DRAFT --PLEASE DO NOT CITE OR CIRCULATE WITHOUT PREMISSION

# Gender Earnings' Differentials: Disaggregating Human Capital Characteristics

Elia De la Cruz Toledo[1]

Loyola Marymount University

Abstract

I establish that including a more precise measure of work experience and measures of cognitive ability, and non-cognitive traits increases the percentage attributed to endowments in the estimation of the gender gap in weekly wages in Mexico. The inclusion of these variables decreases the share of the gap attributable to returns. My results show a gender gap of 15% in weekly wages. A quarter of the gender gap in weekly wages is explained by differences in women's and men's endowments. Another quarter is explained by differential returns to employees' characteristics, while half of the gap remains unexplained. The inclusion of the additional measures of human capital increases the share attributable to endowments by eight percentage points, and decreases the share attributable to returns by ten percentage points. The share of endowments explained by human capital increases by fivefold once work experience, and cognitive and non-cognitive traits are included in the model. Previous studies on gender wage differentials have found a similar gap, but fail to include, and consequently, disaggregate the contribution of the different human capital characteristics to the gender gap in weekly earnings.

[1] Contact email: elia.delacruztoledohe@lmu.edu
I am indebted to Jane Waldfogel for her guidance and for sparking my interest in this topic. I thank Ronald Mincy, Neeraj Kaushal, Lena Edlund, Rajiv Sethi, and Jorge U. Colin Pescina for helpful comments that allowed me to improve this draft. I am also thankful for the feedback received at the 2015 Population Association of America Annual Meetings. I acknowledge support from the Mexican National Council of Science and Technology (CONACYT) and from the Columbia Population Research Center at Columbia University. All remaining errors are my own.

*I. Introduction*

In his seminal paper, Jacob Mincer (1958) argued that education and experience play a fundamental role in the determination of a worker's earnings. A wage differential between men and women could mirror a difference in workers' endowments of important aspects of human capital, in particular education and work experience, but it could also be a sign of differential returns to men's and women's characteristics. Hence, it is important to include accurate measures of these variables to estimate wage gaps. Literature analyzing earnings' differentials in Mexico has found mixed evidence on the overall composition of the gender wage gap.[2] Previous studies have failed to incorporate key aspects of human capital that impact workers' earnings. I establish that human capital endowments account for a larger share of the gender earnings' gap once a more precise measure of work experience and measures of cognitive ability and non-cognitive skills are added to the earnings' model.

In this study, I use detailed data on jobs, schooling, IQ (proxy for cognitive ability), self-reported measures of non-cognitive traits, and earnings of full time employees from the Mexican Family Life Survey (MxFLS) for 2002. The empirical strategy relies on the Oaxaca-Ransom decomposition to analyze the gender gap in two outcomes: weekly wages and weekly wages plus employer-provided benefits. Under this decomposition methodology, the component attributable to endowments measures the expected change in female's mean earnings, if females had males' human capital, and other relevant work-related characteristics. The component attributable to returns measures the expected change in females' mean earnings, if females were treated as males. Wages and labor income are only observed for people who are working, and this is not necessarily a random sample of the population. Thus, it is common, and relevant, to include a

[2] Brown et al. (1999), Pagan and Ullibarri (2000) and Sánchez et al. (2001) have found that most of the wage gap is due to differences in human capital endowments. Yet, Popli (2008) found the most of the gender wage gap is explained by discrimination.

correction for sample selection bias in the wage equations, based on the procedure developed by Heckman (1976). Based on this, I apply a selection adjustment to account for selection into employment prior to the decomposition calculations.

Once more precise human capital characteristics are incorporated into the earnings' equation; five main conclusions are derived from the analysis. First, the gender gap in weekly wages is 15%, while the gender gap in weekly wages plus benefits is 45%. Second, a quarter of the gender gap in weekly wages is explained by differences in women's and men's endowments. Another quarter is explained by differential returns to employees' characteristics, and half of the gap remains unexplained. Third, the share of the gender gap in weekly wages attributable to endowments increases by eight percentage points, and the share attributable to returns decreases by ten percentage points when more complete measures of human capital (i.e. projected work experience, cognitive skills, and non-cognitive skills) are added to the model. Fourth, the explanatory power of human capital characteristics on gender differences in the returns is 15%, while human capital explains 50% of gender differences in endowments. Fifth, work experience accounts for half of the human capital's share in endowments (7 out of 15 percentage points) and half of the share in returns (26 out of 49 percentage points). Disaggregated results for the different components of the gender gap in weekly wages that also adds employer-provided benefits yield similar conclusions.

The rest of the paper is organized as follows. Section II briefly describes the prior research on gender differences in earnings in Mexico. Section III presents the data and methods for this study. Section IV presents the analysis of the results. The discussion can be found in section V and section VI concludes.

*II. Prior Literature*

Several authors have attempted to explain the sources of gender wage differentials in Latin America, exploring issues such as differences in individual characteristics and human capital endowments, labor market regulations, and occupational segregation (Tenjo, 1992, 2004; Brown, et al., 1999; Lim, 2002; Rendón, 2003, 2004; Cruces & Galiani, 2007; Deutsch et al., 2004; Atal et al., 2009), among others. The literature has also attempted to relate gender wage gaps to differences in income generating opportunities available in urban and rural areas, but no clear link can be found (Hertz et al., 2008).

In Mexico, Brown et al. (1999) and Sánchez et al. (2001) found that most of the gender wage gap could be explained by gender differences in human capital endowments. Both studies relied on the National Urban Employment Survey (ENEU) and both used a decomposition analysis. Pagan and Ullibari (2000) analyzed the gender gap in weekly earnings across heterogeneous socio-demographic groups using data from ENEU for the year 1995, potential work experience as a proxy for actual work experience, and an additively decomposable index (Jenkins inequality measure[3]). Pagan and Ullibari (2000) found an unexplained gender gap in weekly earnings of 10.4%, being the largest among individuals with either high or low levels of educational attainment. Meza (2001) estimated gender differentials in hourly wages in Mexico from 1988 to 1998. Meza (2001) used data from ENEU for full time workers, and Juhn, et al. (1991) decomposition methodology to measure changes in wage structures, and gender wage differences throughout the wage distribution of male and female workers. This author estimated

[3] The Jenkins (1994) index summarizes the distribution of the unexplained gender wage gap based on the difference between two generalized Lorenz curves. These two curves represent the predicted distributions of female earnings and the counterfactual distribution of female earnings, under the assumption that they are treated as males.

a gender wage gap for 1998[4] of 6% between the 50th percentile of the female wage distribution compared to the 50th percentile of the male wage distribution, and a gender gap of 11% for the 25th-25th percentiles of the wage distribution. The author also found that the gender wage gap also dropped in Mexico between 1988 and 1998.

Most relevant to this study, Popli (2008) estimated the most recent gender gap in hourly wages in the Mexican labor market for 2002. This study used data from the National Income and Expenditure Household Survey (ENIGH), age dummies to proxy work experience, and three different methodologies: a decomposition analysis, the Jenkins measure, and a non-parametric distribution approach. The estimate of the gender wage gap in this study is 21% in 1984 and 16% in 2002. Using a non-parametric approach that created wage counterfactuals, Popli (2008) estimated that half of the gender gap in wages was due to differences in characteristics and half due to difference in returns, with this later component being the only significant share of the two, thus attributing most of the gap to labor market discrimination. Previous literature on the Mexican gender wage gap has failed to include a precise measure of work experience. Instead, these studies rely on age dummies or potential work experience as proxies for work experience and seniority. Prior literature in this topic has not included cognitive or non-cognitive traits in the estimation of earnings' differentials.

## *III. Data and methods*

### ***Data***

Data for this study come from MxFLS. This is a longitudinal survey that collects a wide range of information on demographics, employment decisions, family dynamics, mental health and emotional wellbeing among others. In addition, this is the only database in Mexico that

[4] Meza (2001) provides estimates of the gender gap in hourly wages for all the years from 1988 to 1998, but I only mention results for the latest year available.

contains information on cognitive ability. The survey was conducted during 2002 and has information on monthly income, hours worked and work experience for years 2000 and 2002. The sample size consists of approximately 4,534 individual observations with information in all the relevant variables including earnings. The analytic sample is restricted in two ways. First, it includes workers ages of 25 to 54 years of age to abstract from school enrollment and retirement decisions. Second, as women tend to be overrepresented in part-time, flexible jobs, the analytic sample is restricted to full-time workers, to have more comparable men and women. From the sample of workers previously mentioned, 2,870 are full time workers with complete information on the relevant variables, 24% of them are women and 76% are men. The MxFLS has information on self-reported monthly income and also on the disaggregated sources of income, including wages and employer-provided benefits. While there are 2,870 full time workers with information on total monthly income, there are only 850 full time workers with information on wages.

***Outcome Measures: wages and wages plus employer-provided benefits***

In this study, two measures of earnings are analyzed: weekly wages and total weekly wages that also include employer-provided benefits. Wages are reported on a monthly basis, and then adjusted using self-reported weeks worked to generate weekly wages. Weekly and not hourly wages are preferred to decrease the measurement error introduced by using working hours that might be under or over reported. As opposed to using weeks worked, the measurement error for hours worked can be greater. The second measure of earnings is added to have a broader understanding of differences in the types of jobs that men and women have. This measure of weekly wages plus employer-provided benefits adds up information from wages, piecework, tips, extra hours, meals, housing and transportation allowances, and medical benefits. Among these benefits, only transportation allowances benefit the worker exclusively, the rest of the benefits potentially benefit the whole family. Income from the main and secondary jobs from the aforementioned sources is added up. One relevant aspect of compensation that is not accounted for is job flexibility. Women might be working in more amenable workplaces that allow for remote work options or other flexible work arrangements, where they can achieve a better work-family balance. This type of flexibility is not measured in the compensation package, thus differentials in wages plus benefits might overestimate the actual gender differences.

***Selection correction***

Wages and labor income are only observed for people who are participating in the labor force and this might be a selective group. Thus, it is common to include a correction for sample selection bias in the wage equations based on the procedure by Heckman (1976, 1979). The most straightforward approach to account for selection bias in a decomposition analysis is to deduct the selection effects from the overall differential and then apply the standard decomposition formulas to this adjusted differential (Reimers 1983; Dolton and Makepeace 1986; Neuman and Oaxaca 2004).

The following models use Heckman selection correction to adjust for selection into employment. As a first step I use a probabilistic model of labor force participation for women, controlling for age, age squared, marital status (married or cohabitating, single and divorced, separated or widowed), education (incomplete primary education, primary or some secondary education, secondary or some high school, and some college and beyond), a four-level dummy that captures the effect of number of children (one-child, two children and three or more children, the base category is no children)and urban residence. As a second step predicted values of employment are calculated based on the coefficients of the predictors, holding predictors at their mean. The third step is to calculate the inverse mills ratio (IMR) which is equal to the conditional expectation of a standard normal random variable. The IMR answers the question "what is the probability of an event given that the event has not already occurred" (Autor, 2003). The fourth step is to estimate earnings' differences using the OR decomposition, including the IMR as one of the predictors.

***Empirical strategy***

The most common regression-based method used in the literature to measure earnings' gaps is the counterfactual decomposition technique (and its variations) developed by Blinder (1973) and Oaxaca (1973). In this method, mean differences in workers' earnings can be disaggregated into two components: endowments and returns. The "endowments effect" measures the expected change in female's mean earnings, if females had males' predictor levels. This is the part of the differential explained by group differences in the predictors. The second component known as the "coefficients effect" measures the expected change in female's mean earnings, if females had males' coefficients. This component measures the portion of the earnings' differential due to returns in a worker's characteristics.

The main criticism to the Blinder-Oaxaca decomposition is that using a single gender earnings' structure as a norm for measuring discrimination and productivity differentials is too extreme. Extensive research on earning's decomposition has resulted in improved models for measuring the wage gap (Cotton, 1988; Oaxaca and Ransom, 1994; Dinardo et al., 1996; Juhn et al., 1991, 1993; Machado and Mata, 2005; Chernozhukov, 2010). Particularly, Oaxaca and Ransom (1994) developed a pooled method where the wage structure obtained from the pooled regression is interpreted as an estimate of a competitive norm. This decomposition methodology consists in estimating three separate constrained linear wage regressions: one for males, one for females and one for the pooled sample of all workers. This last equation would include a gender intercept that would shift along with the identification restriction. Namely, the pooled regression is an estimate of what the wage structure would be if there was no wage discrimination (Fortin, 2008).

Under the Oaxaca-Ransom (OR) decomposition[5], worker's earnings for males (m), females (f) and for the pooled sample of males and females (p) are represented in the following linear model:

$$Y_{ig} = \alpha_{0g} + \alpha_{ig}\underline{X}_{ig} + e_{ig}, \quad g=f,m,p \tag{1}$$

In this model, $Y_i$ represents a logged measure of weekly wages or weekly total wages plus employer-provided benefits The $\underline{X_i}$ is a vector that includes all the control variables for the *i*-th individual's characteristics. Controls include dummies for: age, age squared, education (incomplete primary education, primary or some secondary education, secondary or some high school, and some college and beyond), region (urban and rural), formal job contract, family status (i.e. whether the worker is head of the household). State fixed effects are included.

There is an ongoing debate in labor economics regarding inclusion of variables that might be endogenous to wages, such as occupations (Blau and Ferber, 1987). The main models do not include controls for occupations, but additional analyses including fifteen dummies for different occupations, based on the North American Classification System (SCIAN), are included in the appendix. The independent and identically distributed error is represented as $e_i$. Errors are clustered at the state level.

The average earnings' differentials, based on the OR decomposition can be decomposed as:

[5] The detailed disaggregation of the decomposition methodology can be found in the Appendix.

$$\bar{Y}_{im} - \bar{Y}_{if} = \Delta X\hat{\beta}_p + \left[\bar{X}_m(\hat{\beta}_m - \hat{\beta}_p) + (\hat{\beta}_{0m} - \hat{\beta}_{0p})\right] - \left[\bar{X}_f(\hat{\beta}_f - \hat{\beta}_p) + (\hat{\beta}_{0f} - \hat{\beta}_{0p})\right] \quad (1a)$$

Where the first term represents changes in endowments, the second term represents the "advantage of men", in the form of higher returns to their characteristics, and the last term represents the "disadvantage of women", in the form of lower returns to their characteristics (Fortin, 2008). A residual portion of the wage gap that is not explained by returns or endowments of workers' characteristics is deemed as unexplained.

***Augmented models***

One problem that arises when measuring the gender wage gap is that if important characteristics are omitted when calculating this gap, the unexplained component will capture the effect of unobserved group differences in productivity and tastes. Thus, the unexplained portion of the gap might be overestimated to the extent that unexplained pay differentials between men and women are due to gender differences in unmeasured qualifications. Measurement error in wage gaps can also cause an overestimation of the portion due to returns. If women's work experience is measured with error, as it tends to be the case when potential work experience is used, the return to experience might be lower for women, suggesting that gender wage gaps are mainly explained by differences in returns. Once a more precise measure of work experience is used, the portion due to returns might decrease while, at parallel, the portion due to endowments increases.

In order to provide more precise estimates of the shares attributable to endowments and returns, I include human capital variables, that previous studies on wage differentials in Mexico had fail to incorporate: years of work experience, a measure of IQ (as a proxy for cognitive ability) and non-cognitive traits (that reflect emotional wellbeing).

***Work experience***

Women are more likely than men to interrupt their careers to bear and raise children. For this reason, on average, women have less work experience than men. But within women there is also heterogeneity in their labor choices. Mothers have lower work experience than non-mothers (Light and Ureta, 1995; Waldfogel, 1997; Budig and England, 2001). Thus, it would be imprecise to assume continuous work profiles for all women. In order to have a correct model of women's earnings it is crucial to have information on work experience and job interruptions. Wage models ideally should have measures of job continuity and tenure with each employer.

In the case of Mexico, researchers have relied on age dummies or age and age squared to measure the effects of seniority (Sanchez et al., 2000; Meza, 2000; Popli, 2008). Others have used potential work experience as a proxy for actual work experience by subtracting years of education and to years of age –adjusting for the number of years before entering school (Brown et al., 1999; Pagán and Ulibarrri, 2000).

On average, men have strong labor force attachment and potential work experience is a reasonable proxy for men's actual experience (Oaxaca, 1973). Women have more job interruptions and the use of potential experience is an inadequate proxy to measure women's job-related skills (Oaxaca, 1973; Light and Ureta, 1995; Waldfogel, 1997; Altonji and Blank, 1999). Potential experience generally overstates women's work experience since it does not account for the time that women spent out of the labor market for child rearing. Theory predicts that women's potential experience-wage profile is flatter than the actual experience-wage profile (Oaxaca, 1973). Waldfogel (1997) calculates that, on average, the ratio of actual–to–potential work experience for women in the U.S. is a little over two thirds.[6]

[6] Waldfogel (1997) also shows that the ratio is 66% for married mothers v. 77% for women without children. The difference between potential and actual experience is even greater (59%) for never married women with children.

The MxFLS collected information on an individual's first job (day, month and year). Based on this information, *projected work experience* is computed subtracting the self-reported year of an individual's first job to the year of the survey. In addition, I adjusted women's level of work experience to account for the time off for childbearing. Since there is no self-reported measure of this time off, the adjustment is based on the number of self-reported pregnancies and the minimum maternity leave period (3 months) for workers at a formal job. The analytic sample consists of full time workers and the majority (59%) has a formal job contract. After the adjustment, a woman's average work experience decreases by 0.59 years in comparison to the average work experience pre-adjustment, this translates into approximately 7 months. Although this differs from a perfect measure, is to date the most accurate measure of work experience available in the literature of gender wage gaps in Mexico. Given the nature of this variable, the estimates of the gap in projected work experience among male and female workers should be below the gap in actual work experience, but above the 2-year gap in potential work experience found in the literature (Pagan and Ullibarri, 2000). As described, this type of work experience constitutes a general measure of human capital. A measure of specific human capital is available in the data (i.e., job tenure at last job); however the number of observations with information on job tenure is not large enough to allow the use of this variable. This distinction is important as it matters who is willing to pay for training (Becker, 1964).

***Cognitive ability and non-cognitive traits***

Market equilibrium wages and promotion policies depend on efficient job assignments (Lazear & Rosen, 1990). Worker's ability plays an important role in job assignment because it is efficient to assign high-skilled workers to the most productive jobs. When cognitive ability is not observed, employers find proxies to determine workers' tasks and promotion paths and discrimination arises. However, when a wage model does not include a measure of ability it is tempting to claim that a worker has been discriminated when in fact she might have not.

In this study a measure of IQ is used as a proxy for cognitive ability and should be highly correlated with wages. Cognitive scores are measured through Raven's Progressive Matrices Test (RPM). Raven's test was set out with the specific intent of developing tests which would be easy to administer and also easy to interpret in a clear, theoretically relevant way (Raven, 1936; Watt, 1998). This test is a well-validated measure of basic cognitive function; it is designed to measure the person's cognitive ability and does not require the person to be literate. The theoretical framework which guided the development of the tests has since been confirmed in numerous studies (Horn, 1994; Matarazzo, 1990; Ree, Earles, & Teachout, 1994; and Snow, Kyllonen & Marshalek, 1984).

The version of the RPM test applied to Mexican women in the MxFLS is made up of a series of diagrams or designs with a part missing. Those taking the tests were asked to select the correct part to complete the designs from a number of options printed beneath (Raven, 2000). In this study, the RPM score is included in the earnings' model as a continuous variable from 0 to 100.

Non-cognitive skills strongly influence schooling decisions and also affect wages (Heckman et al., 2006; Cawley et al., 2001). Heckman et al. (2006) showed that a change in non-cognitive skills from the lowest to the highest level has an effect on behavior comparable to or greater than a corresponding change in cognitive skills. According to Fortin (2008) some of the features that influence the level of job effort –and that are linked to job responsibility– are a person's self-esteem and the external locus of control. Some of the mechanisms through which non-cognitive skills raise wages are through productivity and, indirectly, also through schooling and work experience Heckman et al. (2006). The literature on personality and earnings usually incorporates non-cognitive factors from the Rosenberg self-esteem[7] and the Rotter locus of control[8] scales (Heckman et al. 2006; Fortin, 2005, 2008; Manning and Swaffield, 2008).

In the MxFLS, there are a set of questions asked to individuals about their own perceptions on emotional aspects of their lives, i.e., feelings of depression, feelings of accomplishment, difficulty concentrating and poor performance assessment, among others. These variables are coded as 1 if the person expressed having negative feelings all or most of the times in three categories: *poor performance*, *feelings of insecurity* and *pessimistic feelings*.

The augmented models that progressively incorporate years of work experience, cognitive ability and non-cognitive traits are the following:

$$Y_{ig} = \alpha_{0g} + \alpha_{ig}\underline{X}_{ig} + \beta_{1g}Projected\ work\ experience + \beta_{4g}IMR_{ig} + e_{ig} \tag{2}$$

$$Y_{ig} = \alpha_{0g} + \alpha_{ig}\underline{X}_{ig} + \beta_{1g}Projected\ work\ experience + \beta_{2g}Cognitive\ ability_{ig} + \beta_{4g}IMR_{ig} + e_{ig} \tag{3}$$

$$Y_{ig} = \alpha_{0g} + \alpha_{ig}\underline{X}_{ig} + \beta_{1g}Projected\ work\ experience + \beta_{2g}Cognitive\ ability_{ig} + \beta_{3g}Noncognitive\ traits_{ig} + \beta_{4g}IMR_{ig} + e_{ig} \tag{4}$$

Earnings' differentials using equations (2) – (4) are calculated through the same OR decomposition method described in equation (1a).

*IV. Results*

Descriptive statistics indicate that, on average, working men have higher earnings than working women, with weekly wages 15% above those of women's and a gap in weekly wages

[7] Rosenberg (1979) posited four principles of self-concept formation: reflected appraisals, social comparisons, self-attribution, and psychological centrality.

[8] Rotter (1954) suggested that people generally identify either an internal or external locus of control in their lives. Those with an internal locus of control tend to believe in their own ability to control events, whereas people with an external locus of control believe other people or events determine their own circumstances.

plus benefits of 45% among men and women (See appendix table 1). Gaps in wages and income are often attributed to differences in experience and education. Descriptive data from Table 1 show mixed evidence on this hypothesis. On average, men have seven more years of projected work experience than women, and men work two more hours per week, compared to women.[9] However, on average women are more educated than men. A greater percentage of women have completed college and beyond, in comparison to similar men (17% v. 12%, respectively). As for participation in the informal economy, 59% of women and 55% of men have a written or verbal job contract. On average, the majority of the workers live in urban regions (82%), around 83% of the workers have at least one child and the average number of children is between two and three. There are significant differences in family composition by gender. While most of the working men are married (95%), less than half of the full-time female employees are married. On the other hand, only 3% of full-time male workers are single, while 36% of full-time female workers are single. It is worth notice that, regarding non-cognitive traits, there are a larger proportion of women who have negative feelings of self-perception. Among this sample of full time workers, women tend to feel more often a *poor performance at work*, *feelings of insecurity* and *pessimistic attitudes* in comparison to men. Differences in these self-reported measures are around 8-14 percentage points (pp). Even though women have a higher average score in the cognitive ability test, the difference among men and women is only 1 point out of 100. Finally, while 99% of the male full-time workers are head of their households, only 22% of comparable female workers are the main breadwinners. A table with the descriptive statistics for the sample that has information on wages plus employer-provided benefits is found in the appendix. This sample does not differ significantly from the one described.

[9] In comparison, Pagan and Ullibarri (2000) report a difference of two years of potential work experience among male and females working in Mexico in 1995.

**Table 1 Descriptive statistics of full time workers of 25-54 years of age with information on wages**

| | All | | Women | | Men | |
|---|---|---|---|---|---|---|
| *Employment variables* | | SE | | SE | | SE |
| Log weekly wage | 4.73 | (0.001) | 4.23 | (0.128) | 4.34 | (0.031) |
| Weekly hours worked | 50.64 | (0.376) | 49.18 | (0.601) | 51.38 | (0.474) |
| Yearly weeks worked | 49.10 | (0.005) | 48.27 | (0.011) | 49.49 | (0.005) |
| Projected work experience | 18.20 | (0.347) | 13.62 | (0.542) | 20.53 | (0.407) |
| Formal contract | 0.560 | | 0.588 | | 0.546 | |
| *Cognitive / non-cognitive traits* | | | | | | |
| Cognitive ability (IQ score) | 51.353 | (0.909) | 52.007 | (1.600) | 51.020 | (1.104) |
| Feelings of poor performance | 0.139 | | 0.192 | | 0.112 | |
| Feelings of insecurity | 0.194 | | 0.286 | | 0.147 | |
| Pessimistic attitude | 0.201 | | 0.286 | | 0.158 | |
| *Marital Status* | | | | | | |
| Married | 0.795 | | 0.482 | | 0.954 | |
| Single | 0.138 | | 0.355 | | 0.027 | |
| Divorced/separated | 0.067 | | 0.163 | | 0.019 | |
| *Education* | | | | | | |
| Incomplete primary | 0.036 | | 0.024 | | 0.041 | |
| Elementary or some secondary | 0.333 | | 0.294 | | 0.353 | |
| Secondary or some high school | 0.329 | | 0.355 | | 0.315 | |
| Complete high school | 0.164 | | 0.159 | | 0.166 | |
| Some college and beyond | 0.139 | | 0.167 | | 0.124 | |
| *Fertility* | | | | | | |
| No children | 0.173 | | 0.208 | | 0.156 | |
| One child | 0.135 | | 0.196 | | 0.104 | |
| Two children | 0.205 | | 0.204 | | 0.205 | |
| Three or more | 0.487 | | 0.392 | | 0.535 | |
| *Other characteristics* | | | | | | |
| Head of household | 0.729 | | 0.216 | | 0.990 | |
| Urban residence | 0.824 | | 0.861 | | 0.805 | |
| Actual observations | 733 | | 247 | | 486 | |
| Weighted observations | 2,748 | | 872 | | 1,875 | |

Note: 1. Earnings were converted to US dollars using the average exchange rate for year 2002 of 9.46 pesos per dollar. 2. Weighted observations are in thousands.

I next estimate a series of multivariate regression models. As mentioned before, all models adjust for selection correction using the method described above. Results from the OR decomposition show that there is a statistically insignificant gender gap in weekly wages of 14-15% in the Mexican labor market. This gap varies due to the interaction of the Inverse Mills Ratio with the new variables added in each model, causing slightly different predictive wages for female workers.

As observed in table 2, using Popli's (2008) benchmark model that uses age as a proxy for work experience (model 1), the largest share of the gender gap in weekly wages remains unexplained (48%); although this share is statistically insignificant. Under this model specification, 18% of the gender gap in weekly wages is explained by endowments in human capital and other worker characteristics and 34% of the gap is explained by returns to a worker's characteristics. Under model (2) that adds projected work experience, the share explained by returns decreases 11 pp, the unexplained share increases 8 pp and the share explained by endowments increases 3 pp. Under model (3), once cognitive ability as well as projected work experience is added to the model, the shares of the gender gap in weekly wages explained by returns and endowments practically do not change. Under model (4) that adds projected work experience, and cognitive and non-cognitive traits, the share of the gender gap in weekly wages attributable to endowments is 25%, and the share attributable to returns is 26%. The share attributable to returns is statistically significant. Overall, the unexplained share of the gender gap in weekly wages remains close to half.

**Table 2 Weekly wage for full-time workers**

| | (1)<br>Age | (2)<br>Projected work experience | (3)<br>Cognitive ability | (4)<br>Non-cognitive traits |
|---|---|---|---|---|
| Men | 4.38 | 4.38 | 4.38 | 4.38 |
| | (0.031) | (0.031) | (0.031) | (0.031) |
| Women | 4.24 | 4.28 | 4.27 | 4.23 |
| | (0.115) | (0.128) | (0.129) | (0.128) |
| **Difference** | **0.136** | **0.106** | **0.112** | **0.148** |
| | **(0.119)** | **(0.132)** | **(0.132)** | **(0.132)** |
| Decomposition of estimated differentials (in %) | | | | |
| Endowments | 18.1% | 21.4% | 20.7% | 25.4% |
| Returns | 34.1% | 22.7% | 22.3% | 25.9%* |
| Unexplained | 47.8% | 55.9% | 57.0% | 48.7% |
| N | 850 | 850 | 850 | 850 |

Note: 1. Wages are calculated in logarithmic terms and converted into USD. 2. Robust standard errors are clustered at the state level. 3. Significance levels: +p <0.10, *p<0.05, **p<0.01, ***p<0.001. 4. Decomposition of estimated differentials reflects the share of the gender gap in weekly wages under each specification and adds up to 100%.

The share of the gender gap in weekly wages attributable to human capital endowments is 15% (Table 3). Projected work experience accounts for 47% of that share, cognitive ability accounts for 0.6%, non-cognitive traits accounts for 32%, and education accounts for 19% (see appendix table 3). The share of the gender gap in weekly wages attributable to returns to human capital characteristics is 49%. Among the characteristics that impact a worker's returns on weekly wages, projected work experience accounts for 53%, cognitive ability accounts for 8%, non-cognitive traits accounts for 26%, and education accounts for 13%. Adding variables of projected work experience, cognitive ability and non-cognitive traits increases the share explained by human capital endowments from 3% to 15%, and the share explained by human capital returns from 8% to 49%. Among these human capital characteristics, work experience is the largest contributor to the portion explained by differences in endowments and in returns. Thus, adding an adequate measure of work experience is relevant in the estimation of the gender gap in weekly wages.

**Table 3 Contribution of human capital characteristics to gender differences in in weekly wages**

| | **Endowments** | | | | **Returns** | | | |
|---|---|---|---|---|---|---|---|---|
| | Model (1) | Model (2) | Model (3) | Model (4) | Model (1) | Model (2) | Model (3) | Model (4) |
| Education | 0.033 | 0.033 | 0.031 | 0.030 | 0.081 | 0.074 | 0.070 | 0.064 |
| Work experience | | 0.091 | 0.091 | 0.073 | | 0.287 | 0.289 | 0.260 |
| Cognitive Ability | | | 0.001 | 0.001 | | | 0.036 | 0.044 |
| Non-cognitive traits | | | | 0.050 | | | | 0.127 |
| **Share of endowments explained by human capital characteristics** | **0.033** | **0.124** | **0.122** | **0.154** | **0.081** | **0.361** | **0.395** | **0.495** |
| Share of endowments explained by personal characteristics | 0.97 | 0.876 | 0.878 | 0.846 | 0.919 | 0.639 | 0.605 | 0.505 |

Note: 1. Differences were calculated in relative terms and adding up the coefficients of the different variables that were used to portray each characteristic included in the OR models that corrected for selection into employment. 2. The total share of endowments explains 25% of the gender differential in weekly wages. 3. The total share of returns explains 25% of the gender differential in weekly wages. 4. Non-cognitive traits include poor performance at work, feelings of insecurity and pessimistic attitudes.

When using the other measure of earnings, that also includes employer-provided benefits; the gender gap in weekly income triples (15% v. 45%) (Table 4). Under model (1), 5% of the gender differences in weekly wages plus employer-provided benefits are explained by endowments, 43% are explained by returns and 52% remains unexplained. The share of the gender differences explained by endowments once projected work experience is taken into account decreases 3 pp, the unexplained portion raises 1 pp, and the share explained by returns increases 2 pp. Adding cognitive and non-cognitive ability in addition to work experience shows the opposite trend: a decreasing explanatory share of returns, an increasing explanatory share of endowments, and subtle changes in the share of the unexplained residual. In all models, the share explained by the endowments of workers characteristics is not statistically significant. However, the share explained by differences in returns to characteristics is always statistically significant. This suggests that men and women might be receiving different returns to their mean job-relevant characteristics.

**Table 4 Weekly wages + employer provided benefits for full-time workers**

| | (1)<br>Age | (2)<br>Projected work experience | (3)<br>Cognitive ability | (4)<br>Non-cognitive traits |
|---|---|---|---|---|
| Men | 4.34 | 4.34 | 4.34 | 4.34 |
| | (0.018) | (0.018) | (0.018) | (0.018) |
| Women | 3.98 | 3.91 | 3.90 | 3.89 |
| | (0.074) | (0.083) | (0.083) | (0.083) |
| **Difference** | **0.360*****| **0.436***** | **0.445***** | **0.455***** |
| | **(0.076)** | **(0.085)** | **(0.085)** | **(0.085)** |
| Decomposition of estimated differentials (in %) | | | | |
| Endowments | 4.7% | 1.9% | 3.1 | 5.0% |
| Returns | 42.7%*** | 44.7%*** | 43.6%*** | 42.8%*** |
| Unexplained | 52.6% | 53.3%*** | 53.3%*** | 52.2%*** |
| N | 3,326 | 3,326 | 3,326 | 3,326 |

Note: 1. Income is calculated in logarithmic terms and converted into USD. 2. Robust standard errors are clustered at the state level. 3. Significance levels: +p <0.10, *p<0.05, **p<0.01, ***p<0.001. 4. Decomposition of estimated differentials reflects the share of the gender gap in weekly wages plus employer-provided benefits under each specification and adds up to 100%.

Gender differences in returns are explained by human capital characteristics in the following way: work experience accounts for 45%, cognitive ability explains 41%, non-cognitive traits explain 5% and education explains 9% (Table 5). Together, these variables explain 44% of the gender income differences. Gender differences in workers' human capital endowments are mainly explained by work experience (69%), followed by education (19%). Endowments in cognitive and non-cognitive traits do not explain a significant portion of the endowments' share (2% and 11%, respectively). The relative weight of these four variables on the total share of the gender gap in weekly wages plus employer-provided benefits attributable to endowments is 15%.

**Table 5 Contribution of human capital characteristics to gender differences in returns and endowments in weekly wages + employer provided benefits**

| | **Endowments** | | | | **Returns** | | | |
|---|---|---|---|---|---|---|---|---|
| | Model (1) | Model (2) | Model (3) | Model (4) | Model (1) | Model (2) | Model (3) | Model (4) |
| Education | 0.034 | 0.034 | 0.030 | 0.029 | 0.053 | 0.054 | 0.040 | 0.042 |
| Work experience | | 0.075 | 0.092 | 0.107 | | 0.251 | 0.223 | 0.197 |
| Cognitive Ability | | | 0.002 | 0.002 | | | 0.173 | 0.178 |
| Non-cognitive traits | | | | 0.017 | | | | 0.022 |
| **Share of endowments explained by human capital characteristics** | **0.034** | **0.109** | **0.124** | **0.155** | **0.053** | **0.305** | **0.436** | **0.438** |
| Share of endowments explained by personal characteristics | 0.96 | 0.891 | 0.876 | 0.845 | 0.947 | 0.695 | 0.564 | 0.562 |

Note: 1. Differences were calculated in relative terms and adding up the coefficients of the different variables that were used to portray each characteristic included in the OR models that corrected for selection into employment. 2. The total share of endowments explains 25% of the gender differential in weekly wages plus employer-provided benefits. 3. The total share of returns explains 25% of the gender differential in weekly wages plus employer-provided benefits. 4. Non-cognitive traits include poor performance at work, feelings of insecurity and pessimistic attitudes.

## *V. Discussion*

In Mexico studies that have analyzed gender gaps in earnings have failed to incorporate measures of important aspects of human capital –projected work experience, cognitive ability, and non-cognitive traits. With models that relied on age dummies or measures of potential work experience, the previous literature has found mixed evidence on whether endowments or returns explain different or equal shares of the gender wage gap in Mexico. I establish that returns and endowments explain equal shares of the gender gap in weekly wages.

This study differs from previous literature in that it incorporates a more precise measure of work experience, and measures of cognitive ability, and non-cognitive traits to estimate the gender gap in weekly wages for full-time workers in Mexico. A gender gap in weekly wages that includes employer-provided benefits is also estimated to explore gender differences in compensations provided by employers. This study also contributes to the literature in that it disaggregates the contribution of four human capital characteristics to the gender gap in earnings.

I establish that for 2002, the gender gap in weekly wages is 15%, and the gender gap in weekly wages plus employer-provided benefits is 45%. Once the human capital variables are added to the earnings model, a quarter of gender gap in weekly wages is accounted for by endowments and another quarter by returns. However, only the share attributable to returns remains statistically significant. Most of the differences in returns and endowments, among male and female workers, attributable to human capital characteristics are explained by work experience. The main difference is that human capital characteristics explain 50% of the share attributable to returns, while human capital explains 15% of the share attributable to endowments. This result suggests differences in human capital endowments between men and women do not account for much of the wage gap, however male and female workers are compensated differently over those gaps in work experience.

Fourth, the explanatory power of human capital characteristics on gender differences in the returns is 15%, while human capital explains 50% of gender differences in endowments. Fifth, work experience accounts for half of the human capital's share in endowments (7 out of 15 percentage points) and half of the share in returns (26 out of 49 percentage points).

Disaggregated results showed that adding work experience, cognitive ability, and non-cognitive traits increases the share of human capital traits in explaining differences in endowments from 3% to 15%. The inclusion of these variables also increases the share of human capital variables in explaining differences in returns from 8% to 49%. This supports the hypothesis that incorporating more precise measures of work experience, cognitive ability and non-cognitive traits is key to calculate an accurate gender wage gap.

Wider differences in the gender gap in weekly wages plus employer-provided benefits portray the disparities in work benefits among male and female full-time workers. This suggests that women (in particular mothers) may be in jobs that offer different rewards (e.g. more flexible schedules) rather than jobs that provide higher additional benefits (medical benefits and different types of allowances), resulting in a wider income gap when these additional measures of work benefits are added up. However, as previously discussed, job amenities are not accounted in the benefits' package, thus results might overestimate the actual differential. In future work, an estimation of the family gap could show whether this hypothesis is holds.

*VI. Conclusion*

In this study it is established that there is a wider gap in weekly wages + employer provided benefits than in just weekly wages. Thus, on average, male workers have jobs with higher benefits in comparison to females. Results from the models analyzed showed that differences in levels and returns to work experience account for a significant share of the differences in earnings among male and female workers. As previously mentioned, women have weaker job attachments, and in consequence less work experience. Government programs that help working mothers to stay in the labor market have been implemented over the last few years and it would be interesting to compare the differences in earnings among a new cohort of women who have received higher maternal benefits.

**References**

Altonji, Joseph G. and Blank, Rebecca. 1999. “Race and Gender in the Labor Market,” in Orley Ashenfelter and David Card eds, *Handbook of Labor Economics*, vol. 3, Elsevier Science B.V.

Atal, Juan Pablo, Hugo Ñopo and Natalia Winder. 2010. “New Century, Old Disparities. Gender and Ethnic Wage Gaps in Latin America.” IZA Discussion Paper No. 5085.

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

Brown, Cynthia, Jose A. Pagán and Eduardo Rodriguez Oreggia. 1999. ‘‘Occupational Attainment and Gender Earnings Differentials in Mexico,’’ *Industrial and Labor Relations Review* 53, (1): 123–35.

Budig, Michelle J. and Paula England. “The Wage Penalty for Motherhood.” *American Sociological Review*, 2001, 66:204–25.

Cawley, John, James J. Heckman, and Edward J. Vytlacil. 2001. Three observations on wages and measured cognitive ability. *Labour Economics* 8, no. 4:419–42.

Chernozhukov, Victor, Fernandez-Val, Ivan, Galichon, A. 2010. “Quantile and probability curves without crossing”. *Econometrica.* 78, 1093–1126.

Cotton, Jeremiah. 1988. “On the decomposition of wage differentials”, *Review of Economics and Statistics*, 70, 236-243.

Cruces, Guillermo and Sebastian Galiani. 2007. “Fertility and Female Labor Supply in Latin America: New Causal Evidence.” *Labour Economics* 14(3): 565-573.

Deutsch, R. et al. 2004. “Working within Confines: Occupational Segregation by Gender in Costa Rica, Ecuador and Uruguay.” In: C. Piras, editor. *Women at Work: Challenges for Latin America*. Washington, DC, United States: Inter-American Development Bank.

DiNardo, John. 2002. “Propensity Score Reweighting and Changes in Wage Distributions.” Mimeo, University of Michigan. [Section IV].

DiNardo, John, Nicole M. Fortin and Thomas Lemieux. 1996. “Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach”. *Econometrica* , 64 (5): 1001-1044.

Dolton, P. J., and G. H. Makepeace. 1986. “Sample Selection and Male-Female Earnings Differentials in the Graduate Labour Market.” *Oxford Economic Papers* 38: 317–341.

Fortin, Nicole. 2008. “The gender wage gap among young adults in the United States: the importance of money vs. people.” *Journal of Human Resources*, 43 (4), 886–920.

Heckman, James. 1976. “The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models.” *Annals of Econometrics and Social Measurement.* 5: 475–492.

———. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica*. 47: 153–161.

Heckman, James, Jora Stixrud, and Sergio Urzua. 2006. “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior,” *Journal of Labor Economics*, 24, 411–482.

Heckman, James, Sergio Urzua and Edward Vytlacil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity," *The Review of Economics and Statistics*, MIT Press, 88(3): 389-432.

Hertz, T. et al. 2008. “Wage Inequality in International Perspective: Effects of Location, Sector, and Gender.” ESA Working Paper 8/08. Rome, Italy: Food and Agriculture Organization of the United Nations, Agricultural and Development Economics Division (ESA).

Jenkins, S. 1994. “Earnings discrimination measurement: A distributional approach,” Journal of Econometrics, vol. 16, no.1, pp. 81-102

Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce. 1991. “Accounting for the slowdown in black-white wage convergence”. In: Kosters, M.H. (Ed.), Workers and Their Wages: Changing Patterns in the United States. American Enterprise Institute,Washington.

Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce. 1993. “Wage inequality and the rise in returns to skill”. *Journal of Political Economy.* 101, 410–442.

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

Light, Audrey, and Manuelita Ureta. 1995. “Early-Career Work Experience and Gender Wage Differentials.” *Journal of Labor Economics*, 13(1): 121–54.

Lim, L.L. 2002. “Female Labour-force Participation.” Geneva, Switzerland: International Labour Organization, Gender Promotion Programme (GENPROM).

Machado, Jose F. and Jose Mata. 2005. “Counterfactual decomposition of changes in wage distributions using quantile regression”. *Journal of Applied Econometrics.* 20, 445–465.

Machado, Cecilia. 2010. Selection, Heterogeneity and the Gender Wage Gap. Columbia University, Economics Department. mimeo.

Manning, Alan and Joanna Swaffield. 2008. “The gender gap in early-career wage growth.” *Economic Journal*, 118 (530), 983–1024.

Manning, Alan, Saidi, Farzad, 2010. "Understanding the gender pay gap: what's competition got to do with it?" *Industrial and Labor Relations Review*, 63 (4), 681–698.

Meza, Liliana. 2001. "Wage Inequality and the Gender Wage Gap in México", *Economía mexicana*, nueva época, núm. 2, México, CIDe, pp. 291–323

Neuman, Shoshana and Ronald Oaxaca. 2004. Wage decompositions with selectivity-corrected wage equations: A methodological note. Journal of Economic Inequality 2: 3–10.

Ñopo, Hugo and Alejandro Hoyos. 2010. "Evolution of Gender Wage Gaps in Latin America at the Turn of the Twentieth Century: An Addendum to "New Century, Old Disparities." IZA Discussion Papers 5086, Institute for the Study of Labor (IZA).

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14 (October): 693-709.

Oaxaca, Ronald and Michael R. Ransom. 1994. "On Discrimination and the Decomposition of Wage Differentials," *Journal of Econometrics* 61 (1994), 5–21.

Pagán, José A. and Susana M. Sánchez. 2000. "Gender Differences in Labor Market Decisions: Evidence from Rural Mexico," *Economic Development and Cultural Change* 48: 619–37.

Pagán, J.A., and Ullibarri, M. 2000. "Group Heterogeneity and the Gender Earnings Gap in Mexico." *Economía Mexicana* (New Series) 9(1): 23-40.

Popli, Gurleen. 2008. *"Gender wage discrimination in Mexico: A distributional approach."* Working Paper. Department of Economics, University of Sheffield.

Reimers, C. W. 1983. "Labor Market Discrimination Against Hispanic and Black Men." *The Review of Economics and Statistics*. 65: 570–579.

Rendón, T. 2003. "Empleo, segregación y salarios por género." In: E. de la Garza and C. Salas, eds. *La Situación del Trabajo en México.* Mexico City, Mexico: Plaza y Valdés.

Rendón, T., and V.M. Maldonado. 2004. *Vínculos entre trabajo doméstico, segregación ocupacional y diferencias de ingreso por sexo, en el México actual*. Mexico City, Mexico: Instituto de Estudios del Trabajo.

Sánchez, Susan. 1998. "Gender Earnings Differentials in the Microenterprise Sector: Evidence from Rural and Urban Mexico." Washington, DC, United States: World Bank, Latin America and the Caribbean Region, Finance, Private Sector, and Infrastructure Sector Unit

Tenjo, J. 1992. "Labor Markets, the Wage Gap and Gender Discrimination: The Case of Colombia." In: G. Psacharopoulos and Z. Tzannatos, editors. *Case Studies on Women's Employment and Pay in Latin America.* Washington, DC, United States: World Bank.

Tenjo, J., R.R. Medina, and L.F. Bernat. 2004. "Evolution of Salary Differences between Men and Women in Six Latin American Countries." In: C. Piras, editor. Women at Work: Challenges for Latin America. Inter-American Development Bank.

Waldfogel, Jane. 1997. "The Effects of Children on Women's Wages." *American Sociological Review*, 62:209-17.

**Appendices**

## Appendix I. Oaxaca-Ransom decomposition

The Oaxaca-Ransom decomposition methodology consists in estimating three separate constrained linear regressions. One regression for females, one for males and a pooled regression that includes gender intercept shifts along with an identification restriction and constraints for each categorical variable (Fortin, 2008). For simplicity, I will show a more general version of the model for the decomposition analysis, following a similar approach to that of Fortin (2008) and Oaxaca and Ransom (1994, 1999).

$$Y_{ig} = \beta_{0g} + \beta_i \underline{X}_{ig} + e_{ig}, \qquad g=f,m,p \tag{1}$$

Assuming that the expected value of the error term is zero:

$$\bar{Y}_{im} - \bar{Y}_{if} = \bar{X}_m\hat{\beta}_m - \bar{X}_f\hat{\beta}_f + (\hat{\beta}_{0m} - \hat{\beta}_{0f}) \tag{2}$$

Where $\Delta X = \bar{X}_m - \bar{X}_f$ and $\Delta\beta = \hat{\beta}_m - \hat{\beta}_f$

Then, $\bar{Y}_{im} - \bar{Y}_{if} = \Delta X\hat{\beta}_m + \bar{X}_f\Delta\beta + (\hat{\beta}_{0m} - \hat{\beta}_{0f})$ (3a)

Analogously:

$$\bar{Y}_{im} - \bar{Y}_{if} = \Delta X\hat{\beta}_f + \bar{X}_m\Delta\beta + (\hat{\beta}_{0m} - \hat{\beta}_{0f}) \tag{3b}$$

Using a non-discriminatory pooled wage structure, income differences can be expressed as follow (Oaxaca and Ransom, 1994; Fortin, 2008):

$$\bar{Y}_{im} - \bar{Y}_{if} = \Delta X\hat{\beta}_p + \left[\bar{X}_m(\hat{\beta}_m - \hat{\beta}_p) + (\hat{\beta}_{0m} - \hat{\beta}_{0p})\right] - \left[\bar{X}_f(\hat{\beta}_f - \hat{\beta}_p) + (\hat{\beta}_{0f} - \hat{\beta}_{0p})\right] \tag{4}$$

Where $\left[\bar{X}_m(\hat{\beta}_m - \hat{\beta}_p) + (\hat{\beta}_{0m} - \hat{\beta}_{0p})\right]$ represents the “advantage of men” and $\left[\bar{X}_f(\hat{\beta}_f - \hat{\beta}_p) + (\hat{\beta}_{0f} - \hat{\beta}_{0p})\right]$ represents the “disadvantage of women.”

## Appendix II. Characteristics of full-time workers

**Table 1. Descriptive statistics of full time workers of 25-54 years of age with information on *income***

| | All | | Women | | Men | |
|---|---|---|---|---|---|---|
| *Employment variables* | | SE | | SE | | SE |
| Log weekly income | 4.35 | (0.003) | 4.34 | (0.018) | 3.89 | (0.083) |
| Weekly hours worked | 51.71 | (0.003) | 50.29 | (0.006) | 52.29 | (0.004) |
| Yearly weeks worked | 48.43 | (0.003) | 48.67 | (0.005) | 48.34 | (0.003) |
| Work experience | 19.29 | (0.003) | 14.56 | (0.005) | 21.23 | (0.003) |
| Formal contract | 0.365 | | 0.437 | | 0.335 | |
| *Cognitive / non-cognitive traits* | | | | | | |
| Cognitive ability (IQ score) | 51.413 | (0.008) | 50.036 | (0.015) | 51.972 | (0.009) |
| Feelings of poor performance | 0.193 | | 0.269 | | 0.162 | |
| Feelings of insecurity | 0.182 | | 0.276 | | 0.143 | |
| Pessimistic attitude | 0.203 | | 0.271 | | 0.176 | |
| *Marital Status* | | | | | | |
| Married | 0.825 | | 0.498 | | 0.958 | |
| Single | 0.104 | | 0.297 | | 0.026 | |
| Divorced/separated | 0.070 | | 0.205 | | 0.016 | |
| *Education* | | | | | | |
| Incomplete primary | 0.041 | | 0.024 | | 0.048 | |
| Elementary or some secondary | 0.368 | | 0.348 | | 0.376 | |
| Secondary or some high school | 0.296 | | 0.320 | | 0.287 | |
| Complete high school | 0.137 | | 0.141 | | 0.135 | |
| Some college and beyond | 0.157 | | 0.166 | | 0.153 | |
| *Fertility* | | | | | | |
| No children | 0.180 | | 0.212 | | 0.167 | |
| One child | 0.100 | | 0.156 | | 0.077 | |
| Two children | 0.205 | | 0.183 | | 0.214 | |
| Three or more | 0.515 | | 0.448 | | 0.542 | |
| *Other characteristics* | | | | | | |
| Head of household | 0.769 | | 0.232 | | 0.988 | |
| Urban residence | 0.846 | | 0.910 | | 0.820 | |
| Actual observations | 2,870 | | 804 | | 2,066 | |
| Weighted observations | 10,148 | | 2,931 | | 7,217 | |

Note: 1. Income was converted to US dollars using the average exchange rate for year 2002 of 9.46 pesos per dollar. 2. Weighted observations are in thousands.

## Appendix III. Gender gaps in wage and income (including occupation dummies)

In models that incorporate occupation dummies and correct for selection into employment, the share of the gender gap in weekly wages explained by endowments is 17%, the share explained by returns is 30% and 53% remains unexplained. The share of the gap in weekly income explained by endowments is 5%; the share explained by returns is 48% and 47% remains unexplained.

**Table 3. Weekly *wages* -including occupations**

| | (1)<br>Age + age squared | (2)<br>Actual experience | (3)<br>Cognitive ability | (4)<br>Non-cognitive traits |
|---|---|---|---|---|
| Men | 4.39 | 4.39 | 4.39 | 4.39 |
| | (0.032) | (0.032) | (0.032) | (0.032) |
| Women | 4.24 | 4.27 | 4.27 | 4.24 |
| | (0.116) | (0.131) | (0.132) | (0.131) |
| **Difference** | 0.149 | 0.111 | 0.115 | 0.150 |
| | (0.120) | (0.135) | (0.135) | (0.135) |
| Decomposition of estimated differentials (in %) | | | | |
| Endowments | 12% | 15% | 14% | 17% |
| Returns | 43% | 27% | 27% | 30% |
| Unexplained | 45% | 59% | 59% | 53% |
| N | 841 | 841 | 841 | 841 |

Note: 1. Income is calculated in logarithmic terms and converted into USD. 2. Robust standard errors are clustered at the state level. 3. Significance levels: +p <0.10, *p<0.05, **p<0.01, ***p<0.001. 4. Decomposition of estimated differentials reflects the share of the gender gap in weekly income under each specification and adds up to 100%.

The main difference between the specification that includes or excludes controls for job occupations is that including occupations increases the share of the unexplained component in the gender gap in weekly wages. This result might be an indication that including occupation dummies in the earnings' model does not add precision to the estimates of earnings' differentials among men and women. Results also suggest that adding job occupations might be endogenous.

**Table 4. Weekly *earnings* (wage + employer provided benefits) -including occupations**

| | (1)<br>Age + age squared | (2)<br>Actual experience | (3)<br>Cognitive ability | (4)<br>Non-cognitive traits |
|---|---|---|---|---|
| Men | 4.34 | 4.34 | 4.34 | 4.34 |
| | (0.018) | (0.018) | (0.018) | (0.018) |
| Women | 3.99 | 3.93 | 3.92 | 3.91 |
| | (0.073) | (0.082) | (0.082) | (0.083) |
| **Difference** | 0.356*** | 0.413*** | 0.421*** | 0.428*** |
| | (0.075) | (0.084) | (0.084) | (0.085) |
| Decomposition of estimated differentials (in %) | | | | |
| Endowments | 11% | 7% | 6% | 5% |
| Returns | 40% | 48% | 48% | 48% |
| Unexplained | 49% | 45%*** | 46%*** | 47%*** |
| N | 3,300 | 3,300 | 3,300 | 3,300 |

Note: 1. Income is calculated in logarithmic terms and converted into USD. 2. Robust standard errors are clustered at the state level. 3. Significance levels: +p <0.10, *p<0.05, **p<0.01, ***p<0.001. 4. Decomposition of estimated differentials reflects the share of the gender gap in weekly income under each specification and adds up to 100%.