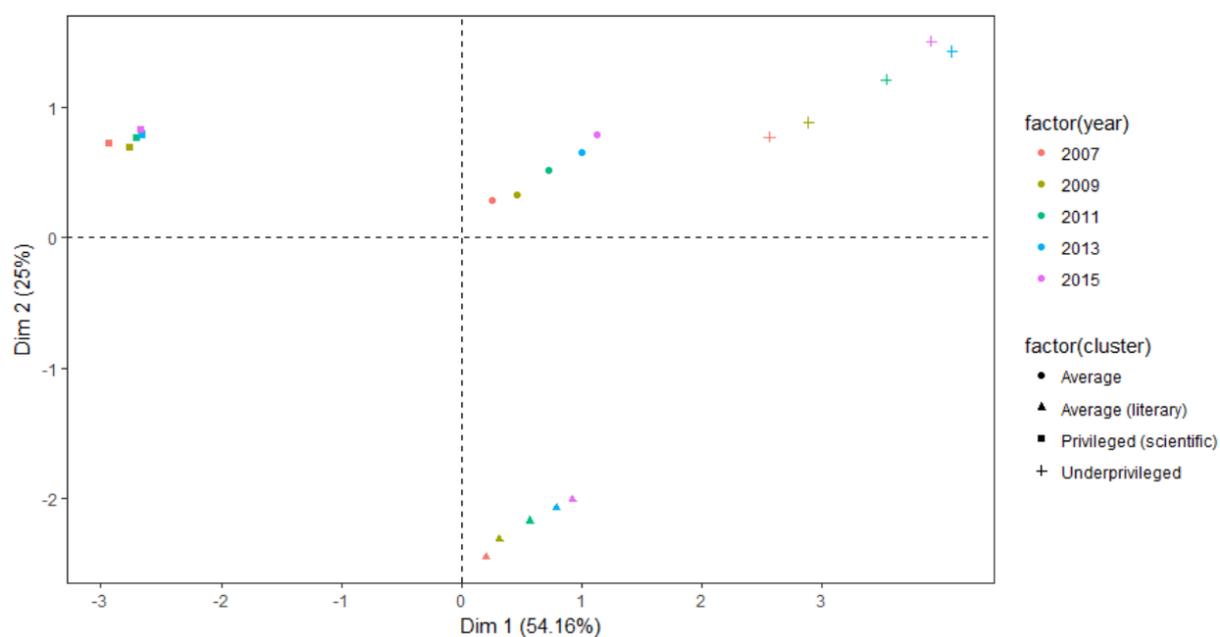


Figure 8: Evolution of the four clusters barycentres between 2007 and 2015

We confirm the legitimacy of the previous conjectures by computing the GINI coefficient as well as the ER polarization measure over the normalized coordinates of universities on the first component of the PCA. The resulting figures are presented in Table 2 for the 3-clusters configuration (3CC), in Table 3 for the 4-clusters configuration (4CC) and in table Table 4 for the modified 5-clusters configuration (M5CC). Whatever the number of clusters retained and the value of the α parameter – i.e. the degree of sensitivity to polarization – the results indicate a trend of increasing polarization among French universities, albeit at a lower rate for higher values of α . If we analyze the evolutions in more detail, it appears that social polarization increases at a high rate between 2007 and 2012, and then increases at a slow pace or stabilizes, depending on the configuration. The MCKENZIE measure of inequality exhibits a very similar trend over time, which supports the validity of the previous results.

Table 2: Evolution of GINI and ER polarization measures between 2007 and 2015 in the 3 clusters configuration (3CC)

Year	GINI	ER ($\alpha = 0.25$)	ER ($\alpha = 0.5$)	ER ($\alpha = 0.75$)	ER ($\alpha = 1$)	MCKENZIE
2007	0.149	0.115	0.090	0.071	0.056	1
2008	0.143	0.111	0.086	0.068	0.054	0.976
2009	0.158	0.122	0.095	0.075	0.060	1.035
2010	0.166	0.128	0.100	0.079	0.063	1.073
2011	0.177	0.136	0.106	0.084	0.067	1.147
2012	0.193	0.149	0.117	0.092	0.073	1.231
2013	0.195	0.151	0.118	0.093	0.074	1.239
2014	0.201	0.155	0.121	0.096	0.076	1.263
2015	0.201	0.155	0.121	0.095	0.076	1.242

Table 3: Evolution of GINI and ER polarization measures between 2007 and 2015 in the 4 clusters configuration (4CC)

Year	GINI	ER ($\alpha = 0.25$)	ER ($\alpha = 0.5$)	ER ($\alpha = 0.75$)	ER ($\alpha = 1$)	MCKENZIE
2007	0.224	0.158	0.112	0.080	0.058	1
2008	0.211	0.148	0.105	0.076	0.055	0.976
2009	0.236	0.166	0.118	0.084	0.061	1.035
2010	0.250	0.175	0.124	0.089	0.065	1.073
2011	0.279	0.196	0.139	0.100	0.072	1.147
2012	0.313	0.220	0.156	0.112	0.081	1.231
2013	0.319	0.224	0.159	0.114	0.083	1.239
2014	0.315	0.221	0.157	0.113	0.082	1.263
2015	0.316	0.222	0.157	0.113	0.082	1.242

Table 4: Evolution of GINI and ER polarization measures between 2007 and 2015 in the modified 5 clusters configuration (M5CC)

Year	GINI	ER ($\alpha = 0.25$)	ER ($\alpha = 0.5$)	ER ($\alpha = 0.75$)	ER ($\alpha = 1$)	MCKENZIE
2007	0.242	0.167	0.117	0.083	0.060	1
2008	0.228	0.157	0.110	0.079	0.057	0.976
2009	0.256	0.176	0.124	0.088	0.064	1.035
2010	0.272	0.187	0.131	0.094	0.068	1.073
2011	0.304	0.210	0.147	0.105	0.076	1.147
2012	0.341	0.235	0.165	0.117	0.085	1.231
2013	0.348	0.240	0.168	0.120	0.087	1.239
2014	0.346	0.238	0.167	0.119	0.086	1.263
2015	0.346	0.238	0.167	0.119	0.086	1.242

5 Discussion

Our purpose was to investigate whether a Matthew effect among French universities as a result of the increasing use of rankings to compare higher education institutions. This master's thesis thus forms part of a body of research which points out the potential for increased social inequalities among universities because of the side-effects of rankings (ALTBACH, 2004 ; VAN PARIJS, 2009 ; HALFFMAN & LEYDESDORFF, 2010).

First, as we wanted to measure and monitor social polarization between universities, we sought to find if there existed clearly defined groups of universities in terms of social composition and main field of study. The resulting typology reveals a highly unequal university landscape in terms of social composition. These findings corroborate and complement accordingly those of BRUSADELLI & LEBARON (2012). In particular, our PCA confirms their prediction according to which would eventually emerge a two tier education system: "on one side, mass universities counting many undergraduate diplomas, and on the other, elite universities, predominantly scientific, at the heart of the reproduction strategies of students coming from high social backgrounds" (see Figures 1 and 2).

Then, we superimposed upon these results the list of French universities appearing in main rankings, in order to determine whether they really measure academic performance as they are supposed to. BILLAUT, BOUYSSOU & VINCKE had already proved in a general manner that because of methodological flaws, rankings ended up measuring entirely different realities than those. We confirm this diagnosis in the case of French universities. More precisely, we find that rankings actually measure social composition, specificity of the main field of study and size of universities rather than merely academic performance.

Finally, after explaining why and how we could use coordinates of universities on the first PC as a social composition index, we used them to compute inequality and polarization measures and tracked their evolution over time. To our knowledge, this is the first study to apply the polarization framework to social composition of universities. We conclude that social polarization between French universities unambiguously increases between 2007 and 2013, irrespective of the tool used to measure it, and then tends to stabilize in more recent years. We thus contradict the results of HALFFMAN & LEYDESDORFF who find no existence of a Matthew effect by tracking over time GINI coefficients computed on ARWU scores.

An important shortcoming of this paper is that the social index we built thanks to the PCA doesn't encompass any information on public scholarships, whereas those are substantial in France and give valuable information on social origin of students. The lack of a variable indicating for each student whether he receives a public scholarship and at which level is obviously due to their sensible nature. We managed to obtain these data by another source, but only between 2007 and 2011. Thus, we didn't encompass them in the analysis as it would have caused a break in series. However, when we use this variable to build the PCA, we find that it is actually positively highly correlated with the fact of being a children of a working class father, so that it doesn't modify the social index to a great extent. Thus, when we perform the analysis between 2007 and 2011 adding the share of students owning a scholarship by university, we find a very similar trend of increasing social polarization. Nonetheless, it would be interesting to have information on public scholarship for subsequent years, in order to determine whether it has an impact on the stabilization of polarization we observe.

Another lead for further research will be to simply extend the study as more recent data become available, in order to determine if the stabilization of social polarization we observe in recent years is confirmed or disproved. Similarly, we could expand the sample of the study by

encompassing engineering schools, business schools, Grandes Écoles, etc. to find out whether there are similar signs of polarization at a more aggregated level. These two leads would likely make the results at hand more robust.

It would also be interesting to test the Matthew effect hypothesis using the same tools and concepts in different education systems. However, the French education system is very particular because of its mainly public nature, thus the analysis and the conclusions would have to be adapted to the particular characteristics of the chosen system. Furthermore, it is possible that the same data would not be available, or not as extensively, which would impair the comparability of the results.

Finally, although we believe that our results support the core hypothesis of this master's thesis, it is possible that the growing social polarization we observe among French universities is actually caused by other phenomena than performative effects of rankings. For instance, we have seen that the inequalities regarding social composition we observed between universities correspond to a large extent to the social inequalities that exist between regions or department. We can thus imagine that what we observe is the consequence of a similarly growing social polarization between regions. Thus, it would be interesting to design a study which would couple our analysis with a spatial segregation study – as done for instance by TRANCART (2012) for secondary schools using segregation measures. As a result, we may be able to distinguish between social inequalities attributable to spatial inequalities and "pure" social inequalities attributable to universities strategies – e.g. offering excellence tracks to attract better students, implementing tuition fees, etc.

6 Conclusion

Over the last decade, university rankings have been increasingly used to evaluate and compare performances of higher education institutions. A simple explanation of this trend is that "university rankings convey the fascination of numbers despite the ambiguity of what is measured" (HALFFMAN & LEYDESDORFF, 2010).

Our results corroborate this statement in the case of French universities: rankings are advertised as measuring academic performance and/or quality of teaching whereas they actually reflect social composition, field of study and size of universities. But what makes the main rankings truly problematic is their often overlooked performative effect. They generate strong incentives for policy makers to align education systems with the model of global research universities. And since it is not suitable for a country to possess more than several world-class universities, it is likely that these universities will get the largest share of resources at the expense of others. The trend of increasing social polarization of French universities we observe between 2007 and 2015 confirms this hypothesis.

As BILLAUT, BOUYSSOU and VINCKE perfectly put it: "an optimistic point of view would be that, after having read our paper, the authors of the ranking would decide to immediately stop their work, apologizing for having created so much confusion in the academic world, and that all political decision makers would immediately stop using "well know international rankings" as means to promote their own strategic objectives" (op. cit., 2010). Given their annual media coverage however, it is unlikely that the production of poorly designed rankings is going to end anytime soon. And although most education policy makers seem to acknowledge the methodological flaws of the main rankings, no one wants to be blamed for having deteriorated the positions of French universities in these.

Against that background, our findings strongly advocate against basing entire education policies on the opportunity of moving up in the main rankings. Instead, they should be coupled with more meaningful indicators. For instance, our social composition index could be used

as a tool to rebalance the funding of universities by giving more weight to universities which exhibit the poorest social composition, and conversely. This approach would have the merit of concentrating funds where they are the most needed, and thus counterbalance the perverse allocating effects of rankings.

References

- Valérie Albouy and Chloé Tavan. Accès à l'enseignement supérieur en france: une démocratisation réelle mais de faible ampleur. *Économie et statistique*, 410(1):3–22, 2007.
- Valérie Albouy and Thomas Wanecq. Les inégalités sociales d'accès aux grandes écoles suivi d'un commentaire de louis-andré vallet. *Économie et statistique*, 361(1):27–52, 2003.
- Philip G Altbach. The costs and benefits of world-class universities. *Academe*, 90(1):20, 2004.
- Gordon Anderson. Polarization measurement and inference in many dimensions when subgroups can not be identified. *Browser Download This Paper*, 2011.
- Haroon Bhorat and Carlene van der Westhuizen. Non-monetary dimensions of well-being in south africa, 1993–2004: A post-apartheid dividend? *Development Southern Africa*, 30(3):295–314, 2013.
- Jean-Charles Billaut, Denis Bouyssou, and Philippe Vincke. Should you believe in the shanghai ranking? *Scientometrics*, 84(1):237–263, 2010.
- Pierre Bourdieu and Jean-Claude Passeron. *Les héritiers: les étudiants et la culture*. Minuit, 2016.
- Isabelle Bruno, Sophie Jacquot, and Lou Mandin. Europeanization through its instrumentation: benchmarking, mainstreaming and the open method of co-ordination toolbox or pandora's box? *Journal of European public policy*, 13(4):519–536, 2006.
- Nicolas Brusadelli and Frédéric Lebaron. Les indicateurs de «performance» universitaire. *Savoir/Agir*, (4):97–104, 2012.
- Alain Desrosières. L'argument statistique. i: Pour une sociologie historique de la quantification. ii: Gouverner par les nombres. *L'argument statistique II*, 2008.
- Domingo Docampo, Daniel Egret, and Lawrence Cram. The effect of university mergers on the shanghai ranking. *Scientometrics*, 104(1):175–191, 2015.
- Jean-Yves Duclos, Joan Esteban, and Debraj Ray. Polarization: concepts, measurement, estimation. *Econometrica*, 72(6):1737–1772, 2004.
- Marie Duru-Bellat and Annick Kieffer. Du baccalauréat à l'enseignement supérieur en france: déplacement et recomposition des inégalités. *Population*, 63(1):123–157, 2008.
- Joan-Maria Esteban and Debraj Ray. On the measurement of polarization. *Econometrica: Journal of the Econometric Society*, pages 819–851, 1994.
- Deon Filmer and Lant H Pritchett. Estimating wealth effects without expenditure data or tears: an application to educational enrollments in states of india. *Demography*, 38(1):115–132, 2001.
- Davidson R Gwatkin, Shea Rustein, Kiersten Johnson, Eldaw Suliman, Adam Wagstaff, and Agbessi Amouzou. Socio-economic differences in brazil. *Washington, DC: HNP/Poverty Thematic Group of the World Bank*, 2000.
- Willem Halfman and Loet Leydesdorff. Is inequality among universities increasing? gini coefficients and the elusive rise of elite universities. *Minerva*, 48(1):55–72, 2010.

- David J McKenzie. Measuring inequality with asset indicators. *Journal of Population Economics*, 18(2):229–260, 2005.
- Robert K Merton et al. The matthew effect in science. *Science*, 159(3810):56–63, 1968.
- Sandrine Nicourd, Olivia Samuel, and Sylvie Vilter. Les inégalités territoriales à l'université: effets sur les parcours des étudiants d'origine populaire. *Revue française de pédagogie. Recherches en éducation*, (176):27–40, 2011.
- Camille Peugny. *Le destin au berceau. Inégalités et reproduction sociale*. 2013.
- Theodore M Porter. *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press, 1996.
- Fabienne Rosenwald. Filles et garçons dans le système éducatif depuis vingt ans. *Données sociales*, 2006.
- Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65, 1987.
- David E Sahn and David C Stifel. Poverty comparisons over time and across countries in africa. *World development*, 28(12):2123–2155, 2000.
- Robert Tibshirani, Guenther Walther, and Trevor Hastie. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):411–423, 2001.
- Danièle Trancart. The impact of social segregation on academic achievement. *Formation emploi*, (4):35–55, 2012.
- Philippe Van Parijs. European higher education under the spell of university rankings. *Ethical perspectives*, 16(2):189–206, 2009.
- Seema Vyas and Lilani Kumaranayake. Constructing socio-economic status indices: how to use principal components analysis. *Health policy and planning*, 21(6):459–468, 2006.
- You-Qiang Wang and Kai-Yuen Tsui. Polarization orderings and new classes of polarization indices. *Journal of Public Economic Theory*, 2(3):349–363, 2000.

Appendices

Appendix A

Table 5: Description of active and supplementary variables used in the 2007 PCA

Variable	Description	N	Mean	St. Dev.
hommes	Pct of male students	77	0.455	0.106
femmes	Pct of female students	77	0.545	0.106
bac_S	Pct of students with a scientific <i>baccalauréat</i>	77	0.388	0.168
bac_ES	Pct of students with an economic <i>baccalauréat</i>	77	0.193	0.080
bac_L	Pct of students with a literary <i>baccalauréat</i>	77	0.143	0.096
bac_pro	Pct of students with a professional <i>baccalauréat</i>	77	0.015	0.011
bac_techno	Pct of students with a technological <i>baccalauréat</i>	77	0.116	0.051
ouvriers	Pct of students whose parents are workers (SPCs 61-69)	77	0.098	0.042
pcs_plus	Pct of students whose parents belong to high SPCs (31-38)	77	0.301	0.095
licences	Pct of undergraduate students	77	0.603	0.202
masters	Pct of graduate students	77	0.349	0.184
doctorats	Pct of PhD students	77	0.047	0.029

Appendix B

Table 6: Percentages of explained variance by principal components of the 2007 PCA

PC	eigenvalue	percentage of variance	cumulative percentage of variance
1	3.791	54.158	54.158
2	1.750	24.999	79.157
3	0.667	9.527	88.684
4	0.365	5.208	93.892
5	0.219	3.130	97.022
6	0.155	2.214	99.236
7	0.054	0.764	100

Table 7: Mean of active and supplementary variables according to HCA 4 clusters in 2007

cluster	hommes	femmes	bac_S	bac_ES	bac_L	bac_pro
Underprivileged	0.431	0.569	0.278	0.224	0.161	0.033
Average	0.458	0.542	0.388	0.198	0.129	0.014
Average (litterary/social)	0.344	0.656	0.187	0.269	0.284	0.012
Privileged	0.560	0.440	0.646	0.091	0.041	0.004

Table 6 (continued): Mean of active and supplementary variables according to HCA 4 clusters in 2007

cluster	bac_techno	ouvriers	pcs_plus	licences	masters	doctorats
Underprivileged	0.183	0.142	0.203	0.785	0.194	0.021
Average	0.126	0.111	0.275	0.656	0.304	0.040
Average (litterary/social)	0.077	0.067	0.348	0.583	0.353	0.064
Privileged	0.071	0.055	0.403	0.349	0.579	0.072

Appendix C

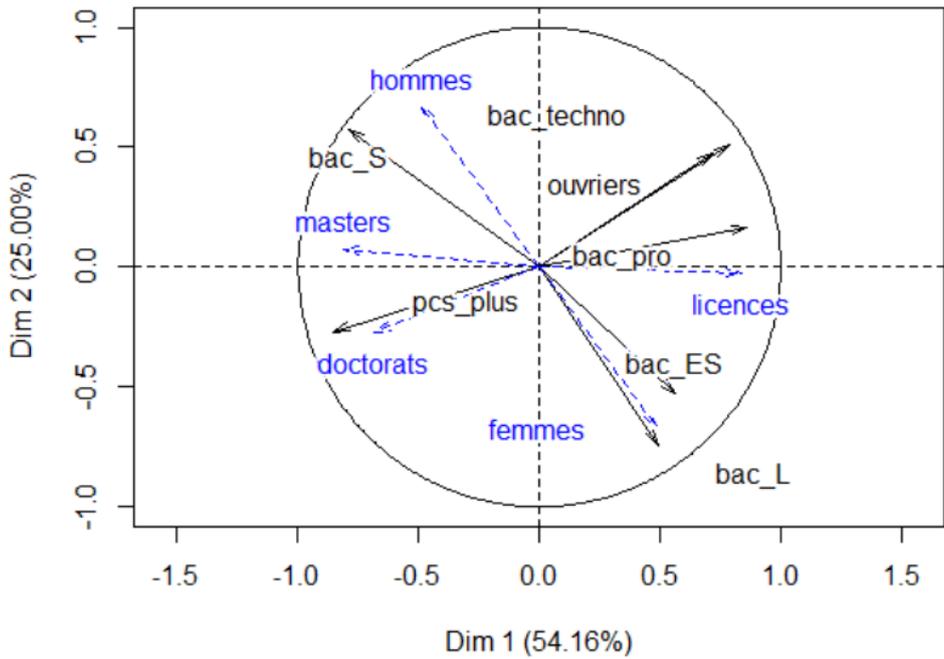


Figure 9: Graph of 2007 PCA active (black) and supplementary (blue) variables on dimensions 1 (horizontal) and 2 (vertical)

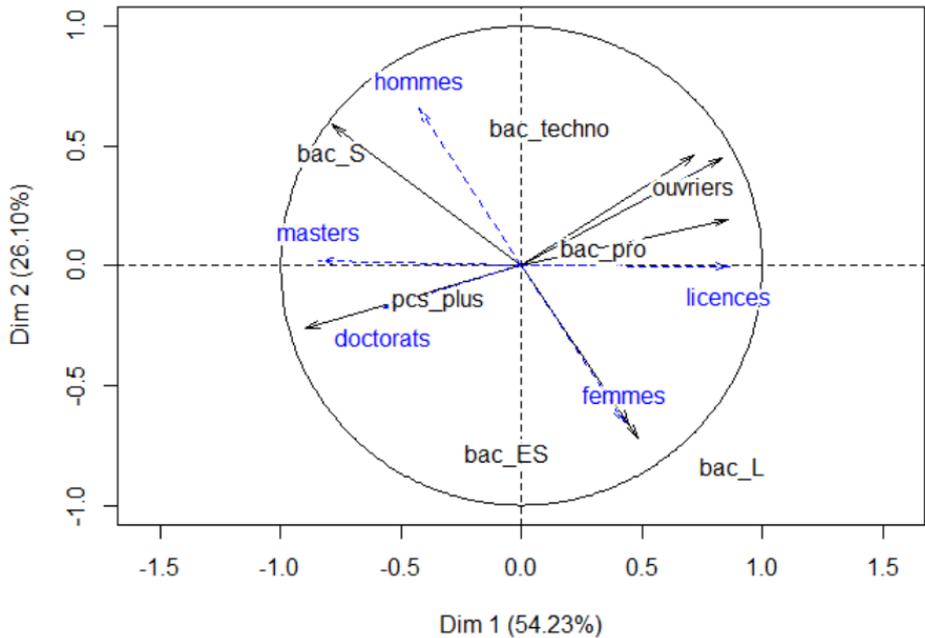


Figure 10: Graph of 2015 PCA active (black) and supplementary (blue) variables on dimensions 1 (horizontal) and 2 (vertical)