Testing for Pay and Promotion Bias in an International Organization

Deon Filmer, Elizabeth King and Dominique van de Walle*

Development Research Group
World Bank

January 2003

Abstract: International organizations pursue multiple objectives in hiring policies — including cultural diversity, reducing costs and avoiding discrimination — among which there can be sharp trade-offs. The paper studies how these trade-offs are resolved in the World Bank's hiring processes. It estimates that half of salary and grade differentials between men and women and staff from high- and low-income countries are attributable to differences in productive characteristics. Alternative explanations for the remainder are explored, including omitted variable bias, quotas and discrimination. It is argued that the first two are not compelling explanations. Discrimination probably exists, though less than would be implied by a cost minimizing hiring policy.

JEL classification: J16; J71
Keywords: Male-female differentials; Discrimination; Internal labor markets

approximate word count: 7490 (without tables); 8990 (with tables)

* The authors thank Rajeev Jain for assistance with the database and the Bank's Human Resources Services Department for its staff files. They are grateful to Margaret Grosh, Berk Ozler, Martin Ravallion and other colleagues and seminar participants for help and excellent comments on earlier versions. The views expressed are those of the authors and do not reflect those of the World Bank's Executive Directors or member countries. EM addresses for correspondence: dfilmer@worldbank.org , eking@worldbank.org and dvandewalle@worldbank.org

1. Introduction

Women are generally remunerated and promoted less than men for their work (Blau and Kahn 1996, 2000). The reasons are subject to much debate as indicated by a spate of recent theoretical and empirical contributions.[1] With controls for occupational segregation — allowing for the fact that women are more likely to be in dead-end jobs with fewer promotion possibilities — and for gender differences in productivity-related characteristics, the wage differentials decline (Lazear and Rosen 1990, Gunderson 1989, Groot and Maassen van den Brink 1996).

Interpreting the remaining male-female earnings gap as evidence of discrimination has been plagued by worries of omitted variable bias. Rather than being due to discrimination, wage disparities may simply reflect systematic gender-related differences in unobserved characteristics.[2] Data are never perfect and the possibility that there are variables that affect productivity and are not controlled for cannot be dismissed lightly.

This paper studies pay and grade differentials by gender and nationality in the internal labor market of an international institution — the World Bank. The Bank is the largest international development agency, part of the United Nations family, with just under 8,000 employees in late 1997 of which about half were professional level staff. Unusually for large organizations, the Bank made available to us its virtually complete personnel data files for 1997.

Large international organizations such as the Bank pursue multiple objectives in hiring policies, including cultural diversity, reducing costs and avoiding discrimination. There can be sharp trade-offs between these objectives. Diversity is enhanced by recruiting from an international labor market. But

---

[1] For example, Juhn, Murphy and Pierce 1991; Booth, Francesconi and Frank 1998; Groot and Maassen van den Brink 1996; Jones and Makepeace 1996; and Kunze 2000.

[2] Recent discussions in the literature include Heckman (1998), Jones and Makepeace (1996), and Blau and Beller (1988).

international organizations face unusually large differences in reservation wages for staff capable of doing the same work given the inter-country differences in wages. One way to reduce costs would be to pay employees their reservation wages, implying unequal pay for equal work, or discrimination. The paper studies how these trade-offs are resolved in the Bank's hiring processes.

The main aspect of the Bank's experience that is of wider interest is it's unusual two-stream entry system, a recruitment approach that could embody differing possibilities for discrimination. Staff are hired either through the Young Professionals (YP) program — which enforces rigorous and closely vetted selection from a large applicant pool — or through a much more loosely controlled general hiring process that may consider only a few candidates for each position. This dual hiring structure allows us to implement a novel test for omitted variable bias. Under the plausible identifying assumption that the YP recruitment process is non-discriminatory, it reveals what part of unexplained earnings disparities in the entire staff body is attributable to omitted variable bias, and by extension, discrimination.

Below we set the stage with a discussion of the Bank's hiring policies, its objectives as an employer and basic descriptive statistics on salaries and grades across staff by gender and nationality groups. Section 3 presents the empirical models we use to explain differentials, while section 4 discusses the data. The following section describes the decompositions of pay and grade differentials — including at entry, current and in growth over time — across staff groups. Section 6 explores possible explanations for our results, while section 7 presents our tests for omitted variable bias using the existence of the YP program. A final section concludes.

2. Hiring Policies in the World Bank and Basic Salary and Grade Statistics

Like other companies and organizations, the Bank wants to cut costs where possible — but it also pursues a number of other, potentially conflicting, objectives as an employer.

i) A preference for cultural diversity. The Bank's charter has long championed a composition of employees that broadly reflects the composition of country membership.[3] In more recent times, this preference has expanded to include sensitivity to the gender balance across nationalities.

ii) An international labor market. Most firms hire locally or, at best, nationally. Due in part to its aim for diversity, the Bank faces a much larger labor market in which supply prices vary enormously. Staff are citizens of so-called 'Part I' member countries — consisting primarily of OECD and "developed" countries that cannot borrow from the Bank — or of so-called 'Part II' member countries — consisting of "developing" and "transition" countries that are entitled to borrow. The labor pool faced by the Bank is therefore one aggregated across diverse and only weakly integrated national labor markets across the world. Wages for the same labor differ among these national labor markets. So the Bank faces unusually large differences in reservation wages for staff capable of performing the same duties.[4] Yet the Bank, like other employers, seeks to reduce its labor costs. In principle, one way to do this would be to pay potential employees their reservation wage. This could well result in substantial discrimination, defined as different wages for workers with the same productivity.

iii) A desire to avoid discrimination. The Bank, in common with many employers, prefers to avoid discrimination. Aside from concerns for equity, it would not be politically feasible for this organization to minimize labor costs at the risk of being adjudged a

[3] Assuming diversity can be measured through the mix of nationalities, the Bank has done well. As of May 1997, 129 nationalities were represented at professional levels, and 177, over all staff.

[4] Reasons besides labor market segmentation are associated with restricted labor mobility internationally. Some potential hires are educated elites, desperate to flee bad economic or political situations, or simply to improve their work environments and bring up their children with more opportunities. To ensure employment, they are willing to accept lower pay than counterparts already settled in the U.S. Others have reduced bargaining power because visas need renewing to stay in the U.S. As is often noted, married women have lower bargaining power as they frequently face lower mobility and choose work conditional on a husband's work location (Hersch and Viscusi 1996). U.S. men (say) who face none of these constraints are likely to have higher bargaining power.

discriminatory employer and of not complying with U.S. law. Such (self-imposed or external) pressure means that the Bank must decide how to hire and compensate workers as it faces international labor markets with widely differing salary levels and skewed salary distributions for similar workers. The Bank's personnel department takes an active role — and one that has become more aggressive in recent years — in trying to ensure that the Bank meets its mandate of achieving staff diversity, while also reaching the goal of equal pay for equal work. A conscious effort has been made through greater centralization of recruitment and close monitoring to reduce horizontal inequity.

iv) A dual recruitment structure. There are two main ways in which staff are recruited to the Bank. Around 20 percent of current staff started out in the Young Professionals Program. The YP program sets an age limit and starts all recruits at the same initial level. It is designed to recruit a young cadre of professionals who have little work experience and recent academic qualifications. In general, YP hires undergo more uniform and rigorous screening, interview, and selection processes than other recruits. The applicant pool is also far larger relative to the number of positions being filled than for other job openings. For these reasons, YPs can be expected to be a relatively homogeneous group in terms of characteristics such as technical ability and communications skills. Non-YP recruitments follow less formal routes that tend also to be less controlled by Personel or by concerns for the greater good of the institution. Some are hired through an open recruitment process that elicits applications from around the world, while others are directly identified and hired by managers. The differences between the YP and non-YP recruitment processes provides an interesting test of discrimination.

Without the third feature of the Bank's hiring structure, and possibly also the fourth, it would not be surprising to find large disparities in salaries between men and women and by country of origin, that

cannot be attributed to differences in the characteristics of employees. The main question addressed in this paper is how far the third and fourth features are able to reduce such disparities.

Table 1 presents descriptive statistics on relative salaries for all staff — and for YP and non-YP employees separately — in professional grades (levels 22-30). (We describe the data source in Section 4.) Staff are classified by gender and by whether they come from a Part I or Part II member country. Salaries are indexed to be 100 for Part I men.[5] Focusing first on the aggregate sample, we see considerable differences in average salaries across the four professional staff groups. Average salary during the year of the data (henceforth "current" salary) for Part II men is 95 percent of that for Part I men; for Part I women, it is 87 percent; and for Part II women, 82 percent. The same rankings are revealed for pay at entry, but with even larger quantitative differences. Some catch-up in relative terms is thus indicated. This especially favors Part II relative to Part I men. Annual salary growth rates are, on average, a little higher for women (3.8 percent versus 3.2 and 3.3 percent for Part I and Part II men, respectively). Mean tenure is slightly lower for women. In May 1997, Part I men accounted for 45 percent of all staff; Part II men, for 29 percent, Part I women, 18 percent, and Part II women, 8 percent.

For both YPs and non-YPs, current salaries follow a pattern similar to that found for the full sample. But this is not true of salaries at entry. Interestingly, YP entry salaries are highest for Part II women, followed by Part I women, Part I men and lastly, Part II men. But the differences are small — the widest (between Part II women and men) is of 7.2 percentage points. Although the starting YP salary range is narrower than for others, entry salaries are also determined based on education, work experience and other factors. In contrast, non-YP starting salaries reveal much greater absolute differences — the

[5] Salaries are net of taxes and expressed in 1990 US dollars. Under the international tax treaties that pertain to Bank employment, non-U.S. citizens are exempt from U.S. income taxes. To maintain comparability among groups, all regular staff are paid salaries net of taxes, with U.S. citizens being given a separate tax allowance.

largest, between Part I men and Part II women, is of 27.3 percentage points — and follow the same pattern as for current salaries. Mean annual pay growth is higher for YP men than for YP women, while the reverse is found for non-YP staff.

A pattern similar to that for salaries is found for the distribution of staff groups across grades (Filmer et al. 1998). Women and Part II staff are frequently recruited at lower grades than men and Part I staff, respectively. The current grade distribution of men is skewed towards the higher grades relative to that for women, even though average employment duration tends to be similar across groups. Male staff recruited as YPs also tend to be in higher grades.

These disparities in pay and grade do not control for differences in observed individual characteristics, such as length of Bank employment, highest education degree, and area of specialization. If individual characteristics that matter to productivity systematically differ across groups, then that could explain the differences in salary and grade. But, if not, we may in fact be seeing the effect of underlying differences in the returns to the characteristics of the four groups.

3. Empirical Models

To measure salary differentials we use a standard approach in labor economics known as the Blinder-Oaxaca decomposition (Blinder 1973, Oaxaca 1973).[6] The reduced-form model for the log of current salary (W_{ij}) for the i^{th} staff member in the j^{th} group can be written as:

$$\ln W_{ij} = X_{ij}\beta_j + \varepsilon_{ij} \qquad (1)$$

where X_{ij} is a vector of individual characteristics (including the constant term), with corresponding parameters β_j. ε_{ij} is a zero mean error term uncorrelated with X_{ij}. The estimated coefficients are used

to calculate predicted salaries for the four groups as well as predicted salaries based on the measured characteristics of one group and the parameters of another and vice versa. These predicted values are next used to decompose the mean salary differentials across groups into two component parts. For example, the salary differential between Part I men (subscripted by *m1*) and women (*f1*) is expressed as:

$$\ln W^*_{m1} - \ln W^*_{f1} = \beta_{m1}(X^*_{m1} - X^*_{f1}) + X^*_{f1}(\beta_{m1} - \beta_{f1}) \qquad (2)$$

[Total difference] [Characteristics] [Structure]

where the $\ln W_j^*$ is the predicted mean (log) current salary of the respective groups. This is simply an identity derived from the original regression model. The first component is the differential attributable to differences in observed characteristics. The second is that attributable to between-group differences in the returns to given individual characteristics; it is in effect, the extra income Part I women would expect to get if Part I men's parameters were used to value their characteristics.[7] The latter is usually referred to as the difference due to "structure".

If the underlying model is correctly specified, then structure is interpretable as discrimination relative to what Part I men are currently paid. If incorrectly specified — for example if there are omitted variables correlated with gender — then the structure component could also pick up the effects of misspecification. To the extent that there are important omitted variables, the coefficient estimates will be biased. If the biases are small or about equal between staff groups, this would not be problematic as they would cancel out. In general, as will be evident below, our models predict current salaries and

---

[6] A previous study of the Bank by Oaxaca and Ransom (1993) also used it to explore current salary differentials and promotion for staff on board in 1988 and 1992.

[7] We use Part I men as the reference wage structure to measure discrimination. As noted by Neumark (1988), the appropriate reference depends on the underlying nature of discrimination. What the wage structure would be like in a different, no-discrimination, situation is of course not known. However, given that Part I men have historically dominated Bank staff and remain the largest staff group, this reference may be the most defensible. It is the reference group used throughout the analysis. Filmer et al. (1998) test sensitivity to this assumption using the coefficients from a pooled regression model as the reference wage structure, as suggested by Neumark (1988) and Oaxaca and Ransom (1994). This made negligible difference to our results.

grade quite well, though biases in the underlying parameter estimates of the decompositions due to omitted variables or endogeneity can never be ruled out conclusively.

We also examine differences in initial salary and in growth in pay since entry. The average annual rate of salary growth for individual *i* in group *j* is

$$r_{ij} \equiv (W_{ij} / W_{ij0})^{1/t_{ij}} - 1 \qquad (3)$$

with tenure t_{ij}, current salary W_{ij}, initial salary W_{ij0}, all indexed to the individual and group. The model for annual salary growth is then:

$$r_{ij} = X_{ij}\delta_j + \upsilon_{ij} \qquad (4)$$

As before, X_{ij} is a vector of characteristics thought to affect salary at the time of entry and over time, and υ_{ij} is a zero mean innovation error. The identity in (3) allows us to back out entry salary predictions and their decompositions.[8] The above analysis is repeated for current grade, grade at entry, and rate of promotion.

4. Data

We use the Bank's Personel Department's data from five years ago, the most recent at our disposal. These data cover *all* regular and fixed-term staff in the professional (or non-support) grades who were on active duty as of May 23, 1997 in the International Bank for Reconstruction and Development part of the World Bank group.[9] The computerized data do not include complete pay and promotion histories, but do include both entry and current levels, so that average annual growth in

[8] We could run an initial salary regression but pre-entry and at time of entry information is sparse.

[9] We exclude staff working for the finance (IFC) and investment (MIGA) parts of the World Bank group. Staff in grades 22 to 30 are included. Some entered at lower grades. In particular, Young Professionals enter at grade 21. There were 3,003 regular staff in grades 22-30 when the data were extracted from the records.

salaries and grades over time can be measured.

The X vector of individual characteristics comprises dummy variables for the year of entry to Bank employment (this captures tenure as well as circumstances specific to the year of appointment), age at entry (in linear and squared forms), dummies for maximum completed level of education (doctorate, masters and bachelors degrees), dummies that interact the education degree with its country of location and discipline, dummies for a number of the main universities where economics PhDs were obtained, years of work experience prior to joining the Bank, current marital status, dummy variables for current nationality, area of specialization, and current affiliation within the Bank (operations; units related to research, knowledge management and dissemination; or others — primarily support units such as accounting, personnel, and finance).[10] The variables are described more fully in Filmer et al. (1998).

Several factors that may influence salary and promotion are not available to us. For example, while data on pay, grade, and departmental affiliation are routinely updated, information on education completed after entry, language and other training, are not. Further, qualitative information about performance and annual performance merit ratings are strictly confidential and not available to us.

Omitted variables also include some pre-entry or time-of-entry information — such as nationality and marital status at entry, and first hiring unit within the Bank. For some staff these attributes will have changed. The data on pre-Bank work experience are also limited. We include years of previous work experience in our regressions, but we do not have data on salary or location for previous jobs.[11]

---

[10] Our only indication of pre-Bank work experience is in the time elapsed between completion of an employee's last degree and their start date. Some finalize degrees after joining, so this variable is sometimes negative. An additional dummy variable captures this. We single out certain nationalities that consistently appear to be different in their effect and each make up at least one percent of total staff.

[11] Oaxaca and Ransom (1993), who had data on the location of previous work experience, found only a marginally significant salary effect of experience outside the U.S. for Part I men (1.8 percent) and of multi-country experience only for Part II men (2.5 percent). This suggests that previous experience does not count for much in initial salaries.

Salary growth is meant to reflect performance over time — although the correlation is far from perfect for a number of reasons. However, annual increases associated with performance ratings are well-structured and have an institution-wide predetermined ceiling each year. Variation in performance is presumably predicted in part by characteristics observed at entry and in part by those unobserved (not closely correlated with observed characteristics) that become apparent to managers with time. Performance is also affected by internal job choices made by staff. The willingness and ability to travel for many months each year, to learn a new language, or to manage a difficult or politically sensitive task are examples of these choices. It is also affected by the willingness of managers to give more responsibility and more rewarding assignments to certain staff. Lastly, the proportion of time spent in operational (lending) versus non-operational (e.g., personnel, accounting, research) jobs can affect one's long-term career. We do not have a history on these choices, but we do include variables on current job affiliation within the organization.

Finally, in common with much of the firm-specific literature, we are unable to deal with possible sample selection bias due to attrition, because of lack of data on staff who left the Bank. To assess the extent of bias in pay and grade, full histories for currently employed staff are not sufficient as they omit those who left. We do not think that this is a major problem however, because attrition tends to be low: around 3 percent per year. Nor can we say anything about discrimination at the point of hiring decisions.

5. Decompositions of Salary and Grade for the Entire Staff Sample

Various decompositions for the entire sample are given in Tables 2 to 6. The total percent difference in predicted salary or grade is given in the first column of each table, followed by the component explained by average individual characteristics and that attributable to structure — or the same characteristics being valued differently across groups. As mentioned, we use Part I men as the

reference wage and grade structure against which discrimination is measured. The underlying regressions are given in Filmer et al. (1998), which also gives standard errors for the decompositions.[12]

Among Part I staff, 61 percent of the 14.2-percent difference in pay between men and women is due to women having different characteristics to men, while the remainder is attributable to structure (see Table 2). Among Part II staff, a smaller percentage (52 percent) of the gender wage gap (14.3 percent) is due to differences in characteristics rather than to structure. Among men, 79 percent of the 5.0-percent pay gap between Part I and Part II staff is due to structure; among women, all of the pay gap, as large as among men, is due to structure rather than characteristics. The largest between-group difference is between Part II women and Part I men (19.3 percent) with more than half of it due to structure. Between-group differences in salaries at entry are larger than the above differences, but follow the same overall pattern (Table 3). About one-half of the 17 percent difference between the initial salaries of Part I men and women are attributable to structure. Similarly, had individual characteristics been valued in the same way for Part II women as for Part II men, the former's starting salaries would have been, on average, 7.7 percent higher. The highest total, and unexplained, differences at 25 and 15 percent, respectively, are again between Part I men and Part II women. Although those with lower entry salaries see some catching up, between-group differences in salary growth are meager so that differences in entry salaries continue to dominate the patterns in current salaries, even in a sample with a mean tenure of 12 years.[13]

Similar to salaries, the greatest difference in grade is observed between Part I men and Part II

[12] As a rule, the regressions for men have higher explanatory power than those for women. F-tests strongly reject the hypotheses that the regression coefficients are the same across the four staff groups.

[13] Part or all of this catch-up is built into the structure of annual salary increases which stipulates that among staff performing at the same level of satisfaction, those in lower salary levels get a slightly higher merit pay increase.

women, with the gap being close to a full grade (Table 4).[14] About half of this difference is due to structure, meaning that Part II women with identical observed characteristics to Part I men are, on average, at approximately half a grade lower. This unexplained gap in grade is even larger — over one full grade — at entry (Table 5). In general, women enter at significantly lower grades than men, as do Part II staff relative to Part I staff. Focusing on differences attributable to structure, women enter at between 0.7 (Part I) and 0.9 (Part II) of a grade lower than do men. Part II men enter at 0.4 of a grade lower than do Part I men, and Part II women 0.5 of a grade lower than Part I women. On average, men can expect a promotion of about 0.19 of a grade each year; Part I women, 0.23; and Part II women, 0.25 (Table 6).[15] Women have been catching up, as indicated by results regarding the contribution of structure — for example, Part II women have been catching up at the rate of 0.04 of a grade relative to Part I men; but these gains are too miniscule to significantly compensate for differences in grades at entry attributable to different valuations of identical included characteristics.

6. Explanations and Biases

The above analysis indicates that a sizeable portion of differences in average salary and grade between men and women and between Part I and Part II staff cannot be explained by differing characteristics as observed in our data set. Although some catching up occurs after entry, salary differences at the time of hiring largely dictate inter-group differences in current pay. And current salary differences are absolutely larger except between Part I and Part II men. The pattern in grades is similar, with disparities in grade assignment at entry largely explaining variation in current grade. Well over one-

[14] The explanatory power of the regressions on grade varies from 30 percent for Part I women to 43 percent for Part II men (Filmer et al. 1998).

[15] We find considerable variance in the promotion models' explanatory power ranging from 16 to 49 percent of total variation explained, for Part I men and Part I women, respectively.

half of a grade difference between men and women at entry is not explained by differences in their mean observed characteristics.[16]

What explains the differences in salaries and grades across staff groups that are not explained by differences in observable characteristics? There are three explanations to consider. The first is that controlling for observed individual characteristics, alternative labor markets for women and Part II staff relative to Part I men are, on average, lower-paying, and this is partially reflected in the Bank's remuneration offers. The ability to pay lower wages is reinforced by a lack of information on the part of new hires — particularly ones from very different labor markets — and cultural factors. For example, there are many countries where salaries are fixed by government, and few countries where wage bargaining is as culturally accepted, and expected, as in the U.S. However, if the Bank is paying some employees less because they are willing to work for less even though they are equally productive, then that is inconsistent with the stated institutional goal of equal pay for equal work.

A second explanation rests on bias on the part of individual hiring managers, reinforced again by imperfect information. The Bank's less formal, non YP recruitment process allows line managers to identify and champion individuals of their choice. Discrimination may thus be introduced as managers follow their preferences for working with, and better rewarding, individuals for whom they have an affinity through being, for example, from the same part of the world and/or the same gender.

A third possible explanation does not involve discrimination. Our unexplained differences in salary, grade and promotion may simply reflect omitted variables, provided that there is a reason for the bias to be consistently stronger in one group (say Part I males) than in another. For example, the effects of education and previous experience may be influenced by the quality of education or individual ability

[16] Qualitatively similar results are obtained when we use pooled model parameters as the reference weights rather than Part I men's wage structure parameters (see Filmer et al., 1998).

that are omitted from our model but were observed at time of hire, promotion or salary raise.[17] The model includes a number of variables that capture aspects of education quality and individual aptitude. But other aspects of education quality, potential earnings in alternative jobs, and communication skills in various languages, are missing and may well be correlated with the included variables. If for some reason, the distribution of omitted characteristics is such that Part I men have higher latent productivity, then the returns to their education (for example) would appear to be higher than for other staff groups or (equivalently) they will have higher salaries at the same levels of observed characteristics. This would not, however, be discrimination.[18]

Binding quotas could strengthen the concern about omitted variable bias. In aiming for a staff mix representing the different cultural backgrounds of member countries, over the past decade the Bank has enforced restrictions on the recruitment of U.S. men and more aggressive recruitment of women, especially from Part II countries.[19] Quotas change the attributes of recruits at the time of hiring and this will be reflected in salaries. Some of those attributes are observed by us, others are not but may have been observed by the hiring manager or the personnel department. Explaining differences due to structure in the face of a quota rests on there being omitted characteristics in our model that systematically apply to those hired under quotas, but not to others. It could be argued that the "quota" on U.S. males has resulted in more selective hiring of U.S. men, and an increase in the quality of recent U.S. male hires, assuming no change in the applicant pool. By contrast, the policy encouraging recruitment of Part II female nationals might imply a less selective hiring process and thus a lower

[17] Including measures of education quality in earnings functions has been found to reduce the returns usually attributed to years of education (Behrman and Birdsall, 1983). Similarly, ability results in a downward adjustment on the return to education (Griliches and Mason, 1972).

[18] That is, it would not be due to "current" discrimination, although it could well reflect the cumulative effects of past societal discrimination.

[19] U.S. men comprise 16 percent of all staff in grades 22-30.

quality of Part II women hires, holding the applicant pool constant.

The quality of recent recruits also depends on the effect of these hiring policies on the quality of the applicant pools.[20] If the quota on U.S. men is widely known, it may discourage applications, reduce the overall quality of applicants and hence result in lower quality of recent U.S. male hires relative to earlier recruits. This seems unlikely, though it is theoretically possible. Similarly, the more aggressive recruitment of women may result in a larger pool of applicants from many more countries, and lead to higher quality of recent women recruits relative to earlier cohorts.

In the next section, we further explore which of these explanations is the most plausible. Since the patterns in current salaries and grades are dominated by those established at entry, the discussion focuses primarily on entry salaries and grades. We compare YPs and non-YPs to learn more about the impact of the hiring process on pay and grade, using the same empirical models as above.

7. Testing for Omitted Variable Bias: YPs Versus Non-YPs

The Bank's dual hiring process provides a unique test for the presence of bias in pay and grade. It can be argued that the Bank's YP hiring process is non-discriminatory — in that it follows a systematic and controlled vetting and checking process that successfully eliminates bias across a set of individual characteristics observable to the YP program administrators but not fully observable to us. Thus, estimating the model on the YP sub-sample provides a natural test of whether observed differentials in pay and grade are due to misspecification, whereby we have failed to control for the right set of variables. Large differences attributable to structure rather than to characteristics would then indicate that key explanatory variables are missing in the underlying model.

The model's explanatory power was uniformly high across the YP regressions, but less so for the

non-YP groups, particularly women. (The regressions are in Filmer et al. 1998.) Tables 7 to 11 give the different resulting decompositions pertaining to the YP and non-YP subsamples.

Differences in the average current salaries of staff who joined as YPs follow a pattern similar to that found for the full sample (Table 7). YP women from Part I and Part II countries are predicted to be paid 17 and 20 percent less, respectively, than Part I YP men. Around 13 to 15 percentage points of this difference are explained by differences in characteristics, leaving about one-third due to structure. There is little total difference, and no difference due to structure, between Part I and Part II YP men. Among non-YPs, differences that cannot be attributed to characteristics are much higher: 42 percent of the 13.7 percent total difference between Part I women and men; 82 percent of the 6.3 percent difference between Part I and Part II men; and 55 percent of the19.4 percent total difference between Part I men and Part II women.

Patterns in salaries at entry are strikingly different. As noted earlier, starting salaries among YPs are highest for Part II women, followed by Part I women, Part I men and lastly, Part II men. However, differentials across the four groups are negligible (Table 8). Identical characteristics are essentially rewarded similarly, with very little difference left unexplained by observed characteristics. Much larger differences in entry level salaries are revealed for non-YPs in a pattern similar to that for the total sample, and with generally larger unexplained components as well. Of the total difference in mean entry salaries of 22 percent for Part I women relative to Part I men, 9 percentage points are due to structure; for Part II men, this portion is 9.2 of 10 percent; and for Part II women, 17 of 32 percent.

What happens after entry? While entry level salaries show negligible differences across staff groups, mean annual growth favors YP men over YP women. As a result, over time, male-female salaries of YP recruits diverge. For example, Part II YP women have mean annual pay raises of 3.4

[20] Developments in alternative labor markets will clearly also affect the applicant pool.

percent. Using the coefficients for Part I men to predict annual growth, their rate of salary growth would have been 3.9 percent. The reverse is found for non-YP staff — women's annual salary raises exceed those of men. Hence, there is a catching-up process or convergence for non-YPs over time.[21] Non-YP Part II women whose average salaries have grown at 3.9 percent annually would have had salaries growing at the slightly lower rate of 3.7 percent, if evaluated using the coefficients of Part I men.

Moreover, while the norm is that all YPs enter at grade 21, over time their distribution of grades diverges by gender. Men who were recruited as YPs are concentrated in grades 24 through 26, while the women are mostly in grades 22 through 24. Part of this difference is due to the lower mean tenure of YP women (11 rather than 16 years) resulting from the relatively recent efforts to hire at least an equal share of women in each YP cohort. The difference in mean current grade of Part I YP men and women is 0.8 of a grade, but only 10 percent of this difference is attributable to structure (Table 9). Among Part II staff, men are 1.1 grades higher, on average, than women, of which 0.4 of a grade (36 percent) is due to structure. The other group differences are much smaller (Table 9).[22]

In the non-YP group, without the "equalizing" influence of the YP recruitment process, generally larger unexplained between-group differences emerge in current grades (Table 9). However, these differences are significantly smaller than for grades at entry (Table 10), suggesting a readjustment over time as more information is revealed about staff's ability to perform (Table 11). For example, the difference due to structure between Part II men and women at entry is cut by more than one-half of a grade at mean tenure of about twelve years.

---

[21] For both the YP and non-YP samples, our salary growth regressions explain from 53 to 72 percent (Appendix Tables 8 and 9, Filmer et al. 1998).

[22] Part II men are 0.1 of a grade higher than Part I men, with most of this difference being due to structure. Among women, Part I YPs are 0.2 of a grade ahead of Part II YPs for reasons other than controlled for characteristics. The difference attributable to structure between Part I men and Part II women YPs is only 0.2 of a grade (22 percent of the total difference), compared with over one-half grade in the full and non-YP samples.

These results suggest that the YP recruitment process largely eliminates inter-group differences in starting salaries for given observed individual characteristics. The more transparent and regulated process of making offers to YPs appears to eliminate structural differences. However, after YPs have entered the common staff pool, a divergence in salaries and grades emerges.

This subsequent divergence could indicate that characteristics observed at entry and valued by the program administrators (presumably a richer set of information than our model can account for) are not in fact good predictors of subsequent performance. However, it seems implausible that this would only apply to women.

A more plausible explanation is that biases slowly emerge which render the work environment less conducive to the progress of women. Men may tend to receive more mentoring, benefit from more role models and be favored in task assignments that have higher potential for promotion. In contrast to that for YPs, annual pay raises for non-YP staff appear to gradually compensate over time for the large unexplained differences in entry level salaries. Indeed, while unexplained differences for YPs are greater for current salaries, for non-YPs they are greater at entry level.

In fact, the YP results indicate only small differences in mean salaries between groups that are not accounted for by characteristics. These differences can be interpreted as the effect of omitted characteristics on our estimates of the effect of structure.[23] For example, under this interpretation, the return to the omitted characteristics of Part I men compared to Part I women would be 2.9 percent of entry level salary differentials (Table 8). If the omitted variables are similar for the YP and non-YP samples, this means that about 3 of the 22 percent premium paid at entry to non-YP Part I men compared

[23] Filmer et al. (1998) compare staff hired between 1980 to 1986 and 1990 to 1997 (when quotas came into force), and find that between-group differences in entry salary and grades have diminished over time as has the component attributable to structure. Again, the comparison does not support the view that quotas are driving our results.

to Part I women may be attributable to unobserved variables. The rest of the entry salary difference due to structure (about 6.4 percent) can then plausibly be attributed to discrimination.

8. Conclusions

Our study has found differences in salaries and grades between men and women and between nationalities in the World Bank that are not readily accountable to differences in the attributes that one would expect to influence productivity. The difference in mean current salaries between men and women (holding nationality group constant) is 14 percent, almost half of which cannot be attributed to differences in observed characteristics such as age, education, and previous work experience. The differences in mean salaries between Part I and Part II staff (independently of gender) are much smaller—5 percent—and are largely attributable to structure rather than to differences in characteristics.

This could reflect omitted variable bias, whereby there are characteristics omitted from our model (but observed by those making salary and promotion decisions) that are correlated with observed variables. For omitted variables to be contributing to the component of unexplained differences in salary and grade between staff groups, they must be distributed across groups such that the bias is strongest for Part I men, followed by Part II men and so on. Our analysis of the YP sample — inarguably a more homogeneous group hired through a highly non-discriminatory process — indicates that only negligible inter-group differences in mean salaries are left unexplained by the variables included in our model. Based on these results, we would argue that omitted variable bias is not a compelling explanation for differences attributable to structure.

A second possible explanation for salary and grade differentials is discrimination. We have identified two ways in which this could arise. In its attempt to reduce labor costs, the Bank may be taking advantage of the fact that the four staff groups face different reservation wages. The other

explanation puts more onus on individual managers and their preferences for discrimination coupled with imperfect information. We find that although YPs all begin at the same level and at salaries that differ only to reflect differences in characteristics, once they enter the general staff pool, their salary increases and promotions begin to mirror the patterns found for the rest of the staff. Our results suggest that bias from both of these sources is present.

Nonetheless, recent evidence suggests a shift is occurring in the hiring process at the Bank. Sources for this shift include: (a) on the supply side, that the applicant pool, especially of women and Part II candidates, has significantly improved in quality; (b) that information gathering during hiring has intensified, thus decreasing room for guess work and the opportunities for stereotypes to bias grade assignment and wage offers; (c) that a renewed effort to balance the composition of staff has meant providing more incentives to staff from minority groups; and (d) that as the institution becomes more diversified it becomes more conducive to high performance by women and Part II nationals, as well as less biased. The result of the shift has been to reduce salary and grade differences that cannot be attributed to observed differences in attributes.

References

Behrman, Jere and Nancy Birdsall. 1983. "The Quality of Schooling: Quantity Alone is Misleading, " *American Economic Review* 73(5): 928-946.

Blackaby, David, Alison Booth and Jeff Frank. 2002. "Outside Offers and the Gender Pay Gap: Empirical Evidence from the UK," CEPR Discussion Paper No. 3549

Blau, Francine and Lawrence Kahn. 1996. "Wage Structure and Gender Earnings Differentials: an International Comparison," *Economica* 63: S29-S62.

Blau, Francine and Lawrence Kahn. 2000. "Gender Differences in Pay," *Journal of Economic Perspectives* 14 (4): 75-99.

Blinder, Alan. 1973. "Wage Discrimination: Reduced Form and Structural Estimates," *The Journal of Human Resources* 8: 436-455.

Booth, Alison, Marco Francesconi and Jeff Frank. 1998. "Women are Promoted but not Well Paid," mimeo, University of Essex.

Cotton, Jeremiah. 1988. "On the Decomposition of Wage Differentials," *Review of Economics and Statistics* 70: 236-243.

Filmer, Deon, Margaret Grosh, Elizabeth King and Dominique van de Walle. 1998. "Pay and Grade Differentials at the World Bank," Policy Research Working Paper 1912, World Bank, Washington DC.

Griliches Zvi and William Mason. 1972. "Education, Income and Ability," *Journal of Political Economy* 80(3): S74-S103.

Groot, Wim and Henriette Maassen van den Brink. 1996. "Glass Ceilings or Dead Ends: Job Promotion of Men and Women Compared," *Economics Letters* 53: 221-226.

Gunderson, Morley. 1989. "Male-Female Wage Differentials and Policy Responses," *Journal of Economic Literature* 27 (March): 46-72.

Heckman, James. 1998. "Detecting Discrimination," *Journal of Economic Perspectives* 12(2): 101-116.

Hersch, Joni and Kip Viscusi. 1996. "Gender Differences in Promotions and Wages," *Industrial Relations* 35(4): 461-472

Jones, David and Gerald Makepeace. 1996. "Equal Opportunities: Pay and Promotion in an Internal Labour Market," *The Economic Journal* 106: 401-409.

Lazear, Edward and Sherwin Rosen. 1990. "Male-Female Wage Differentials in Job Ladders," *Journal of Labor Economics* 8(1): S106-S123.

Neumark, David. 1988. “Employers’ Discriminatory Behavior and the Estimation of Wage Discrimination,” *The Journal of Human Resources* 23(3): 279-295.

Oaxaca, Ronald. 1973. “Male-female Wage Differentials in Urban Labor Markets,’ *International Economic Review* 9: 693-709.

Oaxaca, Ronald and Michael Ransom. 1994. “On Discrimination and the Decomposition of Wage Differentials,” *Journal of Econometrics* 61: 5-21.

Oaxaca, Ronald and Michael Ransom. 1993. “Salary and Promotion Differences Between Gender and Nationality Groups for Specific Grade Levels,” prepared for the World Bank, April.

Table 1: Descriptive statistics on pay and tenure of professional staff, May 1997

All staff	Part I men	Part II men	Part I women	Part II women
Current salaries	100	95.3	86.6	82.5
Salaries at entry	100	91.4	86.4	80.6
Mean annual growth (%)	3.2	3.3	3.8	3.8
Mean tenure (years)	11.9	12.5	10.7	11.2
Number of staff	1356	857	537	250
(% of total staff)	(45%)	(29%)	(18%)	(8%)
YPs				
Current salaries	100	99.5	83.3	81.2
Salaries at entry	100	97.8	103.6	105.0
Mean annual growth (%)	3.83	3.94	3.49	3.38
Mean tenure (years)	16.1	16.5	11.1	10.7
Non-YPs				
Current salaries	100	92.6	87.2	82.4
Salaries at entry	100	90.5	80.1	72.7
Mean annual growth (%)	3.15	3.19	3.90	3.92
Mean tenure (years)	11.1	11.4	10.6	11.3

Note: Data on World Bank (IBRD) professional staff pertain to those in grades 22-30.

Table 2: Between-group percentage differences in current salaries

	Total Difference[a]	Characteristics[b]	Structure[c]
M1-M2	5.0	1.06	3.93
M1-F1	14.23	8.64	5.59
M1-F2	19.33	8.75	10.58
M2-F1	9.23	7.57	1.66
M2-F2	14.33	7.68	6.65
F1-F2	5.10	0.11	4.99

Notes: M1 refers to Part I men; M2 to Part II men; F1 to Part I women and F2 to Part II women.
a - Percentage difference in mean annual salaries between indicated groups.
b - Component of total % difference attributable to differences in included characteristics.
c - Component of total % difference attributable to structure.
b and c may not add up to a due to rounding off.

Table 3: Between-group percentage differences in entry level salaries

	Total Difference	Characteristics	Structure
M1-M2	9.21	2.35	6.86
M1-F1	17.33	8.66	8.66
M1-F2	24.64	10.04	14.6
M2-F1	8.12	6.34	1.80
M2-F2	15.43	7.69	7.74
F1-F2	7.32	1.37	5.94

Table 4: Between-group differences in current grade

	Total Difference	Characteristics	Structure
M1-M2	0.21	.01	0.20
M1-F1	0.66	0.45	0.22
M1-F2	0.93	0.39	0.54
M2-F1	0.45	0.44	0.01
M2-F2	0.72	0.38	0.34
F1-F2	0.27	-0.06	0.33

Notes: units are in grades.

Table 5: Between-group differences in entry grade

	Total Difference	Characteristics	Structure
M1-M2	0.50	0.15	0.36
M1-F1	1.36	0.66	0.70
M1-F2	1.96	0.72	1.24
M2-F1	0.86	0.51	0.35
M2-F2	1.46	0.57	0.89
F1-F2	0.60	0.06	0.54

Table 6: Between-group differences in annual rate of promotion

	Total Difference	Characteristics	Structure
M1-M2	-0.01	-0.01	-0.01
M1-F1	-0.06	-0.04	-0.02
M1-F2	-0.08	-0.04	-0.04
M2-F1	-0.04	-0.03	-0.02
M2-F2	-0.06	-0.03	-0.03
F1-F2	-0.02	-0.00	-0.02

Table 7: Between-group percentage differences in current salaries: YP and non-YPs

	YP			**Non-YP**		
	Total Difference	Characteristics	Structure	Total Difference	Characteristics	Structure
M1-M2	0.07	1.75	-1.69	6.33	1.11	5.22
M1-F1	17.08	13.24	3.85	13.7	8.02	5.68
M1-F2	20.35	15.21	5.13	19.41	8.67	10.75
M2-F1	17.02	11.48	5.53	7.37	6.91	0.46
M2-F2	20.28	13.46	6.82	13.08	7.56	5.53
F1-F2	3.27	1.98	1.29	5.72	0.64	5.07

Table 8: Between-group percentage differences in entry salaries: YP and non-YPs

	YP			Non-YP		
	Total Difference	Characteristics	Structure	Total Difference	Characteristics	Structure
M1-M2	0.20	4.35	-4.15	10.18	0.97	9.21
M1-F1	-5.21	-8.11	2.90	21.89	12.54	9.35
M1-F2	-6.43	-5.23	-1.20	31.99	15.16	16.82
M2-F1	-5.41	-12.46	7.06	11.71	11.58	0.14
M2-F2	-6.63	-9.58	2.95	21.81	14.20	7.61
F1-F2	-1.22	2.88	-4.10	10.10	2.62	7.47

Table 9: Between-group differences in current grade: YP and non-YPs

	YP			Non-YP		
	Total Difference	Characteristics	Structure	Total Difference	Characteristics	Structure
M1-M2	-0.12	0.05	-0.17	0.31	0.02	0.29
M1-F1	0.78	0.69	0.09	0.66	0.43	0.23
M1-F2	0.98	0.75	0.23	0.96	0.43	0.53
M2-F1	0.90	0.64	0.26	0.35	0.41	-0.06
M2-F2	1.10	0.70	0.40	0.65	0.41	0.24
F1-F2	0.20	0.06	0.14	0.30	0.003	0.30

Table 10: Between-group differences in entry grade: YP and non-YPs

	Non-YP		
	Total Difference	Characteristics	Structure
M1-M2	0.58	0.05	0.54
M1-F1	1.68	0.91	0.77
M1-F2	2.42	1.02	1.39
M2-F1	1.10	0.86	0.24
M2-F2	1.83	0.98	0.85
F1-F2	0.74	0.12	0.62

Note: Since all YPs enter at grade 21 there are no differences to decompose.

Table 11: Between-group differences in annual promotion rate: YP and non-YP sample

	YP			Non-YP		
	Total Difference	Characteristics	Structure	Total Difference	Characteristics	Structure
M1-M2	0.03	0.03	0.000	-0.02	-0.003	-0.01
M1-F1	-0.01	0.002	-0.01	-0.06	-0.03	-0.02
M1-F2	0.002	0.005	-0.003	-0.08	-0.03	-0.05
M2-F1	-0.04	-0.03	-0.01	-0.04	-0.03	-0.01
M2-F2	-0.03	-0.03	-0.003	-0.06	-0.03	-0.03
F1-F2	0.01	0.003	0.01	-0.02	0.002	-0.02