Title: **Explaining Gender Differences in Quality Management in Schools:**

A Tale of Two Ceilings

Author(/s): **Miryam Martínez Martínez; Ruth Mateos de Cabo; Manuel M. Molina-López**

Date: **October-2018**

WORKING PAPER SERIES: 5/2018

CÁTEDRA CEU-MM
Madrid (Spain)
ISSN online: 2530-0237

Explaining Gender Differences in Quality Management in Schools: A Tale of Two Ceilings

Miryam Martínez Martínez, Ruth Mateos de Cabo, Manuel M. Molina-López

Universidad CEU San Pablo, Business Economics Department, Madrid (Spain)

Preliminary version: October, 2018

Abstract:

Despite the higher presence of females in the school teacher staff, the percentage of women school principals is lower than the male one. The aim of the present paper is to explore if among the possible causes behind the gender gap observed in school principal positions there are factors that could produce gender bias. We find that female principals are associated with higher management quality in both, people and non-people management. We propose two alternative explanations for this result. One, on the demand-side, is the presence of double standards in the promotion of women to leadership positions. Alternatively, on the supply-side, the presence of a potential lower women's self-efficacy perception could lead them to self-exclude form managerial positions if they consider they are not ready or enough prepared. Both cases could produce that only higher skilled female than male leaders would become principals. We explore which cause is more plausible by looking at the interactions of the principal gender with micro, meso and macro-factors. The fact that only meso and macro-factors are the relevant ones, points to the double standards (supply-side) as the more likely cause of the observed gender differences in management quality in schools.

Keywords: schools; female leadership; management; talent; human resources; gender bias

Acknowledgements: we thank Ricardo Gimeno, Daniela Scur, Siri Terjesen and Alison Cook and the different assistants to the I Workshop in Female Leadership in Madrid, the VI Workshop on Talent Management in Barcelona and EURAM2018 in Reykjavik for their helpful comments. Project of I+D+I FEM2017-83006-R founded by the Ministerio de Economía, Industria y Competitividad-Agencia Española de Investigacióon (AEI) and FEDER (UE) and CEU-Mutua Madrileña Chair: 060516-USPMM-02/17 -

1. Introduction.

In the OECD countries (OECD, 2017) the average of female teachers in secondary education makes up 64% of the teaching population. In lower secondary education 69% of teachers on average across OECD countries are women and at the upper secondary level, the OECD average drops to 59%. But, despite the fact that the number of professionally qualified women in education is larger than those of males their presence in principal positions is still quite low (Novo-Corti et al., 2018). Indeed, while the proportion of male teachers is relatively small in many countries, in comparison there is an over-representation of male principals, in fact, only 45% of principals in the lower secondary schools are women in the OECD average (OECD, 2016).

The reasons that prevent women from reaching leadership positions can be explained both by demand-side and supply-side factors of the market of qualified women (Gabaldon et al., 2016). On the demand side there could be bias that comes from gender bias organizational policies that cause double standards in the promotion of women to managerial positions. In contrast, on the other side there are internal barriers (such as lower self-efficacy perception in the workplace) that could lead women to self-exclude form leadership positions if they consider they are not ready or enough prepared to occupy such positions.

Try to disentangle supply-side from demand-side actors as sources of women's underrepresentation on top positions is difficult because they are usually mixed and distorted by elements of segregation (vertical and horizontal) in the labor market (Mateos de Cabo et al., 2011). Nevertheless, in contrast to other sectors, school principals are former teachers, whereas in other economic sectors top management teams may come from other areas of activity (politicians, CEOs of other companies, academics,...). Thus, given the fact that the top management team of the educational centers are usually selected among the teachers of the

school staff or from other ones, the pool of female candidates for leadership positions is clearly higher in the educational sector than in another ones, so the lack of suitably qualified candidates to occupy these positions could not be used to explain why so few principal positions in schools are held by women.

In this study, we try to identify signals of gender bias in the promotion process to school principals by finding gender differences in the quality of the principal's management. If there were biases in the promotion processes hindering women's access to managerial positions, it would imply that greater qualifications to become a principal would be required to women. As a consequence women principals would perform better as managers than their male counterparts. The presence of a double standard would allow us to explain the gender gap found among school principals, as well as possible gender differences in their management's quality. In our study, double standard implies setting higher bars to evaluate women than men, even when they perform at the same level (Foschi, Lai & Sigerson, 1994).

An alternative explanation for having a lower proportion of female principals and at the same time higher management quality scores could also imply that women tend to have lower self-efficacy perception that lead them to self-exclude when they are offer a principal position. So, in this paper we explore both sides of the school principals market as possible causes for both the reduce presence of women in school principal positions and the higher quality management scores they could produce. On the supply-side we have the self-exclusion whereas on the demand-side are the double standards the ones explaining these stylized facts.

In order to disentangle which one of this causes is the plausible one, we interact the principal gender with factors at the micro, meso and macro level, as defined by Terjesen et al. (2009). We consider that if the demand-side (i.e., double standards in the promotion process) is responsible for the gender differences in quality management, then macro and meso factors (i.e. external factors), such as organizational size or degree of competition, would have a moderator

effect of the principal gender on the management quality. Whereas, if supply-side (self-exclusion due to lower self-efficacy perception) is the cause for those gender differences, then the micro factors (i.e. principal characteristic) would moderate the relationship between gender and management quality.

To test our predictions, we use a management index built upon 20 basic management practice measures divided into two groups: people and non-people management. The management index for each school represents the average of these scores (Bloom, Lemos, Sadun & Van Reenen, 2015). The rest of the paper is organized as follows. Section 2 introduces the literature review and the hypothesis to be tested. Section 3 presents the database and the sample and methodology. Section 4 shows the main results. And, finally, section 5 concludes the paper.

2. Theory and hypotheses

Double standard and women lower self-efficacy perception

Looking at school principals, we observe two stylized facts: i) there are more men than women among school principals. This is even more striking if we take into account that this is a heavily feminized sector (i.e., female teachers are a majority across countries): ii) women working as school principals tend to exhibit higher quality management scores. We will show this fact holds in a later section.

Given these stylized facts we build our theoretical framework around two theories that could accommodate both of them. *Double Standard Theory* and *Women lower Self-Efficacy Perception Theory*. In order to show how both theories work, we present them in a mathematical/formalized way.

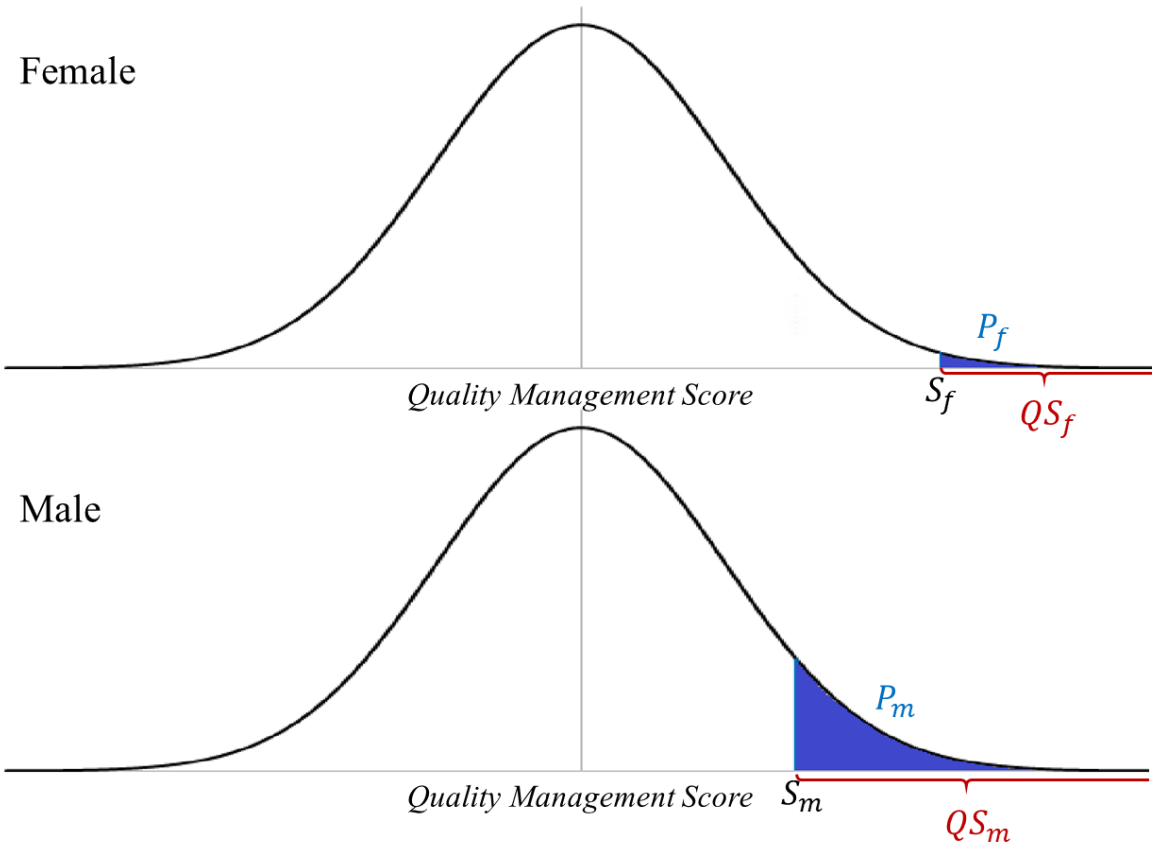
One the demand side: *Double Standard Theory*

Figure 1 gives a schematic representation of the Double Standard Theory. Bell shape curves represent the distribution of Quality Management Scores for both female (top) and male

(bottom) teachers. According to the theory, teachers with best Quality Management Scores are the ones promoted to principal positions. For the shake of simplicity, we assume that both male and female teachers are equally qualified for those positions, as represented by distributions with equal means and standard deviations[1]. Under this theory, the divergences enunciated above as stylized facts appear as a consequence of the double standards applied for female and male candidates to school principal positions. Male candidates are considered for principal positions if their Quality Management Scores (QS_m) are above a minimum standard (S_m), thus: $QS_m > S_m$. However, in the case of female candidates, the minimum standard applied to them is higher than the one applied to men ($S_f > S_m$), i.e., there is a *Double Standard*. This double standard has two consequences: firstly, the number of women that satisfies the condition $QS_f > S_f$ is much lower than that of males, where $QS_m > S_m$, thus, we will end with a lower population of female principals ($P_f < P_m$); secondly, among those female principals that have overcome the double standard, the average quality management scores will be higher than those of their male counterparts ($E\left[QS_f/QS_f > S_f\right] > E[QS_m/QS_m > S_m]$).

[1] Similar outcomes arise even with divergences in the underline distribution of management scores.

Figure 1: Double Standard Theory



The presence of double standards of competence would imply that status characteristics (e.g. gender, ethnicity, socioeconomic class) become a basis to stricter standards for the lower status person (in our case, women are lower-status members) (Foschi 2000). Given that women are members of lower-status group a strict standard for ability in settings involving evaluation, such as a selection process to occupy a managerial position, would mean that they would be required more evidence of competence (this is, a higher minimum quality score than their male counterparts) to be promoted to higher echelons positions. Thus, the existence of double standards for the evaluation of men and women could produce more highly skilled female than male leaders, because they have to overcompensate the obstacles they face when climbing through the corporate ladder (Powell & Butterfield, 1994; Foschi, 2000).

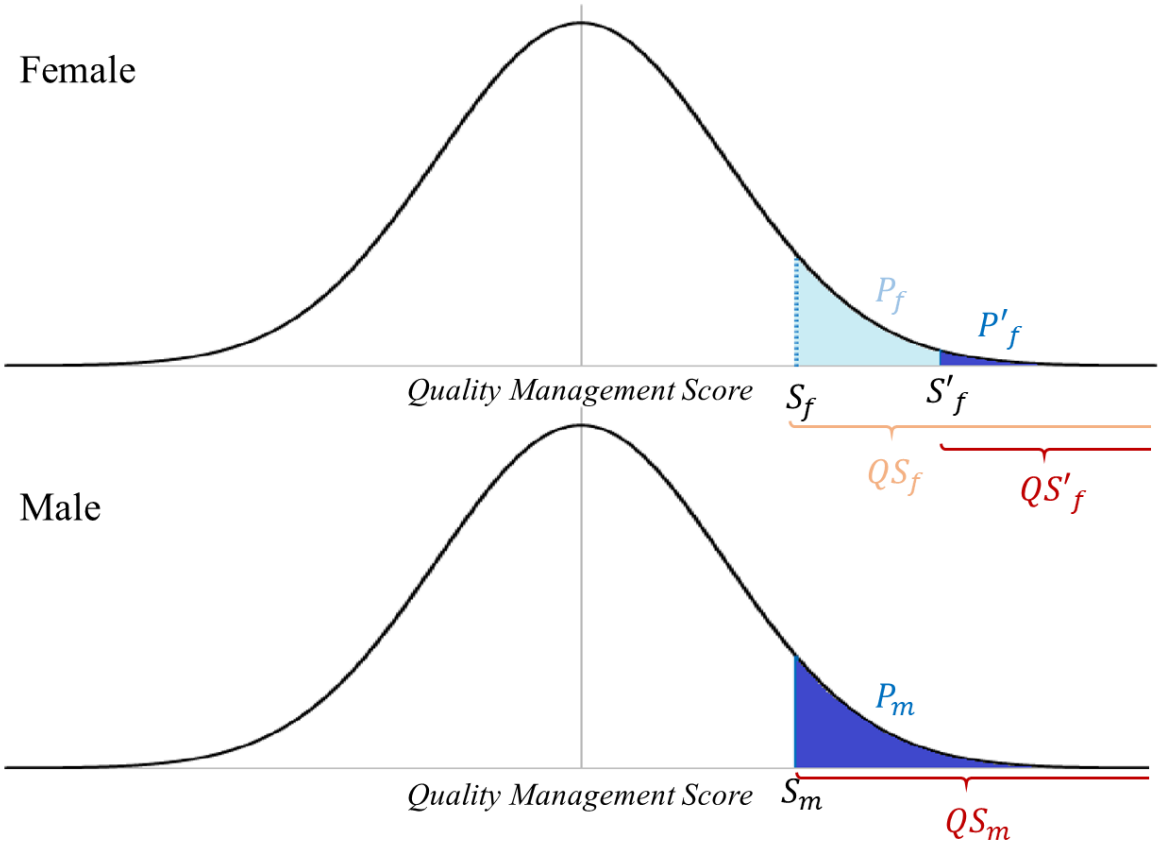
In her study, Foschi (1996) link double standards to gender prejudices (that is, to gender bias or discrimination). There are different types of discrimination. One type is discrimination due to preferences. This occurs when the company discard female minority candidates for its managerial positions even when they are available, because the decision-makers prefer forfeiting income rather than hiring women (Becker, 1957). The second kind of discrimination is statistical discrimination defined by Phelps (1972) that occurs when women are judged according to the average characteristics of their group and not on the basis of their own personal characteristics. In both kinds of discrimination, individuals consciously discriminate, however Bertrand, Chugh and Mullainathan (2005) argue that discrimination may be unintentional and outside of the discriminator's awareness so they propose implicit discrimination to define biases that they people don't think about and may never acknowledge. According to Foschi (1996) double standards could be a consequence of implicit discrimination since the actor does not have to formulate such standards explicitly in order to use them.

On the supply-side: Women Lower Self-Efficacy Perception

Figure 2 gives the schematic representation of the *Female Lower Self-Efficacy Perception Theory*. As in the previous case, bell shape curves represent the distribution of Quality Management Scores for both female (top) and male (bottom) teachers. Again, according to the theory, teachers with best quality management scores are the ones promoted to principal positions. However, in this case the standard applied for both male and female candidates are the same ($S_f = S_m$). Divergences appears because female potential candidates that could be promoted ($QS_f > S_f$) do not consider themselves to be really up to the position offered, thus the *women low self-efficacy perception* appears. Lower women´s self-efficacy perception in the workplace lead them to be tougher when considering themselves for a top-level position than their male counterparts: $S'_f > S_f = S_m$. Thus, even if there is no gender bias in the demand-side of the process, the self-imposed penalty produce the same results than a double standard:

the number of women that satisfies the condition $QS'_f > S'_f$ is much lower than that of males, where $QS_m > S_m$, thus, we will end with a lower proportion of female principals ($P'_f < P_m$), even though they should be similar ($P_f = P_m$); secondly, among those principals that have overcome the higher self-imposed standard, the average Quality Management Scores will be higher than for their male counterparts ($E[QS_f/QS_f > S'_f] > E[QS_m/QS_m > S_m]$).

Figure 2: Female Lower Self-Efficacy perception Theory



The higher women's self-imposed standard when considering themselves to occupy top-managerial positions can come from the fact that women generally judge themselves less suited for many nontraditional occupations than men (Bandura, 1982). Leadership positions are often to believe required agentic qualities (i.e., speaking assertively, competing for attention, influencing others, initiating activity directed to assigned tasks or making problem-focused suggestions) that are usually associated to men. Whereas, communal qualities (i.e., speaking

tentatively, not drawing attention to themselves, accepting other's direction, supporting and soothing others, contributing to the solution of relational and interpersonal problems) that are associated to women are usually not considered as important to success as a leader. Since agentic qualities are the typical qualities to succeed as a leader (Eagly & Carli, 2007), the misfit between the female gender role and the leadership role could raise internal barriers that would prevent more women to consider themselves ready for top managerial positions.

This subjective belief regarding one's self-assessed ability to deal with prospective situations containing many ambiguous, unpredictable, and often stressful elements is known as self-efficacy perception in the psychology literature (Bandura, 1982). Women subjective belief that they lacks skills to perform a traditionally masculinely sex typed task such as a leadership role (i.e. a lower self-efficacy perception to occupy a principal position) may lead them to self-selecting out of upper-echelon positions, if they do not believe they can effectively lead others (Dickerson & Taylor, 2000). This phenomenon has been called as 'concrete ceiling' by Chinchilla and León (2004).

Novo-Corti et al. (2018) explain the behavioral dynamic of self-exclusion of women of managerial positions through a consolidation of the gender stereotypes of women being less capable than men that are transmitted to the educational model, which perpetuates the dynamic of self-exclusion. This fact has conditioned women lack of esteem and confidence in order to run for top managerial positions as it reduces potential successful behaviors and the initiative in the decision-making process. This feeling of lower self-efficacy could evolve in women self-excluding themselves for managerial positions,

Hypotheses development

Both theories posted above are based on the hypothesis that female principals have higher quality management scores than men. Therefore our first hypothesis is:

H1: Female directors present higher management quality scores than their male counterparts.

With regard management quality, we will follow Bloom et al. (2015) that distinguish two types of management quality: talent and non-people management. Talent management includes practices such as hiring, firing, pay and promotions, meanwhile non-people management includes operations, monitoring and target setting. These authors found that people (talent) management scores were notably weaker than non-people management scores, in the eight countries analyzed. This situation leads us to think that the abilities required to manage people are different from the ones required to manage non-people practices, which advises us to deal these categories in a separated way. Therefore, we propose the following sub-hypotheses:

H1a: Female directors present higher non people management quality scores than their male counterparts.

H1b: Female directors present higher talent management quality scores than their male counterparts.

If the previous hypotheses are accepted, our second goal will be to disentangle if the cause comes from the demand (i.e. double standard) or from the supply-side (i.e. female lower self-efficacy perception). This way, we use the potential impact of different moderating factors on the relationship between principal gender and management quality. In order to do this, we draw on Terjesen et al. (2009) that, after reviewing the literature about women on leadership positions, use micro, meso and macro levels (i.e. individual, school and industry factors) to classify the theories that explain the under-representation of women in these positions. Thus, we classify the moderator factors of the impact of women principals on the quality of management school into these three groups. In this way, micro level includes principal

characteristics, meso level school characteristics, and macro level includes the degree of competition of the school (Table 1).

If the cause of both low proportion of female principals and higher management quality scores is lower women self-efficacy perception, then we could expect that the micro factors are the ones that potentially moderate the effect of gender on management quality. For example, female high tenure or background in a male-dominated field (that could mitigate gender differences in behaviors (Grogan & Shakeshaft, 2011; Gornick, 1990) could reduce gender differences in judgments about how well one can occupy a leadership position. Therefore, to take into account these possible interactions we posit the following hypothesis:

H2: Micro variables moderate the effect of the gender of the principal on the management quality scores.

On the other hand, on the demand-side, we could expect that if the cause of both low proportion of female principals and higher management quality is the double standard posed above, then differences in meso (school characteristics) and macro (industry factors) should play a role in the influence of the double standard effect. For example, technical schools could have higher double standards than regular curriculum ones that have higher accountability requirements. At the macro level, for instance, higher level of competition among schools could reduce the room that schools have to apply double standards. Therefore, we hypothesized the following:

H3: Meso variables moderate the effect of the gender of the principal on the management quality scores.

H4: Macro variables moderate the effect of the gender of the principal on the management quality scores.

As mentioned before, and taking into account the two main management quality components established by Bloom et al (2015), H2, H3 and H4 can be divided into two sub-hypotheses:

H2a: Micro variables moderate the effect of the gender of the principal on the non-people management quality scores.

H2b: Micro variables moderate the effect of the gender of the principal on the talent management quality scores.

H3a: Meso variables moderate the effect of the gender of the principal on the non-people management quality scores.

H3b: Meso variables moderate the effect of the gender of the principal on the talent management quality scores.

H4a: Macro variables moderate the effect of the gender of the principal on the non-people management scores.

H4b: Macro variables moderate the effect of the gender of the principal on the talent management scores.

3. Sample and methodology.

In order to test our previous hypotheses, we use data from the World Management Survey (Bloom, Lemos, Sadun & Van Reenen, 2015). This survey investigates the adoption of 20 basic management practices, leadership, organization, ownership and human resources related to 1,800 high schools in eight countries (UK, USA, Sweden, Canada, Germany, Italy, Brazil and India). This survey methodology has been previously employed in manufacturing (e.g. Bloom, Lemos, Sadun, Scur & Van Reenen, 2016; Bloom, Lemos, Sadun, Scur & Van Reenen, 2014 ;

Bloom & Van Reenen, 2007), retail (Lemos & Scur, 2012), health care (Bloom, Propper, Seiler & Van Reenen, 2015), and schools (Bloom, Lemos, Sadun & Van Reenen, 2015).

Data was collected through telephone interviews to school principals and in 69% of cases interviews were conducted with the presence of two interviewers. Open ended questions were marked in a scoring grid from 1 ('worst practice') to 5 ('best practice'). The management index for each school represents the average of these scores.

Although the survey covers several issues, we have focused on the management section that includes 20 questions about: operations (measures teaching methods), monitoring (measures how school performance is tracked), target setting (measures how goals are set and if they are appropriate) and people management (measures how school deals with employees).

Our base model is as follows:

$$Y_i = \alpha + \beta \cdot Female_i + \Gamma_p \cdot P_i + \Gamma_s \cdot S_i + \Gamma_C \cdot C_i + \varphi \cdot Country_i + \psi \cdot Interviewer_i + \nu_i$$

where Y_i is the management score for each principal, *Female* is a dummy variable with a value 1 if the principal is a woman and 0 otherwise, P_i is a set of the micro variables (tenure, STEMB background, personnel autonomy and academic autonomy), S_i is a set of meso variables (number of students, pupil per teacher ratio, the school ownership, type of school curriculum (regular vs. vocational) and if the school select pupils based on academics), and finally C_i represents the macro characteristic (the number of competitors). We have also controlled by country and by the interviewer.

Table 2 includes the main descriptives of all variables we use. Among principal characteristics we consider the gender, his or her tenure (number of years at the school), whether he or she has a STEMB degree (science, technology, engineering, mathematics and business degree), the

academic content autonomy and the personnel one[2]. The considered characteristics about the school are the number of pupils, the pupil per teacher ratio, the school ownership (if it is a private for-profit school, private non-for-profit[3] or a public one), the school curriculum character (i.e. if the school has a regular curriculum vs. a technical education) and student selection (if the schools makes its pupil selection based on academics). Finally, the macro level includes the number of competitors the school has[4]. Correlations between the different variables are shown in Table A1 of the annex.

The average management score used by Bloom et al. (2015) is computed as a simple average on 20 management practice questions. However, since these questions consider a wide range of aspects, we have explored if this measure could be decomposed in several components. For that reason, a Principal Component Analysis (PCA) has been conducted.

This way, the PCA was conducted on the 20 management practices questions with orthogonal rotation (varimax). The Kaiser–Meyer–Olkin measure verified the sampling adequacy for the analysis, KMO = .97, and all KMO values for individual items were > .95, which is well above the acceptable limit of .5 (Field, 2009). Bartlett's test of sphericity χ^2 (190) = 20715.19, $p < .001$, indicated that correlations between questions were sufficiently large for PCA (see Table 3). An initial analysis was run to obtain eigenvalues for each component in the data. Two

[2] For measuring the academic content autonomy, the principal is asked: "To add a new class – for example, introducing a new language such as Mandarin – what agreement would you need". To measure the personeel autonomy the question was: "To hire a full-time teacher what agreement would you need?". In both cases to measure the degree of autonomy, it is used a 1–5 scale where 1 refers to no authority to make any decision and 5 refers to complete authority to make any decision.

[3] We refer to private non for-profit schools as schools receiving at least partial funding from the government and with at least limited autonomy to follow school-specific charters in one of three areas: establishing the curriculum content, selecting teachers and admitting pupils. In our data, these are *escolas de referencia* in Brazil, separate schools in Canada, private *ersatzschulen* in Germany, private-aided schools in India, *friskolor* in Sweden, academies, foundation and voluntary-aided schools in the UK (equivalent to autonomous state schools), and charter and magnet schools in the US.

[4]The measure of competition is collected during the survey itself by asking the principal: "How many other schools offering education to 15-year olds are within a 30-minute drive from your school?"

components had eigenvalues over Kaiser's criterion of 1 and in combination explained 56.22% of the variance (see Table 4). Given the large sample size, and Kaiser's criterion on two components, this is the number of components that were retained in the final analysis.

Table 5 shows the factor loadings after rotation. The questions that cluster on the same components suggest that component 1 represents a non-people management score, and component 2 a talent management score. For this reason, we will run three different analyses: with the average management score, non-people management score, and with the talent management one.

4. Results

We have constructed regression models to explain the average management score (Table 6), as well as the two factors we have obtained in the previous section from the principal component analysis: the non-people factor (Table 7) and the talent management factor (Table 8). All models include interviewer and country fixed effects as well as the duration of the interview as controls. We have excluded also those observations where the interviewee is not the principal/head teacher/head master to have a homogeneous sample.

In all tables column 1 estimates the basic model in a multivariate framework, columns 2–12 interact Female with the moderating conditions hypothesized in Hypotheses H2 H3 and H4. Concerning principal effects, we have included the three different groups: variables related with principal characteristics (tenure, STEMB background, personnel and academic content autonomy), those related with school characteristics (number of pupils, class size, school ownership, school curriculum character, and pupil selection based on academic merit) and a macro variable (number of competitors).

For the regression on the average management score we find that STEMB background of the principal, Principal personnel autonomy, Number of pupils, Private not for profit school, and

Schools with a regular curriculum are positively associated with average management score, whereas Pupils per teacher ratio is negatively associated with this score. These results hold for all the models.

In our baseline hypothesis (Hypothesis 1), we posit that female directors produce higher management quality scores. The positive and highly significant coefficient for Female variable (in all models except column 5 specification) strongly support this hypothesis. This result indicates that women principals are associated with higher management scores, after adjusting for country, noise controls, principal and school characteristics.

Hypotheses 2, 3 and 4 further identify micro, meso and macro variables that moderates the influence of the principal gender on the management quality to try to disentangle if the cause of the gender gap on management quality scores comes from the demand or the supply side. Consistent with the prediction in Hypothesis 3, we find that the interaction of Female × Regular curriculum is also negative and significant. Besides, the interaction of Female × Pupil per teacher ratio is negative but only marginally significant. Similarly, for the hypothesis 4, the interaction of Female × Number of Competitors is negative and significant. Both results are consistent with gender differences in quality management in schools coming from double standards through the promotion process (demand-side) since they are factors depending on the school and the industry and not on the personal characteristics of the principal. By contrast, we do not find evidence supporting hypothesis 2, since neither Tenure, STEMB background nor management autonomy, all of the them at the micro level, moderate the effect of the Female variable on average quality of management.

Next, we have estimated regression models to explain the non-people factor score (Table 7). Regarding principal effects, we find that STEMB background of the principal, Number of pupils, and Schools with a regular curriculum are positively associated with the non-people

management factor, whereas Pupil per teacher ratio and Private schools (both for profit and nonprofit, being public the reference category) are negatively associated with this score.

With respect to the hypotheses testing, the positive and highly significant coefficient for Female variable (in all models except column 2, 5 and 6 specification) indicates that women principals are associated with higher non people management scores, after adjusting for country, noise controls, principal and school characteristics. This confirms hypothesis 1a.

Regarding the side where this effect may come we have studied the moderating effects stated by hypotheses 2a, 3a and 4a founding that only the interaction of Female × Regular curriculum is negative and significant, supporting hypothesis 3a and suggesting that the effect of Female principal on the non-people management score comes from the demand-side (double standard). Other conditions do not have to seem any effect on the degree the principal gender determines the non-people management quality.

Finally, we have estimated regression models on the talent management factor (Table 8). In these models, we find that STEMB background, Personal and Academic content autonomy, Number of pupils, and Private (both for profit and not for profit) schools are positively associated with average management score, whereas Pupil per teacher ratio is negatively associated with this score.

Regarding the baseline hypothesis (Hypothesis 1b), the positive and highly significant coefficient for Female variable (in all models except column 2, 4 and 5 specifications) suggest that women principals are better than their male counterparts when managing talent in schools.

Moving to the interactions that identify the micro, meso or macro levels that will have an effect on gender differences on talent management quality, the results are quite similar to those found for the regressions on the average management score. Once again, hypothesis 3b is confirmed since we find several interactions coefficients regarding meso variables that are significant.

This way, we find that the interaction of Female × Number of students is negative and significant. Similarly, the interaction of Female × Regular curriculum is also negative and significant. Both results are consistent with gender differences in quality management in schools coming from double standards through the promotion process (demand-side) since they are meso factors depending on the school and not on the personal characteristics of the principal. Finally, for the case of Hypotheses 4b, the interaction of Female × Number of Competitors is negative and highly significant, also pointing for the demand-side as the most plausible cause for gender differences in talent management.

Regarding hypothesis 2b, we only find marginally significant interaction between Female and STEMB background. In this regard, some studies have found that women with professions or careers in male-dominated areas (such as leadership positions or science) would tend to "act like a man", to be considered as one of the group (Grogan & Shakeshaft, 2011; Gornick, 1990). This could lead to soften the gender differences in perceptions and judgments regarding oneself abilities, this way women with STEMB background would exhibit a higher self- esteem and lower self-limiting behavior that the average of their gender group. Nevertheless, the marginality of the result might not provide enough evidence to support this hypothesis.

To sum up, Table 9 shows hypothesis confirmed and not confirmed by the results analysis. These results seem to point out that the gender differences found in quality management scores comes from demand-side factor at the meso (curriculum character of the school) and marco level (degree of competition) rather than to supply-side factor that operate at the micro-level.

5. Discussion and conclusions

Despite representing the majority of teachers' staff, women are still underrepresented in leadership positions on schools even although the results in this paper show that they have higher management quality scores. This could be the result of the presence of double standards

that is hindering women's ascent in school top hierarchies or that they have lower self-efficacy perceptions to perform a leadership position than men. In both cases the results would be that the proportion of women principals in school is lower than the one of their male counterparts and that those women that reach a principal position are more qualified than males as they are assessed in a stricter way either by the recruiters in their school (glass-ceiling) of by themselves (concrete-ceiling) when it comes to occupy top positions.

Therefore, the gender differences found in quality management in schools can be explained by the demand or by the supply-side of the market of school principals. In order to disentangle which side of the market offers the most plausible explanation, we use the moderator effects of micro (principal characteristics), meso (school characteristics) and macro-level (industry characteristic).

Results of the estimated models show that the most plausible explanation for the gender differences in quality management scores in schools seems to come from the demand-side, since meso (school size and regular curriculum school) and macro-factors (school degree of competence) are the ones that play a role in the moderation of the principal effects, mitigating the observed gender gap observed. This way, it seems that whereas women principal exhibit higher management quality scores than men in schools with low competition, this gap is attenuated in high competition schools (both on average management and talent management scores). This fact can be explained by the dynamics of the Becker theory of discrimination (1957) according to whom a company may discard female minority candidates for its leadership positions even when they are available because the decision-makers prefer forfeiting income rather than hiring women (what he called taste-based discrimination). This discriminatory practices of those schools that prefer not to work with women are only possible if the company (or the school in our case) is in a situation where it can forfeit incomes. Otherwise, in a high level of competition context the company cannot indulge their male preferences since higher

costs, in terms of loss of efficiency and lower profits, than those of schools that did not discriminate would lead them to not survive in the long term. This way, competition seems to be playing in favor of women principals.

Also, women principals observe a less pronounced double standard in regular curriculum schools in all the management scores. Technical schools generally have more freedom than regular curriculum ones in terms of their hiring policy and the academics contents, as they are not usually providing a certified or compulsory education (Waslander, Pater & Van der Weide, 2010). This higher freedom could sometimes be traduced in more room for indulging their preferences selecting the people they "like" instead of people they "need", that could lead to bias in the promotion process. This is also in line with findings from Foschi (1996) experiments that found that double standard is most pronounce when accountability for the assessments is low (as in the case of technical or vocational schools that are usually subject to less accountability than regular ones).

In conclusion, results show that female principals are associated with higher management quality both in talent and non-people management what we interpret as a sign of the presence of double standard of competence for female teachers when trying to reach school principal position. This double standard of competence seems to come mainly from the demand side (glass ceiling) since no micro level factor moderates the effect of principal gender on quality managements.

Our contributions are the following: 1) A theory that explains two stylized facts (previous theories that only explain the reduce proportion of women leaders ignoring the higher quality of women management they are losing half of the picture, such as, the reduced pool of female candidates); 2) This theory could be also valid for other sectors and institutions where you observe the two stylized facts and no also to women but to other discriminated groups (e.g. ethnic minorities, or LGTB+ community).

In order to disentangle the ultimate cause that explain gender differences in management quality (double standard theory associated with the glass ceiling phenomenon or women's lower self-efficacy perception liked to the concrete ceiling) future studies should get additional primary data to clarify and complement this result by implementing in-depth interviews, focus groups, etc.

Finally, the main policy implication of our results is that schools, as well as governments, should encourage measures and positive actions aimed to eliminate or, at least, reduce double-standard in female's evaluation that could result in a better use of the available talent pool for leadership positions. In this regard, it would be desirable that an "universal" standard would be applied to all performers during the promotion procedure to secure that demonstrations of ability leave no doubts about the superior quality of the chosen applicants. This, in turn, could contribute to make these institutions obtaining higher management results.

References

Bandura, A. (1982). The self and mechanisms of agency, in J. Suls (Ed.) (1982), *Psychological perspectives of the self*. Hillsdale (NJ): Lawrence Erlbaum.

Becker, G. (1957). *The Economics of Discrimination*, 2nd Ed. Chicago (IL): University of Chicago Press.

Bertrand, M., Chugh, D., & Mullainathan, S. (2005). Implicit discrimination, *American Economic Review*, *95*(2), 94-98.

Bloom, N., Lemos, R., Sadun, R., Scur, D., & Van Reenen, J. (2016). International data on measuring management practices, *The American Economic Review*, *106*(5), 152-156.

Bloom, N., Lemos, R., Sadun, R., & Van Reenen, J. (2015). Does management matter in schools? *The Economic Journal*, *125*(584), 647-674.

Bloom, N., Lemos, R., Sadun, R., Scur, D. & Van Reenen, J. (2014). The new empirical economics of management, *Journal of the European Economics Association*, *12*(4), 835-876.

Bloom, N., Propper, C., Seiler, S., & Van Reenen, J. (2015). The impact of competition on management quality: evidence from public hospitals, *The Review of Economic Studies*, *82*(2), 457-489.

Bloom, N. & Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries, *Quarterly Journal of Economics*, *122*(4), 1351-408.

Chinchilla, N. & León, C. (2004). *La ambición femenina. Cómo re-conciliar trabajo y familia*, 1st ed. Madrid: Aguilar.

Dickerson, A., & Taylor, M. A. (2000). Self-limiting behavior in women: Self-esteem and self-efficacy as predictors, *Group & Organization Management*, *25*(2), 191-210.

Eagly, A. H., & Carli, L. L. (2007). Women and the labyrinth of leadership, *Harvard Business Review*, *85*(9), 62.

Field, A. (2009). *Discovering statistics using SPSS*. Sussex: Sage publications.

Foschi, M. (1996). Double Standards in the Evaluation of Men and Women, *Social Psychology Quarterly*. *59*(3), 237-254.

Foschi, M. (2000). Double standards for competence: Theory and research, *Annual Review of Sociology*, *26*(1), 21-42.

Foschi, M., Lai, L., & Sigerson, K. (1994). Gender and Double Standards in the Assessment of Job Applicants, *Social Psychology Quarterly*, *57*(4), 326-339.

Gabaldon, P., De Anca, C., Mateos de Cabo, R., & Gimeno, R. (2016). Searching for women on boards: An analysis from the supply and demand perspective, *Corporate Governance: An International Review*, *24*(3), 371-385.

Gornick, V. (1990). *Women in science,* New York: Touchstone.

Grogan, M. & C. Shakeshaft. 2011. *Women and educational leadership,* San Francisco (CA): Jossey-Bass.

Lemos, R., & Scur, D. (2012). *Could Poor Management be Holding Back Development?* Working paper, International Growth Centre.

Mateos de Cabo, R. M., Gimeno, R., & Escot, L. (2011). Disentangling discrimination on Spanish boards of directors, *Corporate Governance: An International Review*, *19*(1), 77-95.

Novo-Corti, I., Calvo, N., & Varela-Candamio, L. (2018). Is my professional future biased for gender perceptions? A study for the Spanish case with public policy proposals on education, *Cuadernos de Gestión*, *18*(2), 59-93.

OECD (2016): *Education at a Glance 2016: OECD Indicators*, Paris: OECD Publishing. http://dx.doi.org/10.1787/eag-2016-en

OECD (2017): *Education at a Glance 2017: OECD Indicators*, Paris: OECD Publishing. http://dx.doi.org/10.1787/eag-2017-en

Phelps, E. S. (1972). The statistical theory of racism and sexism, *The American economic review*, *62*(4), 659-661.

Powell, G. N. & Butterfield, D. A. (1994). Investigating the "glass ceiling" phenomenon: An empirical study of actual promotions to top management, *Academy of Management Journal*, *37*(1), 68-86.

Schun, S. C., Bark, A. S. H., Van Quaquebeke, N., Hossiep, R., Frieg, P., & Van Dick, R. (2014). Gender differences in leadership role occupancy: The mediating role of power motivation, *Journal of Business Ethics*, *120*(3), 363-379.

Terjesen, S., Sealy, R. & Singh, V. (2009). Women Directors on Corporate Boards: A Review and Research Agenda, *Corporate Governance: An International Review*, 17, 320–337.

Waslander, S., Pater, C. & Van der Weide, M. (2010). *Markets in Education: An Analytical Review of Empirical Research on Market Mechanisms in Education*, OECD Education Working Papers, 52.

Wolfers, J. (2006). Diagnosing discrimination: Stock returns and CEO gender, *Journal of the European Economic Association*, 4(2–3), 531–541.

Table 1. Micro, meso and macro variables

Micro	Principal tenure
	Principal background (STEMB)
	Principal personnel autonomy
	Principal academic content autonomy
Meso	Number of pupils
	Pupils per teacher ratio
	School curriculum character
	Admission based on academics
	School ownership
Macro	Number of competitors

Table 2. Means and Standard Deviation of Variables.

	N	Minimum	Maximum	Mean	SD
Average Management Score	1,851	1.00	4.15	2.27	0.61
Non-people factor	1,846	-2.71	3.04	0.00	1.00
People factor	1,846	-2.78	3.44	0.00	1.00
Female principal	1,851	0.00	1.00	0.44	0.50
Tenure	1,850	1.00	52.00	6.56	6.22
Principal has STEM background (confirmed)	1,851	0.00	1.00	0.29	0.46
Principal personnel autonomy	1,851	1.00	5.00	2.76	1.67
Principal academic content autonomy	1,847	1.00	5.00	2.89	1.47
Number of pupils	1,851	10.00	5324.00	787.78	633.99
Pupil/teacher ratio	1,851	1.01	187.86	16.45	11.55
Number of competitors	1,844	0.00	1000.00	9.59	28.22
Private. for profit school	1,851	0.00	1.00	0.07	0.26
Private. not for profit school	1,851	0.00	1.00	0.21	0.41
Public school	1,851	0.00	1.00	0.72	0.45
Schools with a regular curriculum	1,851	0.00	1.00	0.92	0.26
Schools with pupil selection based on academics	1,851	0.00	1.00	0.24	0.43
N valid	1,834				

Table 3. KMO and Bartlett´s Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.972
Bartlett´s Test of Sphericity	Approx. Chi-squared	20715.2
	Df.	190
	Sig.	0.000

Table 4. Total Variance Explained

	Initial Values			Sum of square loads		
Component	Total	% variance	% cumulative	Total	% de variance	% cumulative
1	9.940	49.698	49.698	9.940	49.698	49.698
2	1.305	6.525	56.223	1.305	6.525	56.223
3	0.890	4.451	60.673			
4	0.709	3.546	64.220			
5	0.676	3.382	67.602			
6	0.634	3.168	70.770			
7	0.548	2.738	73.508			
8	0.543	2.714	76.222			
9	0.498	2.489	78.710			
10	0.492	2.459	81.169			
11	0.467	2.333	83.502			
12	0.452	2.261	85.763			
13	0.427	2.137	87.899			
14	0.411	2.055	89.955			
15	0.387	1.935	91.889			
16	0.366	1.828	93.717			
17	0.359	1.794	95.511			
18	0.337	1.684	97.195			
19	0.291	1.456	98.652			
20	0.270	1.348	100.000			

Extraction method: main components analysis.

Table 5. Rotated Component Matrix

	Components	
	1	2
Performance Review	0.779	0.208
Performance Tracking	0.773	0.193
Performance Dialogue	0.760	0.263
Target balance	0.739	0.235
Time Horizon of targets	0.705	0.320
Continuous Improvement	0.705	0.298
Target Interconnection	0.701	0.321
Data driven planning and student transition	0.693	0.301
Target Strech	0.685	0.348
Adopting educational best practices	0.673	0.316
Standardization of instructional planning processes	0.617	0.182
Personalization of instruction and learning	0.616	0.407
Consequence Management	0.610	0.396
Clarity and comparability of targets	0.562	0.465
Managing talent	0.265	0.729
Retaining Talent	0.092	0.715
Promoting high performers	0.308	0.685
Removing poor performers	0.280	0.658
Creating a distinctive employee value position	0.504	0.592
Rewarding high performers	0.308	0.549

Extraction method: analysis of main components. Rotation method: Varimax with Kaiser Standardization.
Rotation converged in 3 iterations.

Table 6. Regression on Average Management

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Female principal	0.085***	0.035	0.093***	0.081**	0.051	0.325**	0.269***	0.162***	0.084***	0.087***	0.289***	0.085***
	(0.022)	(0.041)	(0.026)	(0.040)	(0.045)	(0.154)	(0.099)	(0.042)	(0.022)	(0.025)	(0.072)	(0.024)
Principal Tenure (Ln)	-0.015	-0.031*	-0.015	-0.015	-0.015	-0.015	-0.015	-0.016	-0.015	-0.015	-0.014	-0.015
	(0.012)	(0.016)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Principal has STEM background (confirmed)	0.053**	0.054**	0.063**	0.053**	0.053**	0.053**	0.053**	0.054**	0.053**	0.053**	0.055**	0.053**
	(0.022)	(0.022)	(0.028)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Principal personnel autonomy	0.032***	0.032***	0.032***	0.031***	0.032***	0.032***	0.033***	0.032***	0.032***	0.032***	0.033***	0.032***
	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Principal academic content autonomy	0.018*	0.018*	0.018*	0.018*	0.013	0.018*	0.018*	0.019*	0.018*	0.018*	0.017*	0.018*
	(0.010)	(0.010)	(0.010)	(0.010)	(0.012)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Number of pupils (Ln)	0.102***	0.102***	0.102***	0.102***	0.102***	0.118***	0.103***	0.102***	0.102***	0.102***	0.100***	0.102***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.019)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Pupil/teacher ratio (Ln)	0.108***	0.107***	0.109***	0.108***	0.108***	0.107***	-0.078**	-0.111***	-0.108***	-0.109***	-0.108***	-0.108***
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.034)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Number of competitors (Ln)	0.022*	0.023*	0.023*	0.022*	0.022*	0.022*	0.022*	0.041***	0.022*	0.022*	0.024*	0.022*
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.015)	(0.012)	(0.012)	(0.012)	(0.012)
Private, for profit school	0.060	0.059	0.059	0.059	0.055	0.058	0.055	0.064	0.046	0.060	0.064	0.060
	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)	(0.072)	(0.052)	(0.052)	(0.052)
Private, not for profit school	0.089**	0.089**	0.088**	0.089**	0.088**	0.088**	0.088**	0.088**	0.089**	0.091**	0.091***	0.089**
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.040)	(0.035)	(0.035)
Schools with a regular curriculum	0.112***	0.110***	0.113***	0.112***	0.113***	0.110***	0.112***	0.116***	0.112***	0.112***	0.202***	0.112***
	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.051)	(0.041)
Schools with pupil selection based on academics	0.000	0.003	0.000	0.000	0.001	0.002	-0.001	-0.002	0.000	0.000	-0.002	-0.000
	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.049)
Female x Tenure		0.034										
		(0.023)										
Female x STEMB			-0.025									
			(0.046)									
Female x Personnel autonomy				0.001								
				(0.012)								
Female x Academic autonomy					0.012							
					(0.014)							
Female x number of students						-0.038						
						(0.024)						
Female x pupils per teacher							-0.069*					
							(0.037)					
Female x Number of competitors								-0.045**				
								(0.021)				
Female x Private for profit									0.023			
									(0.079)			
Female x Private non for profit										-0.005		
										(0.050)		
Female x Regular curriculum											-0.221***	
											(0.075)	
Female x Admision based on academics												0.001
												(0.055)
Constant	0.359	0.381	0.359	0.362	0.374	0.254	0.245	0.327	0.360	0.358	0.265	0.359
	(0.236)	(0.236)	(0.236)	(0.237)	(0.236)	(0.245)	(0.243)	(0.236)	(0.236)	(0.236)	(0.237)	(0.236)
Observations	1,711	1,711	1,711	1,711	1,711	1,711	1,711	1,711	1,711	1,711	1,711	1,711
Interviewer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-stat	36.94	36.39	36.32	36.30	36.33	36.40	36.44	36.48	36.31	36.30	36.64	36.30
pval	0	0	0	0	0	0	0	0	0	0	0	0
R squared	0.569	0.570	0.569	0.569	0.569	0.570	0.570	0.570	0.569	0.569	0.571	0.569

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Figures inside parentheses represent robust standard deviation of each parameter.

Table 7. Regression on non-people management factor

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Female principal	0.135***	0.065	0.117**	0.194**	0.096	0.208	0.405**	0.182**	0.145***	0.137***	0.468***	0.142***
	(0.042)	(0.078)	(0.050)	(0.076)	(0.087)	(0.297)	(0.191)	(0.081)	(0.043)	(0.047)	(0.139)	(0.046)
Principal Tenure (Ln)	-0.026	-0.047	-0.025	-0.025	-0.026	-0.026	-0.026	-0.026	-0.026	-0.026	-0.024	-0.026
	(0.023)	(0.030)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Principal has STEM background (confirmed)	0.091**	0.093**	0.069	0.092**	0.091**	0.091**	0.091**	0.092**	0.090**	0.091**	0.095**	0.092**
	(0.043)	(0.043)	(0.054)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
Principal personnel autonomy	-0.003	-0.003	-0.003	0.006	-0.003	-0.003	-0.002	-0.003	-0.003	-0.003	-0.002	-0.003
	(0.021)	(0.021)	(0.021)	(0.023)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Principal academic content autonomy	-0.008	-0.009	-0.008	-0.007	-0.015	-0.008	-0.009	-0.008	-0.007	-0.008	-0.010	-0.008
	(0.020)	(0.020)	(0.020)	(0.020)	(0.024)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Number of pupils (Ln)	0.141***	0.141***	0.141***	0.141***	0.141***	0.146***	0.142***	0.141***	0.141***	0.141***	0.138***	0.141***
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.036)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Pupil/teacher ratio (Ln)	-0.125**	-0.123**	-0.125**	-0.126**	-0.125**	-0.125**	-0.081	-0.127**	-0.126**	-0.125**	-0.124**	-0.125**
	(0.058)	(0.058)	(0.058)	(0.058)	(0.058)	(0.058)	(0.065)	(0.058)	(0.058)	(0.058)	(0.058)	(0.058)
Number of competitors (Ln)	0.041*	0.042*	0.041*	0.041*	0.041*	0.041*	0.041*	0.052*	0.041*	0.041*	0.044*	0.041*
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.029)	(0.024)	(0.024)	(0.024)	(0.024)
Private, for profit school	-0.404***	-0.406***	-0.402***	-0.391***	-0.410***	-0.404***	-0.412***	-0.401***	-0.310**	-0.405***	-0.398***	-0.403***
	(0.100)	(0.100)	(0.100)	(0.101)	(0.100)	(0.100)	(0.100)	(0.100)	(0.137)	(0.100)	(0.100)	(0.100)
Private, not for profit school	-0.180***	-0.180***	-0.178***	-0.178***	-0.181***	-0.180***	-0.181***	-0.181***	-0.181***	-0.176**	-0.177***	-0.179***
	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)	(0.077)	(0.067)	(0.067)
Schools with a regular curriculum	0.206***	0.204**	0.204**	0.208***	0.207***	0.206***	0.206***	0.208***	0.208***	0.206***	0.354***	0.207***
	(0.079)	(0.079)	(0.080)	(0.080)	(0.080)	(0.080)	(0.079)	(0.080)	(0.080)	(0.080)	(0.099)	(0.080)
Schools with pupil selection based on academics	-0.094	-0.091	-0.095	-0.096	-0.093	-0.094	-0.095	-0.095	-0.096	-0.094	-0.097	-0.082
	(0.087)	(0.087)	(0.087)	(0.087)	(0.087)	(0.087)	(0.087)	(0.087)	(0.087)	(0.087)	(0.087)	(0.094)
Female x Tenure		0.047										
		(0.044)										
Female x STEMB			0.060									
			(0.088)									
Female x Personnel autonomy				-0.022								
				(0.024)								
Female x Academic autonomy					0.014							
					(0.027)							
Female x number of students						-0.012						
						(0.046)						
Female x pupils per teacher							-0.102					
							(0.070)					
Female x Number of competitors								-0.027				
								(0.040)				
Female x Private for profit									-0.151			
									(0.152)			
Female x Private non for profit										-0.009		
										(0.097)		
Female x Regular curriculum											-0.362**	
											(0.144)	
Female x Admision based on academics												-0.039
												(0.106)
Constant	-2.118***	-2.088***	-2.118***	-2.162***	-2.101***	-2.150***	-2.287***	-2.138***	-2.123***	-2.119***	-2.273***	-2.120***
	(0.453)	(0.454)	(0.453)	(0.456)	(0.454)	(0.471)	(0.468)	(0.454)	(0.453)	(0.454)	(0.457)	(0.453)
Observations	1,707	1,707	1,707	1,707	1,707	1,707	1,707	1,707	1,707	1,707	1,707	1,707
Interviewer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-stat	19.65	19.35	19.33	19.34	19.32	19.32	19.38	19.33	19.34	19.32	19.49	19.32
pval	0	0	0	0	0	0	0	0	0	0	0	0
R squared	0.413	0.414	0.413	0.413	0.413	0.413	0.414	0.413	0.414	0.413	0.415	0.413

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Figures inside parentheses represent robust standard deviation of each parameter.

Table 8. Regression on talent management factor

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Female principal	0.058***	0.008	0.081***	0.016	0.012	0.375**	0.184*	0.161***	0.050**	0.058**	0.210***	0.054**
	(0.022)	(0.041)	(0.026)	(0.040)	(0.046)	(0.156)	(0.100)	(0.042)	(0.023)	(0.025)	(0.073)	(0.024)
Principal Tenure (Ln)	-0.008	-0.023	-0.008	-0.008	-0.008	-0.008	-0.008	-0.009	-0.008	-0.008	-0.007	-0.008
	(0.012)	(0.016)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Principal has STEM background (confirmed)	0.044**	0.045**	0.074**	0.044*	0.044*	0.044*	0.044*	0.046**	0.045**	0.044**	0.046**	0.044*
	(0.023)	(0.023)	(0.029)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Principal personnel autonomy	0.050***	0.050***	0.051***	0.044***	0.051***	0.051***	0.051***	0.050***	0.050***	0.050***	0.051***	0.050***
	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Principal academic content autonomy	0.034***	0.033***	0.033***	0.033***	0.026**	0.033***	0.033***	0.035***	0.032***	0.034***	0.033***	0.034***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.012)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Number of pupils (Ln)	0.092***	0.092***	0.092***	0.092***	0.093***	0.114***	0.093***	0.093***	0.092***	0.092***	0.091***	0.092***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.019)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Pupil/teacher ratio (Ln)	-0.108***	-0.107***	-0.108***	-0.107***	-0.108***	-0.107***	-0.087**	-0.112***	-0.107***	-0.108***	-0.108***	-0.108***
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.034)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Number of competitors (Ln)	0.018	0.019	0.019	0.018	0.018	0.018	0.018	0.043***	0.018	0.018	0.019	0.018
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.015)	(0.012)	(0.012)	(0.012)	(0.012)
Private, for profit school	0.327***	0.325***	0.325***	0.317***	0.319***	0.324***	0.323***	0.332***	0.248***	0.327***	0.329***	0.326***
	(0.052)	(0.052)	(0.052)	(0.053)	(0.053)	(0.052)	(0.052)	(0.052)	(0.072)	(0.053)	(0.052)	(0.052)
Private, not for profit school	0.250***	0.250***	0.248***	0.248***	0.249***	0.249***	0.250***	0.248***	0.251***	0.249***	0.251***	0.250***
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.040)	(0.035)	(0.035)
Schools with a regular curriculum	0.065	0.064	0.068	0.064	0.067	0.063	0.066	0.070*	0.064	0.065	0.133**	0.065
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.052)	(0.042)
Schools with pupil selection based on academics	0.037	0.040	0.038	0.038	0.038	0.039	0.037	0.035	0.039	0.037	0.036	0.031
	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)	(0.045)	(0.046)	(0.045)	(0.046)	(0.046)	(0.045)	(0.049)
Female x Tenure		0.034										
		(0.023)										
Female x STEMB			-0.077*									
			(0.046)									
Female x Personnel autonomy				0.016								
				(0.012)								
Female x Academic autonomy					0.016							
					(0.014)							
Female x number of students						-0.050**						
						(0.024)						
Female x pupils per teacher							-0.047					
							(0.037)					
Female x Number of competitors								-0.061***				
								(0.021)				
Female x Private for profit									0.126			
									(0.080)			
Female x Private non for profit										0.002		
										(0.051)		
Female x Regular curriculum											-0.166**	
											(0.076)	
Female x Admision based on academics												0.020
												(0.056)
Constant	0.140	0.161	0.140	0.171	0.159	0.001	0.061	0.095	0.143	0.140	0.069	0.141
	(0.238)	(0.238)	(0.238)	(0.239)	(0.239)	(0.247)	(0.246)	(0.238)	(0.238)	(0.238)	(0.240)	(0.238)
Observations	1,711	1,711	1,711	1,711	1,711	1,711	1,711	1,711	1,711	1,711	1,711	1,711
Interviewer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-stat	37.13	36.57	36.59	36.55	36.54	36.65	36.55	36.80	36.58	36.49	36.67	36.49
pval	0	0	0	0	0	0	0	0	0	0	0	0
R squared	0.570	0.571	0.571	0.571	0.571	0.571	0.571	0.572	0.571	0.570	0.571	0.570

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Figures inside parentheses represent robust standard deviation of each parameter.

Table 9. Hypotheses and results

Hypothesis		Result	Consequence
H1	Female directors have a higher management quality than their male counterparts	Confirmed	Double Standard + Lower Self-Efficacy Perception
H1a	Female directors have a higher non people management score than their male counterparts	Confirmed	Double Standard + Lower Self-Efficacy Perception
H1b	Female directors have a higher talent management score than their male counterparts	Confirmed	Double Standard + Lower Self-Efficacy Perception
H2	Micro characteristics moderate the effect of the gender of the principal on the management quality	Not confirmed	
H2a	Micro characteristics moderate the effect of the gender of the principal on the non-people management score	Not confirmed	
H2b	Micro characteristics moderate the effect of the gender of the principal on the talent management score	Not confirmed	
H3	Meso characteristics moderate the effect of the gender of the principal on the management quality	Confirmed	Demand-Side cause (potential double standard)
H3a	Meso characteristics moderate the effect of the gender of the principal on the non-people management score	Confirmed	Demand-Side cause (potential double standard)
H3b	Meso characteristics moderate the effect of the gender of the principal on the talent management score	Confirmed	Demand-Side cause (potential double standard)
H4	Macro characteristics moderate the effect of the gender of the principal on the management quality	Confirmed	Demand-Side cause (potential double standard)
H4a	Macro characteristics moderate the effect of the gender of the principal on the non-people management score	Not confirmed	
H4b	Macro characteristics moderate the effect of the gender of the principal on the talent management score	Confirmed	Demand-Side cause (potential double standard)