ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 78 – 2005

Hiring discrimination in the French financial sector: an econometric analysis on field experiment data

Emmanuel DUGUET* and Pascale PETIT**

ABSTRACT. – We study the determinants of hiring gender discrimination in the French financial sector through a controlled experiment. We find that, on the one hand, the access differences to job interviews by women and men are primarily explained by the expectation of a maternity by young women and, on the other hand, we also find that some institutional mechanisms compensate this difference of treatment between genders so that there is no significant discrimination on average.

La discrimination à l'embauche dans le secteur financier français : une étude économétrique sur données d'expériences contrôlées

RÉSUMÉ. – Nous étudions les déterminants de la discrimination à l'embauche dans le secteur financier français, à partir d'une expérience contrôlée. Nous trouvons que, d'une part, les différences d'accès aux entretiens d'embauche par les femmes et les hommes s'expliquent principalement par les anticipations de naissances des jeunes femmes et que, d'autre part, certaines caractéristiques institutionnelles permettent de compenser cette différence de traitement, de sorte qu'il n'existe plus de discrimination en moyenne.

We thank Pierre Cahuc, Jean-Marc Robin, the editor and two anonymous referees for their helpful suggestions and comments. We also thank the members of EUREQua that have participated to the correspondence test, the participants of the EUREQua Labor Market Seminar, the EPEE Research Seminar (Evry), the Journées Doctorales de l'ADRES and the T2M Conference (Orléans). The authors remain responsible of any error or omission.
* EPEE (Equipe d'Accueil 2177) – Université d'Evry Val d'Essonne – 4 boulevard François Mitterand – 91025 EVRY Cedex. Email : emmanuel.duguet@univ-evry.fr
** P. PETIT : DARES and EUREQua, UMR CNRS 8594, University of Paris I Panthéon-Sorbonne ; email : p.petit@dares.travail.gouv.fr

Introduction

The existence of discrimination against some demographic groups is a growing concern in developed countries. According to HECKMAN [1998], discrimination is said to arise "if an otherwise identical person is treated differently by virtue of that person's race or gender, and race or gender by themselves have no direct effect on productivity".

CAIN [1986] distinguishes two major sources of discrimination. The first, advanced by BECKER [1957] states that all workers have the same skills but that the majority demographic group (workers, employers and customers) are prejudiced against minority demographic groups, so that the majority group minimizes contact with others (taste discrimination). The second is advanced by ARROW [1972a, 1972b and 1973], MCCALL [1972] and PHELPS [1972]. Employers are not able to evaluate perfectly the productivity of each applicant for a job. Hence, in the hiring process, employers combine direct evaluations of candidates (qualifications, experience, tests...), with beliefs concerning the average productivity of their demographic group. In this case statistical discrimination applies: two applicants for a job may be treated differently if they belong to different demographic groups, even though they have identical productive characteristics. This kind of discrimination can produce a self-fulfilling anticipation that contributes to maintain discrimination on the labor market (LOMMERUD and VAGSTAD [2000]).

Since the seminal contribution by BECKER [1957], many empirical studies have attempted to estimate the gender wage discrimination (KUNZE [2000]). A first method, widespread in the literature, is the BLINDER-OAXACA decomposition (BLINDER [1973], OAXACA [1973]). The gender wage gap is decomposed in two parts: first, the differences in productivity related endowments and, second, the unexplained part of the gap, a residual commonly interpreted as a measure of discrimination. A second method estimates discrimination by the difference between the wage gap and the estimated productivity gap. In France, this difference would be, for the whole economy in 1997, around 4% (CRÉPON, DENIAU and PÉREZ-DUARTE [2001]); MEURS and PONTHIEUX [2000]).[1] In the banking sector, this difference would reach 15% (FAKHFAKH, MERLATEAU and MEURS [2002]).

Even though they are interesting for their own sake, these estimations of the wage gap should be complemented by an evaluation of the other types of discrimination, mainly for two reasons. First, in the countries that have laws against discrimination, the discriminatory practices should move toward practices that are less easily detectable than wage discrimination. For instance, discriminatory practices could appear in the hiring process. Second, the reduction of the wage gap increases the labor cost of women compared to men and could therefore decrease the probability that women get jobs.

This paper studies discrimination in the access to hiring interviews. In order to evaluate this type of discrimination, one needs new data sources. Indeed, there is no general data source that would provide information on each application to job offers, including productive and non-productive characteristics. The data must be

1. The gap is around 15% on American data (BLAU and KAHN [1994]; HELLERSTEIN, NEUMARK and TROSKE [1999]).

collected by the researchers themselves. This constraint, however, offers the possibility to adapt the data collection process to the evaluation of discrimination.

The collection method that appeared in the literature is correspondence testing. Its principle is to conduct a controlled experiment in the following way. One builds similar pairs of applications for both genders and sends them to the same job offers. When carefully done, this method should reveal the differences of treatment that are explained by the gender (for a survey, see PETIT [2003]). The access to hiring interviews provides *a priori* an imperfect indicator of job access. Sending the candidates to the interviews would permit to measure their access to employment. However, several arguments can be advocated to justify the limitation of the experiment to correspondence testing (RIACH and RICH [1991]). First, in the case of the financial sector, the recruiting process is costly for firms. Each interview generally requires a recruiter during a significant time period in order to organize discussions, ability tests etc. Therefore, we expect that a candidate receives an interview when the recruiter considers that (s)he has a significant chance of getting the job. Second, using correspondence testing, we are able to exercise an accurate control over the content of applications to be sure that all relevant characteristics other than gender are carefully matched (RIACH and RICH [1991]). More precisely, we are sure that the physical appearance and personality of candidates have not been a hiring criterion and have not affected the outcome. These arguments are supported by several empirical studies: NEUMARK *et al.* [1996] find that the gender gap is 39% for job access and 35% for the access to interviews in favor of men in high-price restaurants; KENNEY and WISSOKER [1994] find a gap of 15% for both the access to interviews and to jobs against the Hispanic candidates. Third, the data collection procedure is less costly, so that it is possible to collect more data.

Correspondence testing has sometimes been criticized. First, following HECKMAN [1998], it is possible that a given productive characteristics does not send the same signal depending on it belongs to a woman or a man. One can extend this argument to the characteristics of the correspondence test itself. A same curriculum vitae, for instance, can send a different signal depending on the gender of the applicant. In this situation, controlling for the content of the applications does not necessarily leads to a proper measure of discrimination. A way to account for this critic is to use regression methods that control for both the productive and non-productive characteristics of the test, instead of comparing the percentages of success directly. This paper extends the practice of the previous literature by controlling for all the characteristics available: application, type of job, type of firm and of the test (date of application, CV type).

A second way to improve on the correspondence test deals with the statistical methods used. The normality of the estimators is used systematically, while it is guaranteed on large samples only. Moreover, this assumption is not tested when it is used. Here, the problem comes from the fact that the correspondence tests are performed on a small number of applications, because of the high collection cost of the data. In this paper, we recommend to use the bootstrap method.

This paper uses the first correspondence test conducted in France (PETIT [2004]). The financial sector has been chosen because it includes a large proportion of women in its labor force (53%, AFB ENQUÊTE EMPLOI [1999]). Even though this large proportion could suggest that there is no gender discrimination in this sector, it remains possible that there is a composition effect, according to which the women would be allocated to the less paid jobs. Notably, 63% of the employees are women while they represent only 31% of the executives. We wish to test this assumption by collecting data of both the jobs requiring low and high qualifications. However, we have not been able to collect data on the executives' jobs for the two following reasons. First, there are few job offers at this level of qualification so that the sample size would have been too small; second, the recruiters may be more careful about the applications, especially by checking the references indicated in the CVs. Therefore we had no other choice than to replace the comparison between employees and executives by a comparison between low-skilled and high-skilled employees. Another reason to study the financial sector is that its collective agreement is especially favorable to women. Compared to the common rule, women have the right to 45 additional days of maternity vacations with full pay that is incurred by the employer. Women can replace it by 90 days of maternity vacations with a half-pay.

We obtained two main findings. First, the probability of a maternity and the stronger involvement of women in children's education would explain their lower access to job interviews. Second, some types of firms appear to favor women against men such that they fully compensate the former discriminatory effect. Overall, there is no significant discrimination.

In the first section, we provide some theoretical motivations for the study and, in the second section, we present the data collection methodology that aims to eliminate the selection biases. The econometric model and its estimation are presented in the third section. The results are discussed in section 4.

1 Model

We consider a simple situation where the total cost of a male worker for the employer is equal to:

$$C_m = w_m + p\ F_m$$

where w_m is the annual labor cost, p the probability of a paternity and F_m the cost of a paternity vacation for the bank. In France, since January 1st 2002, a man can take up to 14 days of paternity vacation with pay, whatever his status (Law of December 4th 2001). The duration of the vacation is extended to 21 days for a multiple birth. The cost of the vacation is paid by the Sécurité Sociale so that the firms only incur an organizational cost. The annual labor cost of a female worker is denoted C_f. The total cost of a female worker for an employer is therefore:

$$C_f = w_f + p\ F_f + M,$$

where w_f is the annual labor cost, p the probability of a maternity vacation and F_f the maternity cost with $F_f > F_m$. In France, women have the right to 112 days of maternity vacations with full pay that are mainly incurred by the Sécurité Sociale. In the financial sector, women have the right to 45 additional days of maternity vacations with full pay that is incurred by the employer. Women can replace it by 90 days of maternity vacations with a half-pay. The total cost therefore includes both the labor cost and the replacement cost of the employee on vacation. The quantity M represents the balance of two effects. The first effect is positive and represents the money equivalent to the reluctance of the employer to hire women, following BECKER [1957]. The second effect is negative and represents the amount of money that the firm is ready to pay to reduce discrimination. Here, this amount is represented by a reduction of the perceived annual labor cost of a woman. This amount is related to the discrimination legislation, to the image of the firm or to the existence of a positive discrimination policy inside the firm.

In the presence of hiring discrimination, we should observe a wage gap between women and men that is not correlated to the productive characteristics of the individuals; we denote it $\delta = w_m - w_f > 0$. The perceived cost difference between women and men is therefore equal to:

$$y^* = C_f - C_m = p\left(F_f - F_m\right) + M - \delta.$$

We will observe a preferential hiring of men when the probability of maternity crosses the following threshold $\overline{p}$:

$$y^* > 0 \Leftrightarrow p > \overline{p} = \frac{\delta - M}{F_f - F_m}.$$

The quantity $\delta/(F_f - F_m)$ represents the benefit to cost ratio of women hiring. Indeed, if there is wage discrimination, it is less costly for the employer to hire a woman. Therefore, we reach the conclusion that wage discrimination should reduce hiring discrimination. This benefit from hiring women should be balanced against the higher maternity cost $F_f - F_m$ and the discrimination component strictly speaking $M/(F_f - F_m)$.

The previous relationship implies three properties. First, the reduction of wage discrimination should increase hiring discrimination. Second, the hiring probability of a woman should decrease with her probability of maternity. Third, since some women value more their professional life than their domestic life, the banks have an interest to implement screening designs so as to attract the women that have the lowest probability of maternity. These designs can take the form of restrictions in the labor contract they offer.

Several type of screening restrictions can be considered. The first one is the duration of the labor contract; a CDD (contrat à durée déterminée) is a fixed-term labor contract and a CDI (contrat à durée indéterminée) is a labor contract without a specified duration. In case of maternity, a CDD is not extended for the duration of the maternity vacation and the employer has the possibility to end the contract at its term, even if it occurs during the maternity vacation. Other screening restrictions are possible: the jobs with a wage that depends on the output or the jobs that clearly indicate possibilities of promotion, aim to attract the candidates that intend to devote themselves to their professional career. These screening measures can be interpreted as a reduction of the parameter M.

Globally, the elements of the labor contract can result from an initial situation of statistical discrimination, such that the firms have an interest to modify the content of the labor contracts in order to attract the women that put their professional career first.

Table 1 describes the relationships between the parameters of the model and the data available in this study. From this simple theoretical representation, the decisions that we can observe can be put under the following form:

$$y_i^* = p_i \left(F_{f,i} - F_{m,i}\right) + M_i - \delta_i, \quad i = 1,...N$$

where N is the number of correspondence tests. The parameters can depend on each application, in a way that is described in Table 1. The discrimination variable that we observe is equal to:

$$D_i = \begin{cases} 1 & if \quad y_i^* > 0 \\ 0 & if \quad y_i^* = 0 \\ -1 & if \quad y_i^* < 0 \end{cases}$$

The average of this variable over all the correspondence tests $1/N\sum_{i=1}^{N} D_i$ simply gives the difference between the success percentages of men and women that have the same productive and non-productive characteristics and that have applied for the same jobs.

In our application, we represent the parameters of the model by a set of explanative variables, denoted X, which is presented in Table 1.

2 Implementation of the correspondence tests

We have decided to compare the access to hiring interviews between several types of pairs in order to better isolate the part of the gap that can be attributed to hiring discrimination.

2.1 Three types of applications

Three types of applications have been built: aged 25, single, childless; aged 37, single, childless; and aged 37, married with three children. The difference between the two first types of applications should reveal the difference of treatment linked to the age of the candidate. The difference between the two last types of applications should reveal the difference of treatment linked to the family constraints. The characteristics of both genders are identical.

TABLE 1
Relationship between the parameters of the model and the data available

Parameter	Variable	Comment
P	Age (25 or 37)	The women of 37 years old have a lower fecundity rate than the women of 25 years old.
	Number of children (none or three)	The women without children have a lower fecundity than the women of the same age with three children.
F	Age (25 or 37)	The women aged 37 that have a job and never known unemployment have a professional experience that is stronger than the one of the women aged 25 with a comparable profile. Their labor cost is therefore higher.
	Qualification (Baccalauréat or BTS, see Table 2)	Women that hold a baccalauréat degree have a lower labor cost than the women with a BTS degree.
	Training paid by the employer	If the firm pays for training, it will tend to hire employees that are present in the firm for a period that is long enough to recoup the training cost.
M	Type of financial establishment	Some establishments have explicit policies against discrimination (like La Poste). This reduces *M*.
	Type of labor contract	Short-term labor contracts offer more opportunities to stop the labor relationship. They have a lower *M* parameter than the long-run labor contracts.
	Possibilities of promotion	This should attract the candidates that value their professional career. This should reduce *M*.
	Wage increasing with the output	This should attract the candidates that value their professional career. This should reduce *M*.
	Negotiable wage	This should attract the candidates that value their professional career. This should reduce *M*.

N.B.: The correspondence tests only concern people that already have a job and do not record any period of unemployment. The data used in these study are experimental, which explains why the available data takes a limited set of values.

These three types of applications send three different signals to the employers. The family status of the candidates aged 25 is expected to change since, in France, the fertility rate of the women aged 25-29 is 130% while the fertility rate of women aged 35-39 is 50%.[2] Young women, in our experiment, are therefore expected to have children. The choice of the age of the candidates is also important in the experiment design. The choice of 37 years old results from a trade-off. On the one hand, it should send a signal of the stability of the family situation; the older a woman

2. On average, the first maternity occurs when women are 29 years old.

is, the lower her probability of a maternity. On the other hand, on the French labor market, when the candidates get older they have smaller chances to get a job.[3]

Overall, the probability of a maternity is the highest for the women aged 25, decreases for the childless women aged 37 and reaches its minimum for the women aged 37 with three children.[4]

Banking establishments offer four types of jobs: lowly qualified administrative jobs, highly qualified administrative jobs, lowly qualified commercial jobs and highly qualified commercial jobs. This is summarized in Table 2. The three types of applicants have been alternatively endowed with the four previous qualifications. Globally, 24 applications have been built.

TABLE 2
Structure of the correspondence test

Qualifications	Commercial jobs	Administrative jobs
Baccalauréat	Lowly qualified commercial jobs:	Lowly qualified administrative jobs:
School-leaving examination granting the right to enter university (A-level)	Receptionist Counter clerk	Administrative technician Administrative clerk Accounting clerk
Brevet de Technicien Supérieur (BTS)	Highly qualified commercial jobs:	Highly qualified administrative jobs:
Vocational training certificate, given after 2 years of post-baccalauréat education	Establishment director Customer consultant Sales manager Bank customer assistant	Executive manager Portfolio manager Recovery manager Accounting manager

2.2 Design of the correspondence test

For each qualification, the three types of applicants have been endowed with a significant experience in the banking sector and without unemployment period. In order to avoid detection, the applications were not strictly identical, but the candidates have been carefully matched in all essential personal and experience characteristics. Therefore, the only effective distinctive characteristic within each pair of candidates is the gender. Thus for each job advertisement, applicants have been endowed with similar qualification, training, past work experience, personal history, hobbies, driving license, typical French first and family names. Addresses have been assigned to residential areas located in Paris at the same distance from the potential workplace and with comparable social and economic characteristics. The applicants' residencies signal to employers the daily travel time and social environment.

3. The probability to find a job decreases after 40 years old.
4. On average, women have two children, so that women with three children will be expected to have completed their maternity projects.

2.3 Implementation of the correspondence test

In order to control for the possibility that the style or the content of a particular application might influence employers' responses, the curriculum vitae were rotated among each pair of candidates. Moreover, the applications did not include any photography. Therefore we are sure that the physical appearance and personality of candidates have not been a hiring criterion and have not affected the outcome. Since the data collection procedure is less costly, we have been able constitute a larger sample: 942 applications have been sent (*i.e.,* 471 for each gender).

2.4 Job advertisement sources

In France, the public organism that centralizes most of clerk and management job offers in the service sector is the ANPE (Agence Nationale pour l'Emploi). Therefore we have used the ANPE job offer database, which is updated daily. In order to get a sample that is also representative of other advertisement sources we have also used private websites specialized in job offers and the firms' own web pages devoted to recruitment (on a daily basis). In addition, we have also sent unsolicited applications. This variety of sources is likely to reflect what a real candidate would do in order to find a job.

2.5 Applications mailing

The job applications have been sent between January and March 2002.

In order to limit the consequences of the monopolistic characteristic of the labor demand in banking, we have chosen to extend the study to several lines of business in the financial sector: commercial banks, trust banks, mutualist banks, savings bank, factoring bank, credit organisms and post office banking services. Secondly, we have chosen to send the applications at the establishment level when they were endowed with their own human resources service, rather than at the level of firms.

TABLE 3
Number of correspondence tests

Number of job offers*	157
Including:	
Low qualified commercial jobs	40
Low qualified administrative jobs	24
Highly qualified commercial jobs	58
Highly qualified administrative jobs	35
Number of establishments	75
Number of firms	59
Number of observations	942

* *For each gender and type of correspondance test, which makes 2x3x157=942 observations.*

We have focused on jobs localized near Paris and on fixed-term and permanent contracts (this definition excludes interim). Some establishments have been used several times when they have offered jobs of different profiles that corresponded to the ones of the experiment. All applications to one specific job have been posted simultaneously to ensure that they would arrive the same day. In addition, applications have been posted in different post offices in order to avoid detection.

Overall, this study has been designed to reduce the potential biases of the correspondence testing methodology that have been highlighted in the literature. As a consequence, this data collection should allow for a better estimation of the hiring discrimination.

As an additional safety, we also control for the characteristics of the applications themselves when we perform the regression analysis.

The answer to the application is recorded as positive when the recruiter invites the candidate to the hiring interview or when the recruiter asks for more information about the availability and qualifications of the candidate. Conversely, the answer is recorded as negative when the recruiter rejects the application or does not answer it. Since each application is sent for both genders, there are four possible outcomes: neither candidate is invited, both candidates are invited, the man only is invited or the woman only is invited. The results are reported in Table 4.

TABLE 4
Answers to the correspondence tests
There are 157 observations for each correspondance test

Type of correspondance test	Neither invited	Both invited	Discrimination against women	Discrimination against men
25, single, childless	75	44	22	16
37, single, childless	118	12	14	13
37, married, 3 children	124	10	12	11

3 Model and estimation

The model and its estimation result from three characteristics that are specific to experimental samples. First, women and men have the same value of the explanative variables. Second, the samples have a small size. Third, by construction, there is no selection bias, since the values of the explanatory variables have been set by the researcher.

3.1 Model

We model the answer of the recruiter to the applications. Let $Y_{k,i}$ the answer to the application i for the gender k (male, female). We have $Y_{k,i} = 1$ if the candidate obtains an interview and $Y_{k,i} = 0$ otherwise. The characteristics of both genders are identical for each i and are regrouped in a common vector X_i. A correspondence test can be summarized in the vector ($Y_{m,i}$, $Y_{f,i}$, Y_i) where m denotes male and f female. Let P_m the probability that a man gets an interview and P_f the corresponding probability for a woman. A first definition of discrimination is simply: $\tilde{D} = P_m - P_f$. In a standard correspondence test, one simply computes the average difference between the success percentages of men and women. This difference can be directly obtained by performing an ordinary least square (henceforth, *OLS*) regression of the differences $D_i = Y_{m,i} - Y_{f,i} \in \{-1,0,1\}$ on a constant term.[5] We have: [6]

$$E(D_i) = E(Y_{m,i} - Y_{f,i}) = E(Y_{m,i}) - E(Y_{f,i}) = P_m - P_f = \tilde{D}.$$

The *OLS* regression of D_i on a constant term provides an unbiased and consistent estimator of discrimination. This first measurement of discrimination is sometimes considered as too crude since the same characteristics can be perceived differently by the recruiters depending on they belong to women or men (Heckman [1998]).

In order to answer this critique, we generalize the model by explaining the probabilities to get an interview by all the characteristics X_i. In order to simplify the interpretation of the regression coefficients, we center the explanative variables, so that $\overline{X} = 0$. We get: [7]

$$E(Y_{k,i} | X_i) = P_k^0 + X_i \beta_k \ , \ \ k = m, f$$

which implies:

$$E(D_i | X_i) = P_m^0 - P_f^0 + X_i (\beta_m - \beta_f).$$

The centering of the explanative variables implies that:

$$E(D_i) = \underset{X}{E}\left[E(D_i | X_i)\right] = P_m^0 - P_f^0 .$$

This measure of discrimination includes two parts. The first component, estimated by the constant term of the model, $P_m^0 - P_f^0$, measures the discrimination at the average point of the sample (defined by $X_i = \overline{X} = 0$). The second component admits three different interpretations. First, when the variable from X_i that is considered is related to discrimination, the quantity $X_i(\beta_m - \beta_f)$ measure *conditional* discrimination, that is a type of discrimination that only applies to a subgroup of the population under study (*i.e.,* only to a part of the women). Second, when the variable from X_i that is considered is related to the productivity of the candidate (or to the labor cost), the quantity $X_i(\beta_m - \beta_f)$ represents the differences of productivity or labor cost that are perceived by the recruiter. Third, when the variable from X_i that

---

5. In general, an *OLS* regression of a variable on a constant term gives its sample mean.
6. The binary variable Y_k is distributed according to a Bernoulli distribution with parameter P_K ($k = m,w$).
7. The centering of the explanative variables does not imply any loss of generality since there is a constant term in the model. Notice that *we do not* center the left-hand variable.

is considered deals with the characteristics of the correspondence test itself (type of CV, date of mailing), the quantity $X_i(\beta_m$-$\beta_f)$ represents the biases that would have appeared if no correction had been made.

Standard correspondence tests often rely on the assumption that $\beta_m = \beta_f$. We can test this assumption by examining whether there is a significant explanative variable in the regression of D_i on X_i.

3.2 **Estimation**

The estimation method can be simplified thanks to the data collection methodology. We avoid two sources of complications. First, there is no self-selection since the researcher makes the applications; only the answers of the recruiter are real. Second, there is no unobserved heterogeneity, so that panel data methods are not needed.

Therefore, it is possible to use the *OLS* estimator provided that the predictions lies in the interval]-1,+1[, since we evaluate the difference of two probabilities P_m-P_f.[8] Here, there is a specific reason why the predictions should lie in the right interval: all the explanative variables are dummy variables belonging to {0,1} so that they do not take extreme values. This property of the explanative variables explains why the predictions should not take extreme values for a reasonable value of the regression coefficients. Table 5 provides the extreme values of the *OLS* predictions.

TABLE 5
Admissibility of OLS regressions

Type of correspondence test	$\hat{D}$	
	Minimum	Maximum
Aged 25, single, childless	–0.585	+0.555
Aged 37, single, childless	–0.272	+0.331
Aged 37, married with three children	–0.403	+0.456

We have also estimated ordered Logit and Probit models and they provide comparable results to the main ones of the linear model. Overall, the efficiency of these estimators seems to be lower than the ones we have retained. They are presented in Appendix 1.

But another caution should be taken. The sample size is small and a linear model on a difference of two probabilities can produce heteroskedastic disturbances.[9] Therefore, we estimate the standard errors by the bootstrap method.[10] We perform separate regressions for each type of correspondence test (see Table 4). Therefore we have 157 observations for each regression. There remains to set the number of

8. The choice of *OLS* also avoids to use non linear methods that are valid on large samples only.
9. On this point, see MADDALA [1983].
10. On this method, see EFRON and TIBSHIRANI [1993].

bootstrap repetitions. We use the method of Andrews and Buchinsky [2000], presented in appendix 2. Overall, the bootstrapped standard errors tend to be higher than the *OLS* asymptotic standard errors. All the estimations have been performed with *SAS*.

4 Results

Due to the size of our samples, we will comment on the effects that are significant at the 5% and at the 10% levels.[11]

A direct comparison of the means does not reveal significant differences of access to hiring interviews between women and men (Table 8). Similarly, the constant terms of the regressions are never significant (Tables 9 to 11). Therefore, there is no significant discrimination at the average point of the sample. On the contrary, there are significant *conditional* differences between women and men. The absence of discrimination on average conceals an interesting compensation effect.

A higher qualification increases the maternity cost incurred by the firms (F_f) because it is proportional to the labor cost. We find that the women childless aged 25 with a high qualification have a lower probability than similar men to access hiring interviews. When the age reaches 37, this inequality reverses and the childless women aged 37 have a higher probability than men to access hiring interviews. Therefore, the success of women is increasing with their age. This can be related to the fact that women aged 37 have a lower maternity probability (p) than the women aged 25. For women aged 37 with three children, we find the success probability is the same than men. This implies that the number of children would in fact reduce the probability of qualified women to access interviews. This result is likely to come from the fact the employer expects less flexibility (a higher M) from women with a significant family burden. A similar result is found for the funding of training by the firm. Childless women aged 25 have a lower probability than men to obtain an interview on a job that explicitly include training paid by the employer. This difference vanishes for women aged 37, with or without children. We have also computed the *overall* effect of qualification on the probability to get an interview (Table 6). Our basic result is confirmed: young qualified women have a lower probability to get an interview at the 5% level.

The other results deal with the screening through the clauses of the labor contract, which aim to reduce the M parameter. The mainstream of screening concentrates on childless women aged 25. They have fewer chances to access interviews when the job is a long-run contract. This effect cancels for 37 years old women. Childless women aged 37 are not globally subject to discrimination: they even have more chances than men to access interviews on commercial jobs. This result is confirmed for the women aged 37 with three children: they have more chances than men to access interviews on jobs with a possibility of promotion.

11. When the normality tests are rejected, we use the percentages of positive and negative estimates, that have been computed independently from the standard errors. They are given in the column « Proba[Coeff>0] or Proba[Coeff<0] » depending on the coefficient is positive or negative. A value of 0.95 or higher means that the coefficient is significant at the 5% level.

TABLE 6
Conditional Effect of Qualification
Estimation of $E(D_i|\text{highly qualified})$

Type of correspondence test	Bootstrap Estimate	Bootstrap Standard Error	Proba[Coeff>0] or Proba[Coeff<0]	Normality Cramer-Von Mises p-value
Aged 25, childless	0.131**	0.063	0.984	0.250
Aged 37, childless	-0.024	0.045	0.695	0.005
Aged 37 with three children	0.023	0.045	0.679	0.005

*** significant at the 5% level.*

TABLE 7
Sample Statistics

Variables	Percentage
Baccalauréat	42.1
BTS	57.9
Administrative	37.2
Commercial	62.8
Application sent directly	63.5
Application sent through ANPE	36.5
Short-term labor contract	20.7
Long-term labor contract	79.3
Promotion indicated	23.0
Training funded by the firm	15.2
Negotiable wage	38.9
Wage depending on output	11.0
Gender of the recruiter	
Woman	68.5
Man	31.5
Line of business:	
Commercial banks	*36.8*
Trust banks	*19.3*
Mutualist banks	*15.6*
Savings banks	*3.4*
Factoring banks	*7.1*
Credit organism	*14.3*
La Poste	*3.4*
January 2002	*12.4*
February 2002	*47.6*
March 2002	*40.0*
CV Type A	63.9
CV Type B	36.1

TABLE 8
Comparison of the means
Left-hand variable: difference of answers between men and women (D_i).
Sample of 157 observations. Estimation by the bootstrap. The number of simulations is determined by the method of ANDREWS and BUCHINSKY [2000] with *pdb*=5 and τ=0.05.

Type of correspondence test	OLS Estimates	Average of Bootstrap Estimates	Optimal number of simulations	Bootstrap Standard Error	Pr [coeff>0]	Normality Cramer-Von Mises
Childless, single, aged 25	0.041	0.041	385	0.044	0.808	0.008
Childless, single, aged 37	0.014	0.012	1052	0.035	0.596	0.005
Aged 37, married with three children	0.007	0.007	1262	0.032	0.551	0.005

Globally, the main differences of treatment concentrate on childless young women. Institutional mechanisms, however, counterbalance this negative effect. Childless young women have more chances than men to access interviews at La Poste or in credit organisms. The first result can be explained by the fact that La Poste is a public firm that promotes social objectives. In the case of credit organisms, which mostly operate by phone, the explanation could rely on the preferences of customers (that reduces *M*).

Finally, passing through the ANPE seems to have a positive effect on the chances to access interviews for women aged 37 with three children. Therefore, the existence of some institutional practices seems to compensate the cost differences between women and men.

Conclusion

Data collection by field experiments does not exempt from the use of econometric methods. We show that one should take into account all the characteristics of the application, of the type of job and of the firm.

The main differences of treatment between women and men originate in the presence of children and on the maternity expectations of the recruiters. These informations, that already play an important role on the labor supply, would also play an important role on the labor demand in the financial sector. This first effect, negative, on the access to interviews, is compensated by a positive effect taking place at La Poste, credit organisms and ANPE.

Since the access to hiring interviews drives the access to jobs, we should see a reallocation of the female workforce toward the financial lines of business where the working conditions are the most compatible with family obligations. Even though it is difficult to generalize from this study, our results suggest that the equilibrium on the financial labor market would depend on the evolution of demography.

TABLE 9

Childless, single, aged 25

Left-hand variable: difference of answers between men and women D_i.

Sample of 157 observations. Estimation by the bootstrap. The number of simulations is determined by the method of ANDREWS and BUCHINSKY [2000] with *pdb*=5 and τ=0.05.

We use 969 bootstrap repetitions for all the estimates.

The *OLS* standard errors are corrected for possible heteroskedasticity.

Variable	*OLS* Estimates	*OLS* Corrected Standard Error*	Average of Bootstrap Estimates	Optimal Number of Repetitions	Bootstrap Standard Error	Pr[coeff>0] or Pr[coeff<0]	Normality Cramer-Von Mises
Intercept	0.041	0.039	0.039	345	0.043	0.640	0.058
Characteristics of the candidate (ref. Baccalauréat and Administrative):							
– BTS	0.225**	0.078	0.223	16	0.088	0.990	0.250
– Commercial	-0.120	0.090	-0.122	204	0.100	0.882	0.250
Characteristics of the job offer (ref. Short-run, no promotion, no training, non negotiable wage, wage independent of output)							
– Long-run	0.266**	0.113	0.266	30	0.129	0.980	0.250
– Promotion expected	-0.043	0.100	-0.046	861	0.113	0.666	0.250
– Training funded	0.212*	0.125	0.225	89	0.142	0.949	0.244
– Negotiable wage	0.084	0.090	0.086	394	0.099	0.818	0.168
– Wage depending on output	0.124	0.141	0.114	476	0.154	0.769	0.007
Characteristics of the firm (ref. Female recruiter, commercial bank)							
– Male recruiter	0.067	0.090	0.064	569	0.102	0.730	0.213
– Trust bank	-0.146	0.135	-0.143	303	0.146	0.835	0.250
– Mutualist bank	-0.192	0.157	-0.186	223	0.169	0.873	0.250
– Savings bank	-0.071	0.114	-0.061	968	0.152	0.647	0.177
– Factoring bank	0.057	0.146	0.068	969	0.163	0.659	0.235
– Credit organism	-0.242**	0.123	-0.235	98	0.140	0.961	0.105
– La Poste	-0.285*	0.139	-0.269	207	0.183	0.925	0.005
Characteristics of the correspondence test (ref. Application sent directly, January 2002, CV type A)							
– Through ANPE	0.133	0.120	0.140	272	0.136	0.849	0.025
– February 2002	0.051	0.156	0.055	918	0.174	0.641	0.024
– March 2002	0.157	0.160	0.169	336	0.183	0.824	0.134
– CV type B	0.035	0.080	0.034	882	0.087	0.640	0.135

* *significant at the 10% level*

** *significant at the 5% level*

TABLE 10

Childless, single, aged 37

Left-hand variable: difference of answers between men and women D_i.

Sample of 157 observations. Estimation by the bootstrap. The number of simulations is determined by the method of ANDREWS and BUCHINSKY [2000] with *pdb*=5 and τ=0.05.

We use 1463 bootstrap repetitions for all the estimates.

The *OLS* standard errors are corrected for possible heteroskedasticity.

Variable	*OLS* Estimates	*OLS* Corrected Standard Error	Average of Bootstrap Estimates	Optimal Number of Repetitions	Bootstrap Standard Error	Pr[coeff>0] or Pr[coeff<0]	Normality Cramer-Von Mises
Intercept	0.014	0.034	0.013	832	0.038	0.636	0.250
Characteristics of the candidate (ref. Baccalauréat and Administrative):							
- BTS	-0.132*	0.070	-0.121	103	0.077	0.937	0.250
- Commercial	-0.151*	0.081	-0.143	104	0.090	0.936	0.250
Characteristics of the job offer (ref. Short-run, no promotion, no training, non negotiable wage, wage independent of output)							
- Long-run	0.008	0.098	0.014	1359	0.111	0.558	0.250
- Promotion expected	-0.003	0.103	-0.016	1463	0.120	0.548	0.132
- Training funded	-0.037	0.109	-0.022	1120	0.123	0.578	0.114
- Negotiable wage	-0.109	0.074	-0.097	188	0.083	0.886	0.030
- Wage depending on output	0.130	0.113	0.125	366	0.132	0.838	0.005
Characteristics of the firm (ref. Female recruiter, commercial bank)							
- Male recruiter	-0.034	0.087	-0.034	844	0.095	0.645	0.123
- Trust bank	-0.129	0.108	-0.134	232	0.115	0.885	0.250
- Mutualist bank	-0.154	0.159	-0.153	360	0.169	0.810	0.138
- Savings bank	-0.046	0.096	-0.051	1090	0.125	0.658	0.059
- Factoring bank	-0.205	0.168	-0.205	216	0.188	0.876	0.190
- Credit organism	-0.022	0.122	-0.015	1352	0.132	0.532	0.131
- La Poste	0.129	0.307	0.138	844	0.382	0.645	0.005
Characteristics of the correspondence test (ref. Application sent directly, January 2002, CV type A)							
- Through ANPE	-0.029	0.102	-0.039	981	0.109	0.648	0.250
- February 2002	0.124	0.158	0.118	507	0.176	0.752	0.170
- March 2002	0.075	0.161	0.066	809	0.177	0.655	0.250
- CV type B	-0.104	0.080	-0.100	229	0.090	0.870	0.250

* *significant at the 10% level*

** *significant at the 5% level*

TABLE 11

Aged 37, married with 3 children

Left-hand variable: difference of answers between men and women D_i.

Sample of 157 observations. Estimation by the bootstrap. The number of simulations is determined by the method of ANDREWS and BUCHINSKY [2000] with *pdb*=5 and τ=0.05.

We use 1198 bootstrap repetitions for all the estimates.

The *OLS* standard errors are corrected for possible heteroskedasticity

Variable	*OLS* Estimates	*OLS* Corrected Standard Error	Average of Bootstrap Estimates	Optimal Number of Repetitions	Bootstrap Standard Error	Pr[coeff>0] or Pr[coeff<0]	Normality Cramer-Von Mises
Intercept	0.007	0.031	0.006	1198	0.034	0.563	0.011
Characteristics of the candidate (ref. Baccalauréat and Administrative):							
- BTS	0.037	0.060	0.032	738	0.068	0.678	0.250
- Commercial	-0.026	0.075	-0.024	963	0.082	0.608	0.250
Characteristics of the job offer (ref. Short-run, no promotion, no training, non negotiable wage, wage independent of output)							
- Long-run	-0.074	0.080	-0.075	456	0.093	0.780	0.250
- Promotion expected	-0.120*	0.069	-0.119	175	0.077	0.942	0.250
- Training funded	0.025	0.059	0.033	1047	0.071	0.690	0.250
- Negotiable wage	-0.018	0.063	-0.018	1062	0.070	0.612	0.250
- Wage depending on output	0.035	0.133	0.033	1043	0.153	0.593	0.250
Characteristics of the firm (ref. Female recruiter, commercial bank)							
- Male recruiter	-0.050	0.080	-0.050	594	0.088	0.713	0.250
- Trust bank	-0.088	0.103	-0.082	444	0.113	0.773	0.243
- Mutualist bank	-0.134	0.119	-0.127	267	0.131	0.848	0.197
- Savings bank	0.329	0.223	0.336	193	0.271	0.889	0.005
- Factoring bank	-0.121	0.193	-0.110	672	0.211	0.693	0.250
- Credit organism	0.048	0.120	0.054	830	0.132	0.649	0.250
- La Poste	-0.053	0.074	-0.054	978	0.098	0.714	0.005
Characteristics of the correspondence test (ref. Application sent directly, January 2002, CV type A)							
- Trough ANPE	-0.157*	0.078	-0.162	74	0.087	0.933	0.250
- February 2002	-0.136	0.109	-0.139	243	0.119	0.884	0.005
- March 2002	-0.180*	0.129	-0.181	188	0.144	0.914	0.250
- CV type B	0.042	0.076	0.048	659	0.082	0.685	0.250

* *significant at the 10% level*

** *significant at the 5% level*

Since the unfavorable treatment of women that we observe is related to the difference of expected labor costs, one should be tempted to conclude that there is no discrimination. However, the conclusion may be different if we think to fairness.

Indeed, the children obviously result from a joint decision of women and men and our results tend to show than only women would bear the consequences. The fact that the decision is joint suggests that a fairer situation could be achieved by spreading the costs over both women and men. This would include the maternity costs and the replacement costs. ■

References

Arrow K.J. (1972a). – « Models of job discrimination », In Pascal A.H. ed., *Racial Discrimination in Economic Life,* Lexington Mass.: Lexington Books, p. 83-102.

Arrow K.J. (1972b). – « Some mathematical models of race in the labor market », In Pascal A.H. ed., *Racial Discrimination in Economic Life,* Lexington Mass.: Lexington Books, p. 187-204.

Arrow K.J. (1973). – « The theory of discrimination », In Ashenfelter O.A. and Reeds A. eds., *Discrimination in Labor Markets*, Princeton University Press, p. 3-33.

Andrews D. and Buchinsky M. (2000). – « A three-step method for choosing the number of bootstrap repetitions », *Econometrica,* vol. 67, p. 23-51.

Becker G.S. (1957). – « The Economics of Discrimination », *Chicago: University of Chicago Press,* 2nd Ed., 1971.

Blau F. and Kahn L.M. (1994). – « The impact of the wage structure on trends in U.S. gender wage differentials: 1975-1987 », *Working paper NBER,* no. 4748.

Blinder A.S. (1973). – « Wage discrimination: reduced forms and structural estimates », *Journal of Human Resources,* Vol. 8, no. 4, p. 436-55.

Cain G. (1986). – « The economic analysis of labor market discrimination: a survey », *Handbook of Labor Economics,* vol. 1, p. 694-785.

Crépon B., Deniau N. and Pérez-Duarte S. (2001). – « Wages, Productivity and Worker Characteristics: A French Perspective », mimeo CREST-INSEE (www.ensae.fr).

Efron B. and Tibshirani J. (1993). – « An introduction to the bootstrap », *Monographs on Statistics and Applied Probability,* no. 57, Chapman & Hall (ISBN 0-412-04231-2).

Fakhfakh F., Merlateau M.-P. and Meurs D. (2002). – « Les disparités de rémunération entre hommes et femmes : la situation de quatres branches professionnelles : Une étude qualitative et économétrique », *DARES,* mimeo.

Heckman J.J. (1998). – « Detecting Discrimination », *Journal of Economic Perspectives,* Vol. 12, no. 2, p. 101-16.

Hellerstein J.K., Neumark D. and Troske K.R. (1999). – « Wage, Productivity and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations », *Journal of Labor Economics,* Vol. 17, no. 3, p. 409-46.

Kenney G.M. and Wissoker D.A. (1994). – « An Analysis of the Correlates of Discrimination Facing Young Hispanic Job-Seekers », *American Economic Review,* Vol. 84, no. 3, p. 674-83.

Kunze A. (2000). – « The Determination of Wages and the Gender Wage Gap: A Survey », *IZA Working Paper,* no. 193, (www.iza.org).

Lommerud K.E. and Vagstad S. (2000). – « Mommy tracks and public policy: on self-fullfiling prophecies and gender gaps in promotion », *Working paper CEPR*, no. 2378.

MacCall J.J. (1972). – « The simple mathematics of information, job search and prejudice », In Pascal A.H. ed., *Racial Discrimination in Economic Life,* Lexington Mass.: Lexington Books, p. 205-24.

Maddala G.S. (1983). – « Qualitative and limited-dependent variables in econometrics », *Econometric Society Monograph*, no. 3, Cambridge University Press.

MEURS D. and PONTHIEUX S. (2000). – « Une mesure de la discrimination dans l'écart de salaire entre hommes et femmes », *Économie et Statistiques,* n° 337-38, p. 135-58.

NEUMARK D., BANK R.J. and VAN NORT K.D. (1996). – « Sex discrimination in restaurant Hiring: An Audit Study », *Quaterly Journal of Economics,* p. 915-41.

OAXACA R. (1973). – « Male-female wage differentials in urban labor markets », *International Economic Review,* Vol. 14, p. 693-709.

PETIT P. (2003). – « Comment évaluer la discrimination à l'embauche ? », *Revue Française d'Économie,* Vol. 17, n° 3, p. 55-87.

PETIT P. (2004). – « Discrimination à l'embauche : étude d'audit par couple dans le secteur financier », *Revue Économique,* à paraître. Cahiers de la MSE – EUREQua – 2003(16).

PHELPS E.S. (1972). – « The statistical theory of racism and sexism », *American Economic Review,* vol. 62(4), p. 659-61.

RIACH P.A. et RICH J. (1991). – « Testing for Racial Discrimination in the Labour Market », *Cambridge Journal of Economics,* Vol. 15, p. 239-56.

VUONG Q.H. (1989). – « Likelihood ratio tests for model selection and non-nested hypotheses », *Econometrica,* 57(2), p. 307-33.

APPENDIX 1

Ordered Logit Regressions

In order to check that the estimation method does not drive our results, we have estimated ordered Logit and Probit models (using SAS). We chose the Ordered Logit model since it performs better according to Vuong's test (1989). The non-nested tests are presented in Table A.1 and the ordered Logit regressions in Table A.2. Our results remain unchanged by the application of this different estimation method.

TABLE A.1
Vuong test of Ordered Logit versus Ordered Probit

Type of correspondence test	Vuong t-statistic	p-value	Preferred model at the 5% level
Aged 25, single, childless	1.65	0.101	Equivalent
Aged 37, single, childless	2.18	0.030	Ordered Logit
Aged 37, married with three children	3.92	0.001	Ordered Logit

TABLE A.2

Ordered Logit regressions

Left-hand variable: difference of answers between men and women $D_i \in \{-1,0,1\}$.
Maximum likelihood estimation.

Type of correspondence test	Childless, single, aged 25		Childless, single, aged 37		Aged 37, married with 3 children	
	Estimates	*p*-value	Estimate	*p*-value	Estimate	*p*-value
First Intercept	-1.990^{**}	0.001	-2.476^{**}	0.001	-2.798^{**}	0.001
Second Intercept	2.432^{**}	0.001	$2\ .653^{**}$	0.001	2.890^{**}	0.001
Characteristics of the candidate (ref. Baccalauréat and Administrative):						
- BTS	1.169^{**}	0.008	-0.824^{*}	0.086	0.280	0.578
- Commercial	-0.607	0.180	-0.998^{*}	0.057	-0.145	0.795
Characteristics of the job offer (ref. Short-run, no promotion, no training, non negotiable wage, wage independent of output)						
- Long-run	1.308^{**}	0.027	0.089	0.891	-0.472	0,509
- Promotion expected	-0.215	0.696	-0.025	0.967	-0.852	0.219
- Training funded	0.939	0.119	-0.234	0.739	0.136	0.858
- Negotiable wage	0.438	0.335	-0.690	0.190	-0.078	0.887
- Wage depending on output	0.606	0.431	0.867	0.337	0.290	0.771
Characteristics of the firm (ref. Female recruiter, commercial bank)						
- Male recruiter	0.294	0.521	-0.241	0.654	-0.370	0.519
- Trust bank	-0.784	0.206	-0.860	0.226	-0.748	0.350
- Mutualist bank	-0.927	0.158	-1.047	0.163	-0.986	0.223
- Savings bank	-0.354	0.760	-0.295	0.824	1.959^{*}	0.077
- Factoring bank	0.220	0.804	-1.386	0.197	-0.942	0.446
- Credit organism	-1.193^{*}	0.102	-0.146	0.856	0.383	0.659
- La Poste	-1.339	0.222	0.848	0.452	-0.424	0.765
Characteristics of the correspondence test (ref. Application sent directly, January 2002, CV type A)						
- Trough ANPE	0.659	0.229	-0.175	0.780	-1.135^{*}	0.092
- February 2002	0.302	0.675	0.855	0.299	-0.924	0.289
- March 2002	0.800	0.308	0.530	0.553	-1.172	0.224
- CV type B	0.134	0.771	-0.695	0.185	0.276	0.646
% concordant predictions	71.2		68.9		73.2	

* *significant at the 10% level*

** *significant at the 5% level*

APPENDIX 2

Number of Bootstrap repetitions

The criterion that determines the number of bootstrap repetitions is the following:

$$\Pr\left[100\times\left|\frac{\hat{\theta}_B-\hat{\theta}_\infty}{\hat{\theta}_\infty}\right|\leq pdb\right]=1-\tau$$

where $\hat{\theta}_B$ is the bootstrap estimator on B simulations and $\hat{\theta}_\infty$ the "ideal" bootstrap estimator that would be obtained with an infinite number of repetitions. The quantity *pdb* represents the upper bound on the error, expressed as a percentage of the "ideal" estimator, and the quantity 1-τ is the probability with which one reach that result. For this application, we have set *pdb*=5 and τ=0.05, which implies a maximum error of 5% with a probability equal to 95%. The parameter on which the number of bootstrap repetition is determined also matters. We have chosen the *p*-values of the *OLS* estimators.

More precisely we use the probability that positive *OLS* coefficients are strictly positive and the probability that negative *OLS* coefficients are strictly negative, in order to avoid an arbitrary large number of simulations (ANDREWS and BUCHINSKY [2000]).

The estimation method includes three steps.[12] In the first step, we compute an initial number of simulations for each of the estimated parameter, equal to:

$B_1=\left[\frac{10000\times z^2_{1-\tau/2}}{pdb^2}\times\frac{1-\hat{p}}{\hat{p}}\right]$ for a positive coefficient and to $B_1=\left[\frac{10000\times z^2_{1-\tau/2}}{pdb^2}\times\frac{\hat{p}}{1-\hat{p}}\right]$ for a negative coefficient, where $\hat{p}$ is an estimator of $\Pr[\theta>0]$. With the previous values of *pdb* and τ we get:

$B_1=\left[1536\times\frac{1-\hat{p}}{\hat{p}}\right]$ for a positive coefficient and $B_1=\left[1536\times\frac{\hat{p}}{1-\hat{p}}\right]$ for a negative coefficient.

In this first step $\hat{p}$ is computed by using the asymptotic normality of the *OLS* estimation. This convention does not affect the validity of the final number of bootstrap repetition. It is just a starting value. In practice, the *OLS* gives a set of estimated parameters so that we take the *maximum* number of simulations to perform the *OLS* regressions.

12. In what follows, z_q represents que q-th percentile of the standard normal distribution and [.] denotes the integer part of a number.

In the second step of the method, we make $\overline{B}_1 = \max\{B_1\}$ simulations and compute a non-parametric estimate of $\hat{p}$, by taking the proportion of strictly positive estimates or the proportion of strictly negative estimates depending on the explanative variable. On denote this new estimation of p by $\hat{\pi}$. The new number of simulations is given by:

$B_2 = \left[1536 \times \frac{1-\hat{\pi}}{\hat{\pi}}\right]$ for a positive coefficient and $B_2 = \left[1536 \times \frac{\hat{\pi}}{1-\hat{\pi}}\right]$ for a negative coefficient.

Accounting for all the regression coefficients, the actual number of simulations is equal to $\overline{B}_2 = \max\{B_2\}$. The optimal number of simulations is defined by:

$$B^* = \max\{\overline{B}_1, \overline{B}_2\}$$

If the number of simulations made in the first step is sufficient (*i.e.* $\overline{B}_1 \geq \overline{B}_2$), we stop the method at the second step and compute the statistics on the $\overline{B}_1$ repetitions already made. Otherwise, we pass the third step.

In the third step, if $\overline{B}_1 < \overline{B}_2$, we make $B^* - \overline{B}_1$ additional simulations, that are added to the first step simulations. Then, we compute the statistics from the B^* simulations. The optimal numbers of simulations are presented in the Tables 8 to 11, knowing that *all* statistics are computed on the *maximum* of the number of simulations.