The Gender Wage Gap among Young Adults in the United States: The Importance of Money vs. People [1]

Nicole M. Fortin
Department of Economics
University of British Columbia

February 2007

Abstract

Using two single-cohort longitudinal surveys, the NLS72 and the NELS88, I investigate the impact of four non-cognitive traits—self-esteem, external locus of control, the importance of money/work and the importance of people/family–on wages and on the gender wage gap among these young workers. I find that gender differences in these non-cognitive factors, especially the importance of money/work, have a modest but significant role in accounting for the gender wage gap. In 1986, this role is almost as important in magnitude as the role of gender differences in labor market experience and job tenure combined. The analysis also shed some light on whether the gender convergence in pay over the last two decades of the 20th century comes from within-cohort changes or between-cohort changes. Methodologically, this paper proposes a correction to the Oaxaca-Blinder-Ransom decomposition that results in a truly decomposable approach compatible with the simple pooled regression that includes a gender dummy.

Keywords: Altruism, money, gender wage gap, regression-compatible Oaxaca-Blinder decomposition.

[1]I would like to thank Francine Blau, David Card, Peter Dolton, Thomas Lemieux, Markus Mobius, Nathaniel Wilcox, anonymous referees, and seminar participants at the 2002 and 2006 AEA Meetings, at the 2005 SOLE Meetings, at UC-Berkeley and at the CIAR Workshop on Social Interactions, Identity and Well-Being for comments on this and earlier versions of the manuscript. The early stages of the project benefited from Michael Coelli's diligent research assistance. Financial support was provided by SSHRC Grants #410-99-0959 and #501-2002-0118.

1. Introduction

After spectacular gains in the 1970s and 1980s, the gender wage gap has shown slower progress during the 1990s (O'Neill (2003); Blau and Kahn (2004)). Given the realized promises of the 1980s that convergence in human capital (O'Neill and Polachek (1993), O'Neill (2003)) would be key to the closing of the gender pay gap, this stagnation was disappointing. In the 1980s and 1990s, dramatic increases in educational attainment among younger cohorts allowed women to enter professions, especially managerial professions, previously dominated by men. These gains were concurrent with increases in labor market experience among older cohorts, which also contributed to the narrowing of gender gap (Blau and Kahn (2000)). Thus perhaps unsurprisingly, Blau and Kahn (2004) attribute the 1990s slowdown in the gender convergence in pay to factors, other than educational attainment and labor market experience, that contributed to changes in the unexplained gender pay gap. These factors include changes in labor market selectivity, changes in gender differences in unmeasured characteristics and labor market discrimination, as well as less favorable supply and demand shifts.[2]

Following a renewed interest in the impact of behavioral and personality traits on earnings (Bowles, Gintis, and Osborne (2001)), this paper attempts to move some of the typically unmeasured characteristics into the measurable dimension.[3] The main goal of the paper is to assess the contribution of non-cognitive factors in accounting for the gender pay gap among workers in their early thirties in the mid-1980s. A second goal of the paper is to shed some light on the sources of changes in the gender pay gap over the last two decades on the twentieth century among young workers.

The recent literature on personality and earnings has incorporated non-cognitive factors from the Rosenberg self-esteem and the Rotter locus of control scales (Goldsmith, Veum, and Darity (1997); Osborne Groves (2005); Heckman, Stixrud, and Urzua (2006); Manning and Swaffield (2005); Waddell (2006)), and from the Five-Factor Model (FFM) of personality structure (Mueller and Plug (2006); Nyhus and Pons (2005)).[4] Like the papers in this first group, I

[2]Much recent research on the gender pay gap has focused on the impact of labor market selectivity, an issue outside of the scope of the present paper.

[3]Early research in this area goes back to the late 1970s and early 1980s; a review can be found in Bowles, Gintis, and Osborne (2001) and Osborne Groves (2005).

[4]The Rosenberg self-esteem scale is a widely used measure of self-esteem. The Rotter scale indicates

consider the role of personality traits such as self-esteem and external locus of control, but the gender differences in these traits are not very large.

Because this paper is concerned with the gender wage gap, I put more emphasis on non-cognitive factors where there are important gender differences, namely those associated with differentiated gender roles. In this sense, this study complements that of Mueller and Plug (2006) which focuses on personality traits that are prototypically masculine and feminine. Consistent with the traditional roles of men as "breadwinners", responsible for the financial welfare of the family, and of women as "homemakers" or "caregivers", entrusted with the care of children and the support of relationships, I pay special attention to gender differentials in valuing money and work on the one hand, and in valuing people, usefulness to society, and family on the other.[5] These traits are also easily tied to the economic literature on the household division of labor (Becker, 1985), as argued in section 2. In addition, an important focus of the recent experimental literature has been on gender differences in altruism.[6]

There is ample anecdotal evidence to the effect that women give greater importance in their choice of job to its value to society and to personal interactions, and less importance to money and prestige than men, but few studies investigate this path directly.[7] Yet, a recent hypothesis in the gender pay gap literature, called the negotiating divide hypothesis (Babcock and Laschever (2003)), argues that women may have less of a sense of entitlement to higher wages and are less likely to initiate negotiation or bargain.[8]

I utilized data from two widely used high school-cohort surveys: the National Longitudinal Study of the High School Class of 1972 (NLS72) and the National Education Longitudinal Study of 1988/94 (NELS88). The longitudinal aspect of these surveys allows me to use non-cognitive

whether an individual is more external or fatalistic, believing that hard work and effort are not rewarded, but rather that luck, fate or good fortune determine success and achievement. The five personality traits included in the FFM are extroversion, agreeableness, conscientiousness, neuroticism and openness to experience.

[5]Various terminologies have been used to describe these potentially conflicting values: career-life choices, work-family balance, etc.

[6]The experimental literature also focuses on gender differences in risk taking and competiveness, since experimental methods are more appropriate to study these traits than career-life choices.

[7]Filer (1986) evaluates the impact of various personality traits and job characteristics, including "contribution of job to society", on occupational structure. Chevalier (2004) investigates the impact of gender differences in motivation on the gender wage gap in the United Kingdom among recent graduates.

[8]Säve-Söderbergh (2003) indeed finds that not only were Swedish female university graduates asking for lower wages, but they were also receiving lower counter-offers.

traits measured when the individuals were in their teens (or early twenties), in fact many years before the labor market data was collected. The wage data was recorded when the individuals were in their mid-twenties (in 1979 for the NLS72, in 2000 for the NELS88) and in their early thirties (in 1986 for the NLS72). The measured non-cognitive traits cannot suffer from a reverse causality problem since they are recorded before the respondents' wages are observed. Further there is relatively little time for these non-cognitive traits to affect the labor market outcomes of individuals in their mid-twenties by an indirect behavioral route, such a more altruistic mother taking longer maternal leaves, working part-time and getting lower pay then a career-oriented mother working full-time with minimal work interruptions and getting higher pay. I thus argue that for these young workers the impact of non-cognitive traits is relatively untainted by direct feedback effects from the labor market, rather these traits can be thought off as pre-determined factors.[9] While considering these younger workers presents an advantage with respect to the identification of the impact of non-cognitive traits, it is a disadvantage in terms of a deeper understanding of the gender pay gap in the whole economy. Finally to the extent that there are cohort effects in the gender pay gap, looking at single cohorts rather than cross-sectional data arguably provides a cleaner assessment of changes in the gender pay gap within and between cohorts.[10]

I assess the importance of these non-cognitive traits on wages, appealing to a hedonic wage function framework, and of gender differentials in these traits on the mean gender wage gap. Because of the nature of the non-cognitive traits, I cannot claim to identify the "causal" effects of these factors on wages, rather I am interested in their impact as omitted variables on potential biases in the gender dummy of a standard wage regression.[11] To assess the impact of the traits on gender wage differentials, I propose a correction to the Blinder-Oaxaca-Ransom wage gap decomposition that makes it fully compatible with the simple pooled regression that includes a gender dummy, and for which standard errors are easily computed. The correction also solves the problems resulting for the choice of the non-discriminatory wage structure (Oaxaca and

[9]This case, of course, is not as clear for workers in their early thirties. In their mid-twenties, women in my samples have on average about 1 year of parental experience, but more than 5 years of parental experience in their early thirties.

[10]On this issue, see for example Blau and Kahn (2000) and Weinberger and Kuhn (2006).

[11]To simplify the exposition, at times the word "effect" or "impact" are used when what is estimated is merely a partial correlation.

Ransom (1994)) and includes recent propositions (Gardeazabal and Ugidos (2004)) to correct the identification problem caused by the left-out reference group, acknowledged in Oaxaca and Ransom (1999).

The main finding of the paper is of a modest but significant role (1.9 log points out of 22.9 log points) for gender differences in non-cognitive traits, in particular valuing money/work and people/family (1.7 log points), to account for the gender wage gap in 1986 (among workers in their early thirties from the NLS72). This role is larger than the role of gender differences in educational attainment and cognitive skills (math score) combined, about 1.2 log points. It is almost as important in magnitude as the role of gender differences in labor market experience and job tenure (2.4 log points), which is however lower than typically found in cross-sectional data.

Considering changes in the gender pay gap, from 1979 to 2000, among workers in their mid-twenties, the raw gender pay has decreased by a 5.6 log points (from 23.7 to 18.1 log points). The human capital variables accounted for 2.8 log points of the gap in 1979 and -2.2 log points in 2000, thus accounting for -5.0 log points (-2.2−2.8) of the change in the gap. The non-cognitive traits accounted for an additional -0.6 log points (0.9−1.5) of the change. Essentially the gender convergence in educational attainment, cognitive skills, valuing money and work can explain away all of the gender convergence in pay between young workers from the NL72 and NELS88 cohorts. This result suggests that the clear progression across cohorts of women's preparation for the labor market is responsible for the decrease in the gender pay gap. The within-cohort decrease in the gender gap pay of less than 1 log points (from 23.7 log points to 22.9 log points for the NLS72 cohort) from age 25 to age 32 further suggests a rather small role for within-cohort relative earnings growth. These findings help clarify the deliberations of Blau and Kahn (2000) on which factor,—within-cohort, between-cohort changes or lower discrimination–, is more important in accounting for the gender convergence in pay among young workers.

The remainder of the paper is organized as follows. The next section reviews the literature on gender differences in work values and presents a hedonic wage function framework in which non-cognitive factors are thought to influence wages. Section 3 addresses detailed data issues and presents the wage regression results. Section 4 presents the update on the Oaxaca-Blinder-Ransom methodology, and decomposition results are discussed in section 5. Finally, I conclude

with a review of the findings in section 6.

2. The Influence of Work Values on Economic Behavior

Historically gender differences in work values, which were perceived as a rationalization for occupational gender segregation, have been de-emphasized in the sociological literature (Bielby and Bielby (1984); Rowe and Snizek (1995)). More recently, in an effort to account for the unexplained part of the gender pay gap and increasing returns to unobserved skills, there has been a renewed interest in economics for "soft" skills or non-cognitive traits.

Experimental studies in behavioral economics have noted important differences between men and women in individual attitudes towards altruism and greed (Andreoni and Vesterlund 2001), leadership and competitiveness (Gneezy, Niederle, and Rustichini 2003). Empirical studies have found a substantial impact of these and related traits on wages. Kuhn and Weinberger (2005) find sizeable positive returns, among white males, to leadership skills ranging from 4 to 24 percent and an even larger impact for workers in managerial occupations.

Another trait at the opposite end of the leadership spectrum is the "external locus of control", that is a belief that one's outcomes are controlled more by external forces than by internal forces (or that chance and luck are more important for success than one's own efforts). The effect of external locus of control on earnings, often together with the effect of self-esteem, has been investigated by many authors (Goldsmith, Veum, and Darity (1997), Osborne Groves (2005), Heckman, Stixrud, and Urzua (2006), Manning and Swaffield (2005), Waddell (2006) among others). For example, using the same surveys that I use, Waddell (2006) (using the NLS72) and Coleman and DeLeire (2003) (using the NELS88), find that low self-esteem and an external locus of control are associated with lower investment in human capital, among other negative indicators of future labor market performance. On the other hand, Goldsmith, Veum, and Darity (1997) who include both predicted values of self-esteem and locus of control in their wage regressions, find that only self-esteem was statistically significant.

Psychologists believe that work effort reflects motivation, which is governed in part by these personality factors. Individuals with an external locus of control believing that hard work and effort will be not rewarded are less likely to display high levels of effort on the job. On the positive side, self-esteem have been found to be strong predictor of job performance (Judge,

Tippie, and Bono 2001). Brockner (1988) argues that workers with high self-esteem will tend to be more confident about making decisions, will use their time more effectively and more productively. Psychologists also expect that higher wages will enhance self-esteem, but that poor self-esteem may lead to a decline in productivity and wages which may in turn erode a person's perception of self. Thus theories of self-esteem formation advanced by psychologists suggest that wages and self-perception may be jointly determined, which could lead to an endogeneity problem. To avoid this potential problem, I use measures of personality traits and features obtained many years before the wages are observed.

Because the gender differences in self-esteem and locus of control are relatively small, these commonly studied non-cognitive factors will prove to be less useful for my purpose of explaining the gender pay gap. I thus capture an additional dimension of motivation and work effort by including a measure of the importance of money/work. Psychologists (Furnham and Argyle (1998)) argue that workers beliefs about money are clearly related to their actual economic behavior. The attitude-behavior relation literature (Ajzen and Fishbein (1977), Tang, Kim, and Tang (2000)) further argues that those who do value money and want to make more money (a work-related attitude) may have a higher level of turnover (a work-related behavior) than those who do not. Workers for whom "making lots of money" is "very important" may be more likely to initiate bargaining talks with employers or change jobs for more money. Similarly, workers for whom "success on the job" in "very important" may be more likely to display to work-related behaviors that lead to promotions and higher wages. Kirkcaldy and Furnham (1993) also found that attitudes toward money are predicted by competitiveness or the motive to outperform others. To the extent that there are gender differences in competitiveness, these will translate into gender differences in attitudes toward money.

Gender differences in the importance of people/family are well-known, but there are few attempts to study directly how they affect work-related behaviors.[12] In their analysis of changes in the gender pay gap, Borghans, ter Weel, and Weinberg (2006) model the importance of interpersonal interactions in the context of an assignment model, where there are positive return to interpersonal skills. Here, I combine this dimension of altruism, the importance of

[12]Daymont and Andrisani (1984) who focus their analysis of the gender pay gap on the choice of college major also study the effect of these values.

"working with people rather than things", with one that goes beyond interpersonal skills, the importance of "a job usefulness to others or society" and focus on the implications of women's higher demand for altruistic rewards for earnings. Many more women than men indeed choose lines of work, such as education and health care, where they perform many tasks for the love of others and reap lower market rewards working for public and non-profit organizations.[13] Women are also more likely than men to use their time to volunteer in organizations with some altruistic purpose, such as schools, religious organizations and hospitals.[14]

I show how, in a classic hedonic wage model Rosen (1974), the importance of money/work can have a positive impact on earnings, while the importance of people/family might can have a negative impact. By contrast with the human capital model, the hedonic wage function model is job-oriented and indeed assume the workers choose and change jobs on the basis of the characteristics of the job. The importance of money/work, self-esteem and external locus of control are preferences and personality features that influence the choice of wage level and the level of effort and responsibility linked to the job. The importance of people/family influence the choice of altruistic rewards of the job.

These non-cognitive traits are thought to impact wages through a hedonic wage function resulting from a matching equilibrium between the employers' wage and amenity offers, and the workers' utility maximizing choices of job characteristics. The worker/consumer i maximizes utility over wages, effort and responsibility, and altruistic rewards

$$U_i = U_i(W_i, R_i, A_i), \tag{1}$$

where W_i is the wage rate, R_i is the level of effort and responsibility exerted by the worker/ consumer, and A_i is the level of altruistic rewards, which is derived from both market and non-market activities: $A_i = A_i^M + A_i^H$. The marginal utility of wages and altruistic rewards is positive, but its sign with respect to effort and responsibility is ambiguous. To the extent that leadership and ambition denote different flavors of effort, one can see those as positively linked

[13]There is some evidence that men are becoming more like women on this "usefulness to others" dimension, as found across birth cohorts in the European Value Survey (Fortin 2005).

[14]On the other hand, men are more likely to volunteer as coaches in youth sports teams, in unions, professional organizations, social and political clubs, where they can use and perfect leadership skills.

to utility, but when effort and responsibility exceed some threshold they entail a disutility.

Following Becker (1985) who argues that there is a limited amount of time and energy to be split between market work and homemaking, I assume that there is a total amount of effort and responsibility that an individual can devote to either market or non-market activities, $R_i = R_i^M + R_i^H \leq \bar{R}_i^T$, where $\bar{R}_i^T$ denotes to the total amount available, fixed for each individual, and R_i^H denotes the amount of effort and responsibility allocated to non-market activities, like household work and volunteering activities. To the extent that women carry the double-shift of housework and market work more than men do, this imply that the marginal utility (disutility) of market effort and responsibility is greater(less) for men than for women, $\partial U_{im}/\partial R_{im}^M > \partial U_{if}/\partial R_{if}^M$, and more generally for individuals who value homemaking and volunteering more than others.

The utility function of individuals for whom "working with people" and feelings of "usefulness of a job to others and society" are more important relative to financial rewards will be characterized by a greater marginal rate of substitution between the altruistic rewards and wages than other workers. To the extent this trait characterizes women more than men:

$$\frac{\partial U_{if}}{\partial W_{if}} / \frac{\partial U_{if}}{\partial A_{if}} > \frac{\partial U_{im}}{\partial W_{im}} / \frac{\partial U_{im}}{\partial A_{im}}. \tag{2}$$

For these individuals, working for philanthropic non-profit organizations can take place at a wage cost.[15] Alternatively, individuals could accept a job with lower altruistic rewards and lower effort, but volunteer in altruistic organizations or become a part-time caregiver at home. Because of the quantity constraint on effort and responsibility, when the worker/consumer chooses a job vector $(\overset{*}{W}_i, \overset{*}{R}_i^M, \overset{*}{A}_i^M)$ to maximize his/her utility, he/she simultaneously chooses a level of non-market effort and responsibility, $\overset{*}{R}_i^H = \bar{R}_i^T - \overset{*}{R}_i^M$. Similarly, the individual's demand for the altruistic market amenity may depend on the level already provided by home and volunteering activities. These potential utility maximizing choices generate a job demand locus $\theta_i(W, R^M, A^M)$.

Employers are thought to be cost–minimizers, so that optimizing behavior generates an offer function, $\phi(W, R^M, A^M)$, indicating unit prices that employers are willing to pay on various

[15]For example, in the NELS88, 11.7 percent of women work for non-profit organizations by contrast with 4.4 percent of men.

combinations of wage, effort and altruistic amenities.[16] Because it is time-consuming, the altruistic amenity is by definition costly to provide.[17] It is assumed that employers pay higher wages for higher levels of effort and responsibility.

Subsumed in the worker/consumer job demand functions and the employers' offer functions are the human capital variables which are thought to help eliminate multiple equilibria.[18] I circumvent the identification problems (Ekeland, Heckman, and Nesheim (2004)) associated with hedonic wage functions by including information on usually unobserved tastes for job characteristics. I simply assume that one equilibrium is realized in the data and seek to broadly identify preferences from value statements and reported behaviors in a reduced-form approach. In the equilibrium assignment, we thus find a wage equation,

$$W_i = g(S_i, E_i) + h(R_i^M) + t(A_i^M). \tag{3}$$

where in addition the usual human capital component, $g(S_i, E_i)$, with S_i denoting schooling and E_i labor market experience, there is a component that rewards effort and responsibility $h(R_i^M)$ and another one that taxes altruistic amenities $t(A_i^M)$.

My aim is to first assess the impact of work values and altruistic preferences on wages, then I will consider the impact of gender differentials in these traits on the gender pay gap.[19] These non-cognitive variables will prove that have an impact on wages more distinct from cognitive factors that self-esteem and external locus of control. They also offer a unique possibility to improve the empirical analysis of models of the household division of labor and of wage determination by incorporating variables that pre–date household formation and labor market decisions.[20]

[16]To the extent that some employers (e.g. hospitals) are exert monopsony power in some markets for altruistic jobs (e.g. nursing), the offer functions cannot be seen as perfectly competitive outcomes.

[17]For example, smaller classes are more rewarding to teach.

[18]I could have more explicitly conditioned the utility function on the human capital variables $U_i(W_i, R_i, A_i|S_i, E_i)$.

[19]In a previous version of the paper, I had shown that work values and altruistic preferences have behavioral implications of the expected signs on promotions (or training for career advancement), volunteering in either philanthropic (hospitals, youth organization, etc.) or leadership organizations (sports teams, rotary club, etc.) among men and women, and on the incidence of working part-time or long maternal leave (or the number of dependents) among women.

[20]Although the influences of neurological, environmental, psychological and social factors have been implicated with varying degrees in the development of gender roles, it is generally recognized that gender

3. Data and Descriptive Statistics

To avoid problems of ex post rationalization, I appeal to two single-cohort longitudinal surveys conducted by the National Center for Education Statistics (NCES) to study education issues. The National Longitudinal Study of the High School Class of 1972 (NLS72) known as the grandmother of educational longitudinal surveys, is widely used (Adelman (1994)) outside of the education field. For example Brown and Corcoran (1997) use it, in conjunction with the SIPP, to study the impact of gender differences in the field of the highest degree on the gender wage gap.[21] The NLS72 interviewed students in their senior year of high school in the spring of 1972, with follow-up surveys conducted in 1973, 1974, 1976, 1979, and finally in 1986. While the sample is not fully nationally representative because it ignores students who dropped out before grade 12, Krueger and Dale (2002) argue that it is representative for college graduates.

The NLS72 contains answers to many questions on what is important in life and in selecting a career. For example, respondents are asked, with reference to selecting a career, about the importance of *"Making lots of money"* and *"The chance to be a leader"*, which speaks to the importance of money and ambition in work values. There are also asked about the importance of *"Opportunities to be helpful to others or useful to society"*, *"Opportunities to work with people rather than things"*, this speaks to the altruistic and social aspect of work. The survey also contains questions from the well-known Rosenberg and Rotter scales. Many of these questions are asked repeatedly in each follow-up. Because of the longitudinal aspect of the survey, it is possible to use preferences expressed in a time period that precedes the behavior of interest. Wage information was recorded in the fourth follow-up in 1979, when most respondents were age 25, and in the fifth follow-up in 1986, when most respondents were age 32. Note that in 1986 more respondents were likely to have completed their studies and there was more time for labor market experience to have an effect.

The National Education Longitudinal Study of 1988/94 (NELS88) interviewed students in eight grade in the spring of 1988. A sample of these respondents was then resurveyed through four follow-ups in 1990, 1992, 1994, and 2000, and a refreshed sample was added in the 1990

and sex-role knowledge are acquired early in life (Ruble and Martin 1998).

[21] Here I omit field of study because it is missing for a sizeable portion (more than a quarter) of the sample.

and 1992 waves to ensure a representative sample of high school sophomores and seniors in those years. The wage data was recorded in 2000 when most respondents were age 24, which leaves little time for labor market intermittency to have an effect. To enhance the comparability between the two samples, I select only respondents from the NELS88 who reached the 12th grade. The NELS88 was explicitly administered with the intent of maintaining comparability with the NLS72, so the major components of the two studies are very similar, including the components on values important in life.

Being very detailed and complex also means that these longitudinal surveys are at times difficult to code. In particular, there is detailed information about employment and unemployment spells, which are used to construct an actual work experience variable, but it is not straightforward to code.[22] I retain individuals from the NLS72 (NELS88) who were working in 1979 or in 1986 (2000) and that have valid information on wage, educational attainment, high school math test, and questions from the Rosenberg and Rotter scales.[23]

The means of the human capital variables and other individual characteristics are reported separately by gender for each time period in Table 1. They show that, in the NLS72, when workers are in their mid-20s, women's educational attainment is similar to men's, but by their early 30's, more men than women have a post-graduate degree. In the NELS88, where the workers are in their mid-twenties, this pattern has been reversed with not only more women holding a college degree than men, but also more women holding a post-graduate degree. In these single cohorts of young workers, the differences in the labor market experience of women

[22]In the NLS-72, the information on these spells from October 1972 to October 1979 is available in the first to the fourth follow-up. The information from October 1979 to July 1986 is available in terms of spells of employment for up to four jobs and unemployment for up to eight spells, thus individuals with more than eight spells of unemployment were excluded. In the NELS88, detailed labor force status by month is available from June 1990 to August 1994. In 1997, 1998 and 1999, only information on whether the respondent was unemployed for more than 6 months was available.

[23]I retain only individuals whose wage was between \$1.00 and \$150 in 1986, and the equivalent in 1979 and 2000. The wage rate is computed to approximate as closely as possible an hourly wage measure. In 1986 and 2000, the earnings are reported either as hourly wages, weekly earnings, bi-weekly earnings, monthly earnings or annual earnings and flagged as such. The hourly wages are thus computed correspondingly as weekly earnings divided usual weekly hours of hour, bi-weekly earnings divided usual weekly hours times 2, monthly earnings divided usual weekly hours times 52/12, annual earnings divided by usual weekly hours times 52. In 1979, earnings were recorded only as weekly earnings, wages are thus computed as weekly earnings divided by weekly hours. About 4% of the observations are in the edited categories (with codes in 800 and 1000 numbers) and are recoded using the information in the codebook.

and men are relatively small (at most 13%). There is thus limited scope for labor market intermittency to account for a large share of the gender wage gap. This contrasts with the recent work of O'Neill and O'Neill (2005) were differences in labor market experience and work interruptions for family reasons are found to account for more than half of the gender wage gap.

Cognitive skills are captured using high school math score measured when the respondents were in 12th grade, both in the NLS72 and NEL88. Since the samples exclude earlier drop-outs, this measure of cognitive skills generally precede the realizations of the educational attainment observed. I standardize the measure to have zero mean and a variance of one over the entire sample in each case. Thus in NSL72, the measure for women is negative while the measure for men is positive. This male advantage seemed to have shrunk somewhat in the NELS88.[24] Table 1 also reports the means and standard deviations of demographic characteristics commonly used in wage regressions.

Table 2 reports the descriptive statistics of the four composite non-cognitive traits used in the analysis: self-esteem, external locus of control, the importance of money/work and the importance of people/family. I report the means and standard deviations of the variables by gender, as well as the differences between the male and female averages. First note that all the composite variables are scaled between 0 and 1 inclusively, and that the standard deviations of these variables are in the same range, from 0.141 to 0.286. Also reported in the Table is Cronbach's alpha, a measure of reliability that indicate whether the responses to each question comprising the composite are highly correlated.[25] It is computed over the sample of men and women combined.

The measures of self-esteem and external locus of control are constructed using answers to a subset of four or five questions, asked in the surveys, from the widely used Rosenberg self-esteem scale and the Rotter locus of control scale. The description of the precise statements used and the aggregation method is in the Appendix. As explained in the Appendix, the responses are recoded in terms of a binary variable (1 or 0) reflecting agreement or disagreement with the

[24]This is consistent with the literature on the gender pay in math tests.

[25]The reliability of a scale (or composite variable) is one of three principles (with convergence and stability) used by psychologists to assess their validity. The measure of reliability is based on the correlations between the items that make up the composite, relative to the variances of the items.

statements. In these samples, the reliability indexes for self-esteem and external locus of control seem relatively low, they are however higher than the ones of Waddell's (2006) Attitude and Self-Esteem indexes, also computed with the NLS72 data, which range from 0.36 to 0.50.

The measures of the importance of money/work and of people/family are constructed using subjective questions about values generally important in life, which are asked in the base year and in each follow-up of the NLS72, and in selected follow-ups of the NELS88 and, about values important in selecting a career, which is asked only in the base survey of the NLS72. The values selected to capture the importance of money/work include questions about the importance in selecting a career of *"Making of lot of money"* and of *"The chance to be a leader"* and about the importance in life of *"Being successful at work"* and *"Having lots of money"*. As explained in the Appendix, the composite is constructed as means of the underlying answers coded 1 if "very important", 0 otherwise. Recoding the variable to capture the intensity of preferences is consistent with the idea that workers who really think that money is "very important" are more likely to initiate bargaining talks with employers or change jobs for more money, than workers who simply think that money is "somewhat important".[26]

The importance of people/family composite is constructed using questions about the importance in selecting a career of *"Opportunities to be helpful to others or useful to society"* and *"Opportunities to work with people rather than things"*, and in life of *"Helping other people in the community"*, the *"Ability to give children better opportunities"*, and *"Living close to parents and relatives"*. This composite is constructed using the same recoding as the previous variable. The reliability indexes for these two composites are in the same range as the indexes for the more commonly used self-esteem and locus of control composites. This nevertheless indicates the potential for measurement error in these values, which should lead to an attenuation bias and weaker results than more precise measures would entail.

The gender differences (Δ) in the importance of money/work and people/family, ranging from $|0.043|$ to $|0.065|$, are much larger than the gender differences in self-esteem and locus of control, which range from $|0.008|$ to $|0.026|$. The gender differences in the importance of money/work and of people/family are also much more stable across surveys, despite the dif-

[26]In a previous version of the paper, I indeed found that valuing money was a significant determinant of "having changed for a better job".

ferences across surveys in the questions underlying the composite.[27] Ultimately, these gender differences will have substantial explanatory power towards the gender pay gap. Finally, note that the correlation between the four composite variables is relatively small. If one wanted to construct a grand composite of the four non-cognitive measures, the resulting Cronbach's alpha would be very low around 0.25.[28]

4. Decomposition of Gender Wage Differentials

In this paper, I propose and implement a modification to the Oaxaca-Blinder decomposition that addresses two non-invariance problems of the methodology; these are also discussed in Oaxaca and Ransom (1994) and Oaxaca and Ransom (1999). The first one is the well-known fact that the portion of the gender wage gap explained by differences between male and female characteristics is not invariant to whether the male or female wage structure is chosen as the reference wage structure. Here, I propose to use as reference wage structure the one under which the advantage of men is equal to the disadvantage of women. The second non-invariance problem occurs in the presence of categorical variables and entails that portions of the gender pay gap attributable to gender differences in specific categorical variables will generally not be invariant to the choice of the omitted category. This problem is easily solved by appealing to a restricted least squares estimator. One attractive feature of this modified decomposition is that it is fully compatible with the usual pooled wage regression that includes a dummy for the disadvantaged group.

Let the following denote the log wage equations estimated separately for samples of males (m) and females (f), and for males and females pooled together (p)

$$\ln w_{ig} = \beta_{0g} + \mathbf{X}_{ig}\boldsymbol{\beta}_g + \epsilon_{ig}, \qquad g = f, m, p \tag{4}$$

where the $\mathbf{X}_i$ is a $1 \times K$ vector of explanatory variables and $\boldsymbol{\beta}$ is $K \times 1$ vector of coefficients.

[27] As explained in the Appendix, the greater relative weight of the *"Ability to give children better opportunities"*, which is very important for 70% to 90% of respondents, explains the large value of people/family very important in the NELS88.

[28] Given the regression coefficients are in effect partial correlations, any cross-correlations between the composites will be reflected in changes in the regression coefficients, which are discussed below. It will show that the more important correlations are with math scores.

Here, I am careful to distinguish the intercept from the other explanatory variables. As shown below, this is important in the presence of categorical variables, including the gender dummies.

Under the usual assumption $E(\epsilon_{if}) = E(\epsilon_{im}) = 0$, the difference between the means of male and female log wages will be

$$\overline{\ln w_m} - \overline{\ln w_f} = \overline{\mathbf{X}}_m \widehat{\boldsymbol{\beta}_m} - \overline{\mathbf{X}}_f \widehat{\boldsymbol{\beta}_f} + (\widehat{\beta}_{0m} - \widehat{\beta}_{0f}). \tag{5}$$

In the familiar Oaxaca-Blinder (Oaxaca (1973); Blinder (1973)) decomposition, letting $\Delta\mathbf{X} = \overline{\mathbf{X}}_m - \overline{\mathbf{X}}_f$ and $\Delta\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}_m} - \widehat{\boldsymbol{\beta}_f}$, this expression can be written as either,

$$\overline{\ln w_m} - \overline{\ln w_f} = \Delta\mathbf{X}\widehat{\boldsymbol{\beta}_m} + \overline{\mathbf{X}}_f\Delta\boldsymbol{\beta} + (\widehat{\beta}_{0m} - \widehat{\beta}_{0f}) = \Delta\mathbf{X}\widehat{\boldsymbol{\beta}_f} - \overline{\mathbf{X}}_m\Delta\boldsymbol{\beta} - (\widehat{\beta}_{0m} - \widehat{\beta}_{0f}), \tag{6}$$

depending on whether the male or female wage structure is chosen as the reference or non-discriminatory one, and where the first terms ($\Delta\mathbf{X}\widehat{\boldsymbol{\beta}_g}$) capture differences due to characteristics while the other terms denote differences to the returns to those characteristics.

A first problem with the Oaxaca-Blinder decomposition is that in practice, the choice of the non-discriminatory wage structure (either male or female) will yield different results (Cotton (1988); Oaxaca and Ransom (1994)). One proposition is to use the pooled wage structure as the non-discriminatory one (Neumark (1988); Oaxaca and Ransom (1994)) :

$$\overline{\ln w_m} - \overline{\ln w_f} = \Delta\mathbf{X}\widehat{\boldsymbol{\beta}_p} + [\overline{\mathbf{X}}_m(\widehat{\boldsymbol{\beta}_m} - \widehat{\boldsymbol{\beta}_p}) + (\widehat{\beta_{0m}} - \widehat{\beta_{0p}})] - [\overline{\mathbf{X}}_f(\widehat{\boldsymbol{\beta}_f} - \widehat{\boldsymbol{\beta}_p}) + (\widehat{\beta_{0f}} - \widehat{\beta_{0p}})], \tag{7}$$

where the second term in brackets is interpreted as the advantage of men and the third term in brackets as the disadvantage of women. An important problem with the Neumark method is that the pooled coefficients capture part of the "between" male and female effects, thus overstating the effects of variables with large gender differences. For example, in Table 2 of Neumark (1988), the effects of schooling (0.088) and union (0.172) in the pooled sample are larger than the same effects in either the male (0.072 and 0.138) or female (0.062 and 0.134) samples. This occurs because, as shown in Table 1 of Neumark (1988), males have higher schooling and unionization rates than women and thus their coefficients capture part of the

advantage of males.[29] In addition, if the advantage of men and the disadvantage of women are not equal, the wage structure of reference can hardly be called non-discriminatory.

A better alternative is to include gender intercept shifts, along with an identification restriction, in the regression of males and females pooled together[30]

$$\begin{aligned} \ln w_i &= \gamma_0 + \gamma_{0f} \cdot F_i + \gamma_{0m} \cdot M_i + \mathbf{X}_i\boldsymbol{\gamma} + \upsilon_i, \\ &\text{subject to } \gamma_{0f} + \gamma_{0m} = 0 \end{aligned} \tag{8}$$

where F_i is a female dummy and $M_i = 1 - F_i$ is a male dummy.[31] Then

$$\begin{aligned} \overline{\ln w_m} &= \widehat{\gamma}_0 + \widehat{\gamma}_{0m} + \overline{\mathbf{X}}_m\widehat{\boldsymbol{\gamma}} + E(\upsilon_i|F_i = 0), \\ \overline{\ln w_f} &= \widehat{\gamma}_0 + \widehat{\gamma}_{0f} + \overline{\mathbf{X}}_f\widehat{\boldsymbol{\gamma}} + E(\upsilon_i|F_i = 1), \end{aligned} \tag{9}$$

The resulting decomposition makes the assumptions on the error terms explicit,

$$\overline{\ln w_m} - \overline{\ln w_f} = \Delta\mathbf{X}\widehat{\boldsymbol{\gamma}} + (\widehat{\gamma}_{0m} - \widehat{\gamma}_{0f}) + [E(\upsilon_i|F_i = 0) - E(\upsilon_i|F_i = 1)]. \tag{10}$$

The first term is the familiar portion of the pay gap attributable to differences in characteristics and its standard error is computed as $(\Delta\mathbf{X})'\Sigma_{\boldsymbol{\gamma}}(\Delta\mathbf{X})$, where $\Sigma_{\boldsymbol{\gamma}}$ is the variance-covariance matrix of the $\widehat{\boldsymbol{\gamma}}$. The second term $(\widehat{\gamma}_{0m} - \widehat{\gamma}_{0f})$ correspond to the negative of the coefficient of a female dummy in a familiar wage regression on the pooled sample when male if the omitted category and the intercept thus includes the effect of the male dummy $(=\widehat{\gamma}_0 + \widehat{\gamma}_{0m})$. The last term vanishes,

$$[E(\upsilon_i|F_i = 0) - E(\upsilon_i|F_i = 1)] = Cov(\upsilon_i, F_i) = 0,$$

where the first equality follows from the fact that F is a binary variable (as in Angrist and

[29] A similar problem can be seen in the comparison of black males/white males in Table A2 of Oaxaca and Ransom (1994) with regards to the coefficient for managers and professionals (OC1).

[30] Brown and Corcoran (1997) incorporates a single gender dummy in their pooled regressions and proposes using the term $\Delta = \Delta\mathbf{X}(\widehat{\boldsymbol{\beta}_m} - \widehat{\boldsymbol{\beta}_f})$ as a measure that is invariant to the choice of omitted category. The interpretation of this term is that the elimination of differences in characteristics would lead to a pay increase Δ if estimated at the male returns instead of the female returns; but it is not straightforward.

[31] It is always possible to weight the female and male dummies by their percentage in the sample or to express variables as deviations from the pooled means to obtain a non-discriminatory wage structure that would maintain the same average wage where nepotism and discrimination to be eliminated.

Imbens (1995)), and the second one is a property of ordinary least squares. This decomposition is thus compatible with the conventional practice of including a dummy denoting the group of interest in a pooled regression to investigate the impact of belonging to a disadvantaged group (as in equation (14)). To the extent that $\widehat{\gamma}_{0m}$ is positive and $\widehat{\gamma}_{0f}$ is negative, it is appropriate to say that these expressions are representing the "advantage" of men and the "disadvantage" of women.

The decomposition can also be written in terms of the Neumark-Cotton decomposition as:

$$\overline{\ln w_m} - \overline{\ln w_f} = \Delta \mathbf{X}\widehat{\boldsymbol{\gamma}} + [\overline{\mathbf{X}}_m(\widehat{\beta}_m - \widehat{\boldsymbol{\gamma}}) + (\widehat{\beta}_{0m} - \widehat{\gamma}_0)] - [\overline{\mathbf{X}}_f(\widehat{\beta}_f - \widehat{\boldsymbol{\gamma}}) + (\widehat{\beta}_{0f} - \widehat{\gamma}_0)], \qquad (11)$$

If the coefficients $\widehat{\boldsymbol{\gamma}}$ and $\widehat{\gamma}_0$ truly represent a non-discriminatory wage structure, then the advantage of men will be equal the disadvantage of women. Under this proposed modification to the Oaxaca-Blinder decomposition, it can be easily verified that

$$\widehat{\gamma}_{0m} = [\overline{\mathbf{X}}_m(\widehat{\beta}_m - \widehat{\boldsymbol{\gamma}}) + (\widehat{\beta}_{0m} - \widehat{\gamma}_0)] \quad \text{and} \quad \widehat{\gamma}_{0f} = [\overline{\mathbf{X}}_f(\widehat{\beta}_f - \widehat{\boldsymbol{\gamma}}) + (\widehat{\beta}_{0f} - \widehat{\gamma}_0)]. \qquad (12)$$

and the standard errors on these two terms are obtained from the estimation of equation (8). The right-hand side expressions can be used to assign portions of the unexplained gender wage gap attributable to each variable k, that comes either from the advantage of men, $\overline{X}_{mk}(\widehat{\beta}_{mk} - \widehat{\gamma_k})$, or from the disadvantage of women, $\overline{X}_{fk}(\widehat{\beta}_{fk} - \widehat{\gamma_k})$. Empirically, the results from this regression-compatible decomposition will be in between the female and male based decomposition, like the Cotton decomposition, but not identical. A comparison of results (reported in Appendix Table A3) from the five most common decomposition using as reference wage structure the male coefficients, the female coefficients, the pooled coefficients (*à la* Neumark), the Cotton (weighted average of male and female) coefficients, and the regression-compatible coefficients makes this point and shows that the Neumark method gives extreme results.

A second problem (Oaxaca and Ransom 1999) is that the assignment of the explained part of the gender wage gap to specific variables is not invariant to the choice of the left-out category in the case of categorical variables. One solution suggested in Gardeazabal and Ugidos (2004), who however prefer to use a derived restriction, is to include all categories and impose

a zero-sum restriction on the estimated coefficients of each categorical variable. This is easily implemented via restricted least squares. Letting $\mathcal{C}$ denote the set of categorical variables, and C_k the number of categories for variable k, equation (8) is simply estimated subject to the restrictions:

$$\gamma_{0f} + \gamma_{0m} = 0 \qquad \text{and} \qquad \sum_{j=1}^{C_k} \gamma_{jk} = 0, \qquad k \in \mathcal{C} \tag{13}$$

The categorical restrictions are also imposed in the estimations of the gender specific log wage equations (4).[32] Under these restrictions, the intercepts β_{0m}, β_{0m} and γ_0 will not be contaminated by the effects of any left-out category and it will be possible to assign portions of the explained part of the gender wage gap to specific variables, as is done below.

In summary, the decomposition methodology consists in estimating three separate constrained linear wage regressions: one for each gender and a third regression, estimated on the pooled sample of men and women, that include gender intercept shifts along with an identification restriction. In addition to the constraint on the gender parameters in the pooled regression, additional constraints for each categorical variable are required.[33]

5. Estimation Results

With the estimation of three wage structures (male, female, and non-discriminatory), the proposed decomposition procedure (like that of Neumark (1988)) is more involved than the traditional Oaxaca-Blinder decomposition, yet it is easily performed with modern software. It is also compatible with the regression estimated on the pooled sample where the coefficient of the female dummy is interpreted as a measure of the unexplained gender wage gap, a regression which has long been the simplest way to study wage differentials.

Tables 3a-c report the wage regression results for this standard specification that includes a female dummy, and treat males as the omitted category,

$$\ln w_i = \delta_0 + \delta_{0f} \cdot F_i + \mathbf{X}_i \boldsymbol{\gamma} + \upsilon_i, \tag{14}$$

[32]Restrictions need to apply to all categorical variables, including binary variables such as part-time and full-time. These are easily implemented using a constrained linear regression.

[33]Software in STATA to implement this "Regression-Compatible Oaxaca-Blinder" decomposition is available on the author's web site.

where F is a female dummy, and $\boldsymbol{\gamma}$ is $K \times 1$ vector of coefficients. The coefficients from this standard pooled regression are expected to lie in between the female and male coefficients estimated from the gender specific samples. Rather than reporting all three sets of coefficients, I simply indicate the main source of discrepancies between the coefficients where applicable.

The comparison of Tables 3a (NLS72 in 1979) and 3b (NLS72 in 1986) will reflect within-cohort changes (over time and over the life-cycle) in the determinants of wages, while the comparison of Tables 3a and 3c (NELS88 in 2000) will reflect between-cohort changes (over time and across cohorts) in the determinants of wages.[34] While much research has focused on changes in the returns to education over time and across cohorts (e.g. Card and Lemieux (2001), Fortin (2006)), much less is known about changes in the impact of the cognitive and non-cognitive factors considered here.

The results begin in column (1) which includes only the female dummy without any other control variable: it essentially capture the raw female disadvantage. The raw female disadvantage is of -0.237 in 1979 and -0.229 in 1986 for the NLS72 sample and of -0.181 in 2000 for the NELS88 sample. As different explanatory variables are added in columns (2) to (6), the magnitude of the female dummy becomes larger (or the negative effect becomes smaller). Because the male-female log wage differentials are "explained" by additional regressors, the female dummy can be thought off as capturing of the part of gender differentials "unexplained" by these regressors. To the extent that a substantial part of the gender wage gap is left unexplained by the usual culprits, educational attainment and labor market experience, the estimated female dummy in this wage equation may suffer from an omitted variable bias. Adding cognitive and non-cognitive factors to the regression may help reduce this bias. The cognitive and non-cognitive factors are recorded before, at least several years, the wage realizations and are likely capturing part of an individual fixed effect that may be constant over time. It is also of interest to know whether the "hard" as well as the "soft" factors have the anticipated effects on wages, before considering their effect of the gender pay gap. I discuss the magnitude of the effects of non-cognitive skills to point out that they are not trivial in comparison with educational attainment and cognitive skills, but they do not have a clear interpretation in terms of a marginal

[34] I do not attempt to pool all the data together to isolate the time and cohort effects because of small but possibly confounding differences between the surveys.

effect, like the returns to education for example.

Column (2) adds the four non-cognitive factors together. As sole explanatory variables, they are all generally statistically significant. As found elsewhere, the self-esteem variable is positively linked to log wages, with an effect ranging from 0.05 to 0.2, while the external locus of control has a negative effect ranging from -0.1 to -0.2. In the gender specific regressions (not shown), the point estimates cover a somewhat larger range, but the male and female coefficients for self-esteem and external locus of control are not statistically different.

The other two non-cognitive factors,—importance of money/work and people/family—, also have very significant effect. In the NLS72, these effects are larger in magnitude than the first two factors, in particular the importance of money/work which ranges from 0.2 to 0.3. The effect of the importance of people/family ranges from -0.8 to -0.15 in these simple regressions. In the NELS88, the more important non-cognitive factor in this first regression is external locus of control. In the gender specific regressions (not shown), the point estimates again cover a somewhat wider range and the male and female coefficients are not statistically different, with the exception, in 1986, of the female coefficient on the importance of money which is statistically and economically larger (0.6 vs. 0.15) than the male coefficient.

Column (3) includes the cognitive measure along with the four non-cognitive factors. This absorbs some part of the explanatory power of the non-cognitive factors, reflecting the correlation between these factors and cognitive skills. The changes however are generally not statistically significant, but the importance of people/family is rendered insignificant in the 1986 and 2000 sample. This is consistent with the casual observation that people and math skills are sometimes orthogonal. The impact of math skills on wages, ranging from 0.05 to 0.12, is comparable to the impact of one or two years of post-secondary education and is more precisely estimated. The gender specific regressions reveal that the impact of math skill on female wages is statistically larger (0.17 vs. 0.08 in 2000) than the impact of this skill on male wages for workers in the mid-twenties. It is also interesting to note the impact of math skills on wages has almost double across the two cohorts of workers in the mid-twenties (Tables 3a and 3c), perhaps an implication of the increased used of technology in the workplace. On the other, the impact of math skills on the wages of workers in their early thirties is not significant, possibly absorbed by the post-graduate degree dummy.

Turning to column (4) which includes only human capital and personal characteristics, first note that the college and post-graduate estimated premiums are somewhat lower than estimates found in the literature for that period and these age-groups (Card and Lemieux (2001)). This is attributable to the inclusion of a measure of ability, the math scores, which actually reduces the potential omitted variable bias in the returns to education. The estimates are however consistent with the well-known stylized facts of substantial increases in the college premium, and especially in the post-graduate degree premium over that period. The gender specific regressions, also in accordance with well-known stylized facts, show significantly larger returns to college and post-graduate education for females than for males.

In the NLS72 (Tables 3a and 3b), the impact of actual work experience for these young workers is smaller then usually estimated; indeed in 2000, it is not significant. With tenure also capturing a significant part of the on-the-job training and relatively little age variation across the high school cohort, this is not too surprising. In the NELS88 (Table 3c), there is actually no experience premium among these 24 year olds.[35] To help benchmark the impact of non-cognitive factors, let's note that the impact of tenure (measured in years) on log wages is about 0.02 in both high school cohorts. Finally, let's note that there are no significant differences in the returns to tenure between males and females.

Columns (5) and (6) bring back the non-cognitive factors in two steps. First, in column (5) only self-esteem and external locus of control are included, then all four factors are included in column (6). The coefficients of the non-cognitive factors change very little from columns (5) to (6), reflecting the fact that these factors are not highly correlated. But the comparison of the two specifications shows that the importance of money/work has the greatest effect on the unexplained part of the gender gap, as captured by the gender dummy, which I turned to next.

In summary, the wage regressions in column (6) include the traditional human capital variables, educational attainment and labor market experience, a measure of cognitive skills, four non-cognitive factors, and some personal characteristics as controls. In this specification, the most significant non-cognitive factor that is positively correlated with wages, is the importance of money/work. With this more complete specification, the coefficients of the importance of

[35]The exclusion of workers with missing data on math score may be at play. These include a number of men with trade and vocational degrees, in analyses where these were included a small positive experience premium of 0.007 is found.

money/work from the gender specific regressions no longer show a statistically significant difference between men and women. In 1986 and 2000, self-esteem is the second most significant non-cognitive factor which is also positively correlated with wages. The coefficient of the two other non-cognitive factors—external locus and the importance of people/family—are of the expected sign, that is negative, but they are not significant with the exception of the importance of people/family in 1979. These results are consistent with the model presented in section 2 which argues that the wage rate will be positively correlated with preferences towards money/work and negatively correlated with preferences towards people/family. In the latter case, the results are however only significant in the earlier period. I now turn to the implications of these findings for the gender pay gap.

Tables 4a-c present the results of the regression-compatible Oaxaca decomposition performed by estimating equations (8) and (9) subject to restrictions (13). The decompositions are compatible with the simple wage regression of equation (14) in the sense the sum of the advantage of men and disadvantage of women in the row A3 of Table 4 (expressed in log points) corresponds exactly to the coefficients (times 100) of the female dummy displayed in the first row of Table 3. That is, $\delta_{0f} = -(\gamma_{0m} - \gamma_{0f})$, we also have $\delta_0 = \gamma_0 + \gamma_{0m}$. This unexplained part of the gender pay gap, sometimes attributed to discrimination, does not show much decline across cohorts, going from 18.9 log points in 1979 to 17.8 log points in 2000 among workers in their mid-twenties, or within cohort, going from 18.9 to 17.3 log points from 1979 to 1986, as workers move from the mid-twenties to early thirties. It is important to note however, that this unexplained part is somewhat large in comparison with other studies, because controls for occupational gender segregation, which may be affected by the non-cognitive factors, have been left out.[36]

The part of gender pay accounted by gender differences in all included characteristics is reported in row A2 of Tables 4a-c. Among workers in their early thirties in 1986 (Table 4b), more men than women have a post-graduate degree and there is a close to one year gender gap in labor market experience so that the human capital variables explain as much as 3.8 log points of the gender pay gap. For workers in their mid-twenties, the explanatory power of

[36]Including the occupation and industry dummies would account for an additional 4 to 6 log points approximately.

human capital variables is as high as 2.9 log in 1979 (Table 4b) and -2.2 log points in 2000 (Table 4c). Because in 2000 the educational attainment of women in their mid-twenties is higher than men's, educational attainment has negative explanatory power.

Against this backdrop of human capital and cognitive factors accounting for a relatively low and declining share of the gender pay gap, the non-cognitive factors figure relatively well. These factors account for 32% in 1979 and 34% in 1986 of the explained gender pay gap in the full specification in the NLS72, and for 320% in 2000 in the NELS88. Among workers in their early thirties (1986 NLS72 sample), the non-cognitive factors account for 1.9 log points in the full specification (column 6, Table 4b) and for 3.2 log points by themselves (column 1). The larger share is accounted for by the importance of money/work with 1.7 log points in column 6 and 2.2 log points in column 1.

Among workers in their mid-twenties, the non-cognitive factors account for 1.5 log points in 1979 and close to 1 log points in 2000 in the full specification. Underlying the reduced explanatory power of the importance of money/work in 2000 is a decline in the gender gap in this non-cognitive factor, shown in Table 2. In 2000, young women have become more comparable to young men not only in terms of educational attainment, but also in terms of valuing success at work and money. From 1979 to 2000, the gender wage gap among workers in their mid-twenties has declined by a 5.6 log points (from 23.7 to 18.1 log points). While human capital and cognitive factors account for most of the change, 5.0 (=2.8−(-2.2)) log points, the non-cognitive factors contribute an additional 0.6 (=1.5−0.9) log points, thus accounting for all the change of the pay gap.

6. Conclusion

In this paper, I use two high school-cohort longitudinal surveys, the NLS72 and the NELS88, to study the impact of four non-cognitive factors on wages, and on the gender pay gap in 1979, 1986 (NLS72) and 2000 (NELS88). The four non-cognitive factors include two commonly used measures, the Rosenberg self-esteem and the Rotter locus of control scales, and two other factors closely linked to gender roles in the traditional household division of labor. They are the importance of money/work and the importance of people/family. The comparison of the evolution of the gender gap and its determinants between 1979 and 2000 compares workers in

their mid-twenties across cohorts. The 1979 and 1986 comparison considers a within-cohort evolution of the gender pay gap among workers in their mid-twenties and early thirties.

The non-cognitive factors are thought to account for wage differentials through the workers' demands for wages, effort and responsibility, and altruistic rewards on the job, which interact with the employers' offers of job combinations of wages, effort and responsibility, and altruistic rewards in the context of a hedonic wage function. Like for hours of work, there is a quantity constraint of the amount of effort and responsibility that an individual can allocate between market and non-market activities. Gender differences in levels of effort and responsibility allocated to market activities and in desired levels of altruistic rewards from market and non-market activities may arise from traditional gender roles, which are learned early in life.

I do indeed find more important gender differences in the importance of money/work and the importance of people/family than in the other two, more commonly used, non-cognitive measures. For example, considering the components underlying the importance money/work, men tend to be more ambitious and value money more: the percentage of men who state that *"The chance to be a leader"* is very important in selecting a career and that *"Having lots of money"* is very important to them in their life exceeds that of women by close to 10 percentage points. In the NELS88, I find that there has been some gender convergence in work values. In particular, more women than men state that *"Being successful in work"* is very important to them in their life, but there remains some gender differences in the importance of money, so the importance of money/work composite still display a sizeable gender gap. Considering the components underlying the importance of people/family composite, the percentage of women in the NLS72 who state that *"Opportunities to work with people rather than things"* and that *"Opportunities to be helpful to others or useful to society"* is very important in selecting a career exceed that of men by more than 10 percentage points. By contrast with the previous composite, I do not find much gender convergence in the importance of people/family across the cohorts, but measurement issues may be at play.

The impacts of the non-cognitive factors on wages in these single-cohort samples are of the expected signs. Locus of control and the importance of people/family have negative effects, but these are not always significant. Indeed, the positive factors dominate: self-esteem and the

importance money/work have sizeable positive effects on wages.[37]

To measure the impact of gender differences in these non-cognitive factors on the gender pay gap, I update the Blinder-Oaxaca-Ransom wage gap decomposition introducing a correction to address the non-invariance problems associated with the choice of the non-discriminatory wage structure and with the choice of left-out category for categorical variables. The non-discriminatory wage structure is chosen to be the one where the advantage of men equals the disadvantage of women. In the case of categorical variables, all dummies are included but their coefficients are subjected to a zero-sum condition and estimated via restricted least squares. An attractive feature of the resulting approach is that it is fully compatible with the simple pooled regression that includes a gender dummy and has easily computable standard errors.

The main finding of the paper is that non-cognitive factors account for a small but not trivial part, about 2 log points, of the gender wage gap among workers in their early thirties. This role compares to the one played by gender differences in labor market experience and job tenure in 1986 in this high school cohort. The importance of work/money is the non-cognitive factor that plays the largest role, accounting for 1.7 log points, given both the larger gender differentials in this factor and the larger impact of this factor on wages. Even when workers in their mid-twenties in 1979, this non-cognitive factor accounted for 1.2 log points. This gives some support to the negotiating/bargaining divide hypothesis (Babcock and Laschever (2003)) which argues that women have less a sense of entitlement to higher wages than men. It is however difficult to distinguish this hypothesis from the fact that women may have lower wage expectations because of their foreseen lower involvement in the labor market, lower productivity or negative feedback from the labor market.[38] Among workers in the mid-twenties in 2000, non-cognitive factors play a more modest role, accounting for only close to 1 log points. In the NELS88, gender differences in the importance of money/work have shrunk by a third, so this is not too surprising. Women have closed the gender gap in educational attainment, and are beginning to close the gap in some non-cognitive factors such as the importance of money/work. The

[37]While I do not probe the mechanisms by which theses factors affect wages, in a previous version of the paper I had found that the importance of money/work had positive impacts on promotions or training for career advancement.

[38]Realized wages are affected by investments in human capital and costly signals concerning productivity, which workers undertake on the basis of expectations about future job prospects (Breen and Garcia-Penalosa (2000)).

analysis of non-cognitive factors may thus come to play an increasingly crucial role towards our understanding of gender wage differentials.

Across cohorts of workers in their mid-twenties, from 1979 to 2000, the gender pay gap decreased by 5.6 log points (from 23.7 to 18.1 log points). I find that this decrease can entirely be accounted for by the gender convergence in educational attainment, cognitive skills, valuing money and work.[39] By contrast, I find that the within-cohort change, from 1979 to 1986, in the gender gap pay of 0.8 log points (from 23.7 to 22.9 log points), among workers moving from their mid-twenties to early thirties, is not statistically significant.[40] These results suggest that the clear progression across cohorts of women's preparation for the labor market, rather than a within-cohort progression, is responsible for the decrease in the gender pay gap over the last two decades of the twentieth century. Notwithstanding some smaller non-comparability issues or possible changes in selection biases, I do find that the changes in the part of the gender pay gap unexplained by the largely pre-market factors analyzed here, are non-significant over time and across cohorts: the unexplained part adds up to 18.9 log points in 1979, in 17.3 log points in 1986 and 17.8 log points in 2000.[41] Further analyses of changes in market factors are needed to assess to impact of possible changes in labor market discrimination.

7. Appendix – Construction of the composite non-cognitive factors

7.1. Self-esteem and external locus of control

The composite measures of self-esteem and external locus of control are constructed using answers to the subset of questions, asked in the surveys, from the widely used Rosenberg self-esteem and the Rotter locus of control scales. The precise statements are reported in Table A1, along with the corresponding variable names using in the NLS72 and NELS88 surveys. The self-esteem measure is constructed using four/five questions from the Rosenberg scale. It is based on answers to questions: SQ132A,C,D asked in 1974; FT195A,C,D,H,L asked in 1979

[39]The human capital variables accounted for 2.8 log points of the gap in 1979 and -2.2 log points in 2000, thus accounting for -5.0 log points (-2.2−2.8) of the change in the gap. The non-cognitive traits accounted for an additional -0.6 log points (0.9−1.5) of the change.

[40]Considering only respondents with valid wages both in 1979 and 1986, the change is larger from 23.7 log points in 1979 to 20.1 log point in 1986, but dwarf the between-cohort changes.

[41]The case for pre-market factors is stronger for the workers in their mid-twenties.

for the NLS-72; F2D57A,C,D,H,L and F2S66A,C,D,H,L asked in 1992 for the NELS88. The external locus of control measure is constructed using four/five questions from the Rotter scale. It is based on answers to questions: SQ132B,E,F asked in 1974; FT195B,E,F,K asked in 1979 for the NLS-72; F2D57C,F,G,K and F2S66C,F,G,K asked in 1992 for the NELS88.

In the surveys, the responses to each question are recorded in a four-option format (e.g., "agree strongly", "agree", "disagree", or "disagree strongly"). These are recoded in terms of a dichotomous variable (1 or 0) reflecting only as agreement or disagreement. This binary method of recoding subjective responses is common (e.g. Waddell (2006)) in the creation of such scales and addresses the potential for anchoring, that is the fact that the strength of respondents' reaction to a statement may vary across individuals. It is more important is to capture whether a person "agrees" or "disagrees" in principle with the statement than where anchoring might take place. The alternative of coding the responses as 1, 2, 3 and 4 assigns the same distance between "strongly agree" and "agree" than between "agree" and disagree" which is a strong assumption. Indeed, alternative composites based on this other possible coding were found not to be statistically significant in the wage regressions, although their scale reliability was not much lower. The components of the scale were aggregated using STATA's alpha procedure which consists in taking the mean of the components.

7.2. The importance of money/work and the importance of people/family

The importance of money/work and of people/family composite were constructed using the questions *"How important is each of the following to you in selecting a job or career?"* and *How important is each of the following to you in your life?* where the possible answers were "Not important", "Somewhat important" and "Very important". By contrast with the previous questions, here it is the intensity of feeling that is sought, rather than a dichotomy. The answers to these questions were thus recoded as 1 for "very important", zero for other valid responses. The alternative of coding responses 1, 2, and 3 resulted in a much lower scale reliability of about 0.27 and less significant coefficients in the wage regressions. In addition, the questions are asked repeatedly over time, so that respondents who always answered "Very important" indicated that the value is much more salient for them than respondents who are not

as consistent over time.[42] The gender differences indeed do appear quite salient over time: out of 6 opportunities in the course in 14 years, the percentage of men from the NLS72 who never state that *"Having lots of money"* is very important is 53.2 percent while the corresponding percentage of women is 70 percent. To construct the money/work composite, I use answers to questions about the importance in selecting a career of *"Making of lot of money"* and of *"The chance to be a leader"* and about the importance in life of *"Being successful at work"* and *"Having lots of money"*. As shown in Table A2, these correspond to questions: BQ24A, H, BQ20A,C, FQ20A,C, SQ148A,C and TQ151A,C asked in 1972, 1973, 1974 and 1976 for the 1979 sample; these plus FT197A,C asked in 1979 for the 1986 sample; F1S46A,C, F1D36A,C and F2S40A,C, F2D36A,C and SUCSLWRK, LOTSMONY, asked in 1990, 1992 and 1994 for the 2000 sample. Because there are no question equivalent to BQ24A in the NELS88 and given the importance of leadership as a non-cognitive skill in a wage regression, I use as proxy for that skill in the NELS88, the answer "very" to the question F1S67G *"Do you think other students see you as part of the leading crowd?"* in the first follow-up, when most respondents were age 14, this was echoed by 13 percent of the sample. It is clearly an imperfect proxy since it may capture all forms of leadership rather than leadership in one's career.

The people/family composite is constructed using the answers coded 1 if "very important", 0 otherwise to questions about the importance in selecting a career of *"Opportunities to be helpful to others or useful to society"* and *"Opportunities to work with people rather than things"*, and in life of *"Helping other people in the community"*, the *"Ability to give children better opportunities"*, and of *"Living close to parents and relatives"*. As shown in Table A2, these correspond to questions: BQ24C,I, BQ20G,H, FQ20G,H, SQ148G,H and TQ151G,H asked in 1972, 1973, 1974 and 1976 for the 1979 sample; these plus FT197G,H asked in 1979 for the 1986 sample; F1S46F,G, F1D36F,G and F2S40F,G, F2D36F,G and CHLDOPTY asked in 1990, 1992 and 1994 for the 2000 sample.

The components of these composites are aggregated in the same way as the components of the self-esteem and external locus of control scale.

[42]For example, a voter who voted Republican in the past 6 elections can be seen as a much stronger Republican than a voter who did so in only 3 of those. In a previous version of the paper, I reported the complete list of answers reporting the averages by gender for each of the waves of the surveys.

References

Adelman, C. (1994): *Lessons of a Generation: Education and Work in the Lives of the High School Class of 1972*. Jossey-Bass Publishers, San Francisco.

Ajzen, I., and M. Fishbein (1977): "Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research," *Psychological Bulletin*, 84, 888–918.

Andreoni, J., and L. Vesterlund (2001): "Which is the Fair Sex? Gender Differences in Altruism," *Quarterly Journal of Economics*, 116, 293–312.

Angrist, J. D., and G. W. Imbens (1995): "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association*, 90, 431–442.

Babcock, L., and S. Laschever (2003): *Women Don't Ask: Negotiation and the Gender Divide*. Princeton University Press, Princeton, NJ.

Becker, G. S. (1985): "Human Capital, Effort, and the Sexual Division of Labor," *Journal of Labor Economics*, 3, S33–S58.

Bielby, D. D. V., and W. T. Bielby (1984): "Work Commitment, Sex-Role Attitudes, and Women's Emploment," *American Sociological Review*, 49, 234–247.

Blau, F. D., and L. Kahn (2004): "The US Gender Pay Gap in the 1990s: Slowing Convergence," NBER Working Paper 10853, National Bureau of Economic Research, Cambridge, MA.

Blau, F. D., and L. M. Kahn (2000): "Gender Differences in Pay," *Journal of Economic Perspectives*, 14, 75–99.

Blinder, A. S. (1973): "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources*, 8, 436–455.

Borghans, L., B. ter Weel, and B. A. Weinberg (2006): "People People: Social Capital and the Labor-Market Outcomes of Underrepresented Groups," NBER Working Paper 11985, National Bureau of Economic Research, Cambridge, MA.

Bowles, S., H. Gintis, and M. Osborne (2001): "The Determinants of Earnings: A Behavioral Approach," *Journal of Economic Literature*, 39, 1137–1176.

Breen, R., and C. Garcia-Penalosa (2000): "Bayesian Learning and Gender Segregation," *Journal of Labor Economics*, 20, 899–922.

Brockner, J. (1988): *Self Esteem at Work: Research, Theory and Practice*. Lexington Books, Lexington, Mass.

Brown, C., and M. Corcoran (1997): "Sex-Based Differences in School Content and the Male-Female Wage Gap," *Journal of Labor Economics*, 15, 431–465.

Card, D. E., and T. Lemieux (2001): "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, 116(2), 705–746.

Chevalier, A. (2004): "Motivation, Expectations and the Gender Pay Gap for UK Graduates," Working Paper 1101, Institute for the Study of Labor (IZA), Bonn.

Coleman, M., and T. DeLeire (2003): "An Economic Model of Locus of Control and the Human Capital Investment Decision," *Journal of Human Resources*, 38, 701–721.

Cotton, J. (1988): "On the Decomposition of Wage Differentials," *Review of Economics and Statistics*, 70, 236–243.

Daymont, T. N., and P. J. Andrisani (1984): "Job Preferences, College Major, and the Gender Gap in Earnings," *Journal of Human Resources*, 19, 408–428.

Ekeland, I., J. Heckman, and L. Nesheim (2004): "Identification and Estimation of Hedonic Models," *Journal of Political Economy*, 112, S60–S103.

Filer, R. K. (1986): "The Role of Personality and Tastes in Determining Occupational Structure," *Industrial and Labor Relations Review*, 39, 412–424.

Fortin, N. M. (2005): "Gender Role Attitudes and the Labour Market Outcomes of Women Across OECD Countries," *Oxford Review of Economic Policy*, 21, 22pp.

——— (2006): "Higher Education Policies and the College Premium: Cross-State Evidence from the 1990s," *American Economic Review*, 96(4), 959–987.

Furnham, A., and M. Argyle (1998): *The Psychology of Money*. Routledge, London.

Gardeazabal, J., and A. Ugidos (2004): "More on the Identification in Detailed Wage Decompositions," *Review of Economics and Statistics*, 86(4), 1034–1057.

Gneezy, U., M. Niederle, and A. Rustichini (2003): "Performance in Competitive Environments: Gender Differences," *Quarterly Journal of Economics*, 118, 1049–1074.

Goldsmith, A. H., J. R. Veum, and W. Darity (1997): "The Impact of Psychological and Human Capital on Wages," *Economic Inquiry*, 35, 815–829.

Heckman, J., J. Stixrud, and S. Urzua (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24, 411–482.

Judge, T. A., H. B. Tippie, and J. E. Bono (2001): "Relationship of Core Self-evaluations Traits–Self-esteem, Generalized Self-efficacy, Locus of control, and Emotional Stability–with Job Satisfaction and Job Performance: A Meta-Analysis," *Journal of Applied Psychology*, 86, 80–92.

Kirkcaldy, B., and A. Furnham (1993): "Predictors of belief about money," *Psychological Reports*, 73, 1079–1082.

Krueger, A. B., and S. B. Dale (2002): "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," *Quarterly Journal of Economics*, 117, 1491–1527.

Kuhn, P., and C. Weinberger (2005): "Leadership Skills and Wages," *Journal of Labor Economics*, 23, 395–436.

Manning, A., and J. Swaffield (2005): "The Gender Gap in Early-Career Wage Growth," CEP Discussion Paper 700, Center for Economic Performance (LSE), London.

Mueller, G., and E. S. Plug (2006): "Estimating the Effect of Personality on Male-Female Earnings," *Industrial and Labor Relations Review*, 60, 3–22.

Neumark, D. (1988): "Employer's Discriminatory Behavior and the Estimation of Wage Discrimination," *Journal of Human Resources*, 23, 279–295.

Nyhus, E. K., and E. Pons (2005): "The Effects of Personality on Earnings," *Journal of Economic Psychology*, 26, 363–384.

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

Oaxaca, R. L., and M. R. Ransom (1994): "On discrimination and the decomposition of wage differentials," *Journal of Econometrics*, 61(1), 5–21.

——— (1999): "Identification in Detailed Wage Decompositions," *Review of Economics and Statistics*, 81(1), 154–157.

O'Neill, J., and S. Polachek (1993): "Why the Gender Gap in Wages Narrowed in the 1980s," *Journal of Labor Economic*, 11, 25–33.

O'Neill, J. A. (2003): "The Gender Gap in Wages, circa 2000," *American Economic Review – AEA Papers and Proceedings*, 93(2), 309–314.

O'Neill, J. E., and D. M. O'Neill (2005): "What do Wage Differentials Tell us about Labor Market Discrimination," NBER Working Paper 11240, National Bureau of Economic Research, Cambridge, MA.

Osborne Groves, M. (2005): "How Important Is Your Personality? Labor Market Returns to Personality for Women in the US and UK," *Journal of Economic Psychology*, 26, 827–841.

Rosen, S. (1974): "Hedonic Prices and Implicit Markets," *Journal of Political Economy*, 82, 34–55.

Rowe, R., and W. E. Snizek (1995): "Gender Differences in Work Values- Perpetuating the Myth," *Work and Occupations*, 22, 215–229.

Ruble, D. N., and C. L. Martin (1998): "Gender Development," in *Handbook of Child Psychology*, ed. by W. Damon, and N. Eisenberg, vol. 4, pp. 933–1016. John Wiley, New York.

Säve-Söderbergh, J. (2003): "Are Women Asking for Low Wages? Individual Wage Bargaining and Gender Wage Differentials," Mimeo, Swedish Institute for Social Research, Stockholm University.

Tang, T. L.-P., J. K. Kim, and D. S.-H. Tang (2000): "Does Attitude Toward Money Moderate the Relationship Between Intrinsic Job Satisfaction and Voluntary Turnover?," *Human Relations*, 53, 213–246.

Waddell, G. R. (2006): "Labor Market Consequences of Poor Attitude and Low Self-Esteem in Youth," *Economic Inquiry*, 44, 69–97.

Weinberger, C. J., and P. J. Kuhn (2006): "The Narrowing of the U.S. Gender Earnings Gap, 1959-1999: A Cohort-Based Analysis," NBER Working Paper 12115, National Bureau of Economic Research, Cambridge, MA.

Table 1 – Mean and Standard Deviations of Commonly Used Variables

	NLS72				NELS88	
Year (Age)	1979 (25)		1986 (32)		2000 (24)	
Variables	Male	Female	Male	Female	Male	Female
Log Wage	1.803	1.566	2.521	2.292	2.621	2.440
	(0.418)	(0.395)	(0.585)	(0.693)	(0.469)	(0.467)
Education:						
less then HS	0.006	0.001	0.002	0.003	0.006	0.006
	(0.080)	(0.028)	(0.045)	(0.055)	(0.079)	(0.076)
High School	0.202	0.233	0.178	0.211	0.189	0.133
	(0.401)	(0.423)	(0.383)	(0.408)	(0.391)	(0.339)
Trade	0.112	0.087	0.155	0.133	0.099	0.092
	(0.316)	(0.281)	(0.362)	(0.340)	(0.299)	(0.288)
Some Coll	0.368	0.369	0.297	0.327	0.409	0.362
	(0.482)	(0.483)	(0.457)	(0.469)	(0.492)	(0.480)
College	0.267	0.269	0.252	0.233	0.267	0.362
	(0.443)	(0.443)	(0.434)	(0.423)	(0.442)	(0.481)
Post-Graduate	0.044	0.042	0.115	0.093	0.029	0.047
	(0.206)	(0.200)	(0.362)	(0.290)	(0.169)	(0.211)
Work Experience:						
Years	5.269	4.747	10.851	9.625	4.682	4.461
	(1.440)	(1.560)	(2.354)	(2.917)	(2.234)	(2.028)
Tenure	1.018	0.767	4.916	4.084	2.194	1.988
	(2.000)	(1.738)	(3.816)	(3.616)	(2.225)	(2.057)
Cognitive Skills:						
High School Math	0.091	-0.114	0.098	-0.158	0.010	-0.011
Standardized Score	(0.974)	(0.962)	(0.979)	(0.973)	(1.020)	(0.977)
Personal Characteristics:						
Part-Time	0.048	0.110	0.025	0.164	0.067	0.150
	(0.213)	(0.312)	(0.156)	(0.370)	(0.250)	(0.357)
Black	0.052	0.101	0.052	0.100	0.122	0.101
	(0.222)	(0.301)	(0.222)	(0.300)	(0.327)	(0.301)
Married	0.528	0.524	0.736	0.687	0.368	0.429
	(0.499)	(0.499)	(0.441)	(0.464)	(0.483)	(0.495)
Parental	1.011	1.293	4.507	5.479	0.943	1.174
Experience	(2.071)	(2.354)	(4.767)	(5.476)	(1.992)	(2.293)
Child	0.315	0.333	0.642	0.649	0.306	0.150
	(0.465)	(0.471)	(0.479)	(0.478)	(0.461)	(0.356)
No. of observations	3391	2793	3456	3066	3183	3293

Note: Standard deviations in parentheses.

Table 2 – Mean, Standard Deviations, and Reliability Coefficient of the Non-Cognitive Variables

	NLS72						NELS88		
Year (Age)	1979 (25)			1986 (32)			2000 (24)		
Variables	Male	Female	Δ	Male	Female	Δ	Male	Female	Δ
Self-esteem[a]	0.851	0.877	-0.026	0.935	0.930	0.005	0.757	0.731	0.026
	(0.286)	(0.244)		(0.141)	(0.159)		(0.144)	(0.178)	
	[0.7967]			[0.5629]			[0.6310]		
External locus of control[b]	0.123	0.100	0.023	0.096	0.104	-0.008	0.179	0.156	0.023
	(0.228)	(0.205)		(0.181)	(0.305)		(0.247)	(0.238)	
	[0.4529]			[0.4858]			[0.5975]		
Money/Work[d] very important	0.447	0.390	0.057	0.456	0.391	0.065	0.561	0.518	0.043
	(0.189)	(0.167)		(0.183)	(0.160)		(0.206)	(0.196)	
	[0.5869]			[0.6347]			[0.5487]		
People/Family[e] very important	0.326	0.389	-0.063	0.322	0.382	-0.060	0.629	0.686	-0.057
	(0.196)	(0.199)		(0.193)	(0.196)		(0.253)	(0.254)	
	[0.6268]			[0.6823]			[0.4823]		

Note: Standard deviations in parentheses. Cronach's alpha scale measure of reliability in brackets. The composite non-cognitive variables are generated using STATA's alpha procedure, as explained in detail in the Appendix, using components measured many years before the wage data is gathered. They are on a [0,1] scale. Δ is equal to the male minus the female measure.

[a] The self-esteem measure is constructed using four/five questions from the Rosenberg scale.

[b] The external locus of control measure is constructed using four/five questions from the Rotter scale.

[c] The money/work composite is constructed using the answers coded 1 for "Very important", 0 otherwise to valid answers about the importance in selecting a career of "Making of lot of money" and of "The chance to be a leader in one line of work" and about the importance in life of "Being successful at work" and "Having lots of money".

[d] The people/family composite is constructed using the answers coded 1 if "Very important", 0 otherwise to valid answers about the importance in selecting a career of "Opportunities to be helpful to others or useful to society" and "Opportunities to work with people rather than things", and in life of "Helping other people in the community," the "Ability to give children better opportunities", and "Living close to parents and relatives".

TABLE 3A – IMPACT OF NON-COGNITIVE TRAITS ON 1979 LOG WAGE
(AGE 25) - NLS 72

	(1)	(2)	(3)	(4)	(5)	(6)
Explanatory variables:						
Female	-0.237**	-0.220**	-0.211**	-0.203**	-0.206**	-0.189**
	(0.010)	(0.011)	(0.011)	(0.010)	(0.011)	(0.011)
Non-cognitive traits:						
Self-esteem		0.100**	0.080**		0.054**	0.043**
		(0.019)	(0.019)		(0.019)	(0.019)
External locus		-0.099**	-0.063**		-0.025	-0.036
of control		(0.024)	(0.024)		(0.024)	(0.024)
Money/Work		0.221**	0.232**			0.209**
very important		(0.030)	(0.030)			(0.030)
People/Family		-0.151**	-0.087**			-0.083**
very important		(0.022)	(0.028)			(0.028)
Cognitive Skills:						
High School Math			0.051**	0.035*	0.033*	0.035**
Standardized Score			(0.006)	(0.006)	(0.006)	(0.006)
Education: (HS omitted)						
Less then HS				-0.090	-0.094	-0.085
				(0.081)	(0.081)	(0.080)
Trade				0.051**	0.053*	0.048**
				(0.019)	(0.019)	(0.019)
Some College				0.052**	0.051**	0.044**
				(0.014)	(0.014)	(0.014
College				0.144**	0.140**	0.128**
				(0.017)	(0.018)	(0.018)
Post-Graduate				0.202**	0.196**	0.179**
				(0.022)	(0.029)	(0.029)
Experience:						
Years				0.039*	0.037**	0.036**
				(0.004)	(0.004)	(0.004)
Tenure				0.021**	0.021**	0.021**
				(0.003)	(0.003)	(0.003)
Personal Characteristics:						
Part-Time				-0.030	-0.033	-0.029
				(0.020)	(0.020)	(0.020)
Black				-0.036	-0.034	-0.041
				(0.021)	(0.021)	(0.021)
Married				0.038**	0.036**	0.039**
				(0.011)	(0.011)	(0.011)
Parental				-0.003	-0.003	-0.003
Experience				(0.003)	(0.003)	(0.003)
Child				0.006	0.007	0.009
				(0.017)	(0.017)	(0.017)
Adjusted R-2	0.076	0.093	0.104	0.129	0.131	0.139
No. of observations	6184	6184	6184	6184	6184	6184

Note: Standard errors are in parentheses. Significance at 5% level denoted by **, significance at 10% level denoted by *.

TABLE 3B – IMPACT OF NON-COGNITIVE TRAITS ON 1986 LOG WAGE (AGE 32) - NLS 72

	(1)	(2)	(3)	(4)	(5)	(6)
Explanatory variables:						
Female	-0.229**	-0.197**	-0.187**	-0.188**	-0.189**	-0.172**
	(0.016)	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)
Non-cognitive traits:						
Self-esteem		0.200**	0.168**		0.125**	0.118**
		(0.054)	(0.054)		(0.054)	(0.054)
External locus		-0.162**	-0.118**		-0.058	-0.070
of control		(0.044)	(0.044)		(0.044)	(0.044)
Money/Work		0.328**	0.334**			0.258**
very important		(0.047)	(0.047)			(0.047)
People/Family		-0.150**	-0.066			-0.025
very important		(0.042)	(0.044)			(0.043)
Cognitive Skills:						
High School Math			0.060**	0.011	0.008	0.014
Standardized Score			(0.009)	(0.009)	(0.009)	(0.010)
Education: (HS omitted)						
Less then HS				-0.160	-0.160	-0.138
				(0.155)	(0.155)	(0.155)
Trade				0.031	0.028	0.023
				(0.027)	(0.027)	(0.027)
Some College				0.094**	0.089**	0.081**
				(0.023)	(0.023)	(0.023)
College				0.210**	0.203**	0.189**
				(0.027)	(0.027)	(0.027)
Post-Graduate				0.316**	0.306**	0.289**
				(0.034)	(0.034)	(0.034)
Experience:						
Years				0.008**	0.007*	0.006*
				(0.003)	(0.003)	(0.003)
Tenure				0.019**	0.019**	0.019**
				(0.002)	(0.002)	(0.002)
Personal Characteristics:						
Part-Time				0.039	0.036	0.048
				(0.028)	(0.028)	(0.028)
Black				0.018	0.017	0.011
				(0.031)	(0.031)	(0.032)
Married				0.019	0.013	0.012
				(0.019)	(0.019)	(0.019)
Parental				-0.010**	-0.010**	-0.010**
Experience				(0.002)	(0.002)	(0.002)
Child				0.047	0.046	0.048*
				(0.024)	(0.024)	(0.024)
Adjusted R-2	0.031	0.043	0.050	0.075	0.079	0.081
No. of observations	6522	6522	6522	6522	6522	6522

Note: Standard errors are in parentheses. Significance at 5% level denoted by **, significance at 10% level denoted by *.

TABLE 3C – IMPACT OF NON-COGNITIVE TRAITS ON 2000 LOG WAGE
(AGE 24) - NELS 88

	(1)	(2)	(3)	(4)	(5)	(6)
Explanatory variables:						
Female	-0.181**	-0.173**	-0.169**	-0.188**	-0.186**	-0.178**
	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)
Non-cognitive traits:						
Self-esteem		0.054	0.069*		0.084**	0.082**
		(0.038)	(0.036)		(0.035)	(0.035)
External locus		-0.207**	-0.081**		-0.035	-0.032
of control		(0.054)	(0.054)		(0.024)	(0.024)
Money/Work		0.151**	0.190**			0.165**
very important		(0.029)	(0.029)			(0.028)
People/Family		-0.082**	-0.016			-0.019
very important		(0.023)	(0.023)			(0.022)
Cognitive Skills:						
High School Math			0.120**	0.063*	0.060*	0.064**
Standardized Score			(0.006)	(0.001)	(0.001)	(0.007)
Education: (HS omitted)						
Less then HS				0.069	0.066	0.064
				(0.071)	(0.071)	(0.071)
Trade				0.084**	0.082*	0.086**
				(0.022)	(0.023)	(0.023)
Some College				0.072**	0.073**	0.068**
				(0.017)	(0.017)	(0.017)
College				0.248**	0.246**	0.238**
				(0.021)	(0.021)	(0.021)
Post-Graduate				0.447**	0.443**	0.432**
				(0.036)	(0.036)	(0.036)
Experience:						
Years				0.001	0.002	0.001
				(0.003)	(0.003)	(0.003)
Tenure				0.015**	0.015**	0.015**
				(0.003)	(0.003)	(0.003)
Personal Characteristics:						
Part-Time				-0.207**	-0.206**	-0.202**
				(0.018)	(0.018)	(0.018)
Black				-0.036**	-0.035*	-0.038*
				(0.018)	(0.024)	(0.018)
Married				0.055**	0.054**	0.058**
				(0.011)	(0.012)	(0.012)
Parental				0.002	0.002	0.001
Experience				(0.003)	(0.003)	(0.003)
Child				-0.052**	-0.052**	-0.052**
				(0.017)	(0.017)	(0.017)
Adjusted R-2	0.036	0.053	0.111	0.167	0.168	0.172
No. of observations	6476	6476	6476	6476	6476	6476

Note: Standard errors are in parentheses. Significance at 5% level denoted by **, significance at 10% level denoted by *.

TABLE 4A – DECOMPOSITION OF THE GENDER WAGE GAP 1979 – NLS72

Specification	(2)	(3)	(4)	(5)	(6)
A1. Raw log wage gap			23.71		
in log points			(1.03)		
A2. Differences in characteristics	1.72	2.56	3.39	3.09	4.78
$(\Delta \overline{X}'\hat{\gamma})$	(0.28)	(0.24)	(0.24)	(0.27)	(0.37)
as % of the raw gap	7.3	10.8	14.3	13.0	20.2
A3. Advantage of men−disadvantage of women					
$(\hat{\gamma}_{0m} - \hat{\gamma}_{0f})$	21.98	21.14	20.31	20.62	18.93
	(0.76)	(0.76)	(0.74)	(0.74)	(0.76)
Contribution of explanatory variables to					
B1. Differences in characteristics:$(\Delta \overline{X}'\hat{\gamma})$					
Non-cognitive traits:	**1.72**	**1.51**		**-0.20**	**1.52**
	(0.28)	(0.26)		(0.12)	(0.26)
Self-esteem	-0.23	-0.21		-0.14	-0.11
	(0.05)	(0.05)		(0.05)	(0.05)
External locus	-0.26	-0.14		-0.06	-0.08
of control	(0.05)	(0.05)		(0.05)	(0.05)
Money/work	1.26	1.32			1.19
very important	(0.17)	(0.17)			(0.17)
People/family	0.95	0.55			0.53
very important	(0.17)	(0.18)			(0.18)
Human capital & cognitive skills :		**1.05**	**2.94**	**2.04**	**2.81**
		(0.11)	(0.23)	(0.20)	(0.25)
HS math score		1.05	0.71	0.68	0.72
		(0.11)	(0.13)	(0.13)	(0.25)
education			0.10	0.11	0.10
			(0.08)	(0.08)	(0.08)
experience, tenure			2.12	2.04	1.99
			(0.19)	(0.20)	(0.20)
B2. Advantage of men: $\overline{X}_m(\widehat{\beta}_m - \widehat{\gamma})$					
Non-cognitive traits:	-0.49	-1.11		-7.09	0.15
Human capital & cognitive skills :			-0.19	-8.03	-8.49
B3. Disadvantage of women: $-\overline{X}_f(\widehat{\beta}_f - \widehat{\gamma})$					
Non-cognitive traits:	0.34	1.75		3.93	2.71
Human capital & cognitive skills :			-0.32	4.40	4.97
B4. Sum unexplained: $\overline{X}_m(\widehat{\beta}_m - \widehat{\gamma}) - \overline{X}_f(\widehat{\beta}_f - \widehat{\gamma})$					
Non-cognitive traits:	-0.15	0.64		2.49	2.86
Human capital & cognitive skills :			-3.16	-3.63	-3.52

Note: Standard errors are in parentheses. The bold face numbers sum up the corresponding differences in characteristics. The decompositions are compatible with the regression in Table 3a, with each column corresponding to specification from that table. However, the actual regressions used in the computations follow the specifications of equations (9), (10) and (14), and are thus impervious to the left-category problem.

TABLE 4B – DECOMPOSITION OF THE GENDER WAGE GAP 1986 – NLS72

Specification	(2)	(3)	(4)	(5)	(6)
A1. Raw log wage gap			22.94		
in log points			(1.56)		
A2. Differences in characteristics	3.22	4.26	4.17	4.00	5.70
$(\Delta \overline{X}'\hat{\gamma})$	(0.45)	(0.48)	(0.59)	(0.59)	(0.71)
as % of the raw gap	14.0	18.6	18.1	17.4	24.8
A3. Advantage of men−disadvantage of women					
$(\hat{\gamma}_{0m} - \hat{\gamma}_{0f})$	19.71	18.68	18.79	18.95	17.25
	(1.16)	(1.16)	(1.17)	(1.17)	(1.20)
Contribution of explanatory variables to					
B1. Differences in characteristics:$(\Delta \overline{X}'\hat{\gamma})$					
Non-cognitive traits:	**3.22**	**2.71**		**0.07**	**1.92**
	(0.40)	(0.41)		(0.03)	(0.41)
Self-esteem	0.09	0.06		0.03	0.03
	(0.02)	(0.02)		(0.02)	(0.02)
External locus	0.07	0.06		0.04	0.04
of control	(0.02)	(0.02)		(0.02)	(0.02)
Money/work	2.15	2.19			1.69
very important	(0.31)	(0.31)			(0.31)
People/family	0.92	0.40			0.15
very important	(0.28)	(0.27)			(0.27)
Human capital & cognitive skills :		**1.55**	**3.76**	**3.56**	**3.43**
		(0.22)	(0.56)	(0.56)	(0.56)
HS math score		1.55	0.27	0.21	0.35
		(0.22)	(0.24)	(0.25)	(0.25)
education			0.93	0.90	0.84
			(0.18)	(0.00)	(0.18)
experience, tenure			2.56	2.45	2.36
			(0.52)	(0.47)	(0.47)
B2. Advantage of men: $\overline{X}_m(\widehat{\beta}_m - \widehat{\gamma})$					
Non-cognitive traits:	-0.05	-0.23		5.14	3.06
Human capital & cognitive skills :			-0.01	-22.69	-22.22
B3. Disadvantage of women: $-\overline{X}_f(\widehat{\beta}_f - \widehat{\gamma})$					
Non-cognitive traits:	2.55	2.63		-2.65	-0.08
Human capital & cognitive skills :			0.05	10.21	10.01
B4. Sum unexplained: $\overline{X}_m(\widehat{\beta}_m - \widehat{\gamma}) - \overline{X}_f(\widehat{\beta}_f - \widehat{\gamma})$					
Non-cognitive traits:	2.50	2.61		2.49	2.98
Human capital & cognitive skills :			0.06	-12.48	-12.21

Note: Standard errors are in parentheses. The bold face numbers sum up the corresponding differences in characteristics. The decompositions are compatible with the regression in Table 3b, with each column corresponding to specification from that table. However, the actual regressions used in the computations follow the specifications of equations (9), (10) and (14), and are thus impervious to the left-category problem.

TABLE 4C – DECOMPOSITION OF THE GENDER WAGE GAP 2000 – NELS88

Specification	(2)	(3)	(4)	(5)	(6)
A1. Raw log wage gap			18.11		
in log points			(1.15)		
A2. Differences in characteristics	0.80	1.16	-0.68	-0.51	0.30
$(\Delta \overline{X}'\hat{\gamma})$	(0.24)	(0.24)	(0.59)	(0.29)	(0.35)
as % of the raw gap	4.4	6.3	-3.7	-2.8	1.7
A3. Advantage of men−disadvantage of women					
$(\hat{\gamma}_{0m} - \hat{\gamma}_{0f})$	17.31	16.94	18.78	18.61	17.80
	(0.83)	(0.81)	(0.79)	(0.79)	(0.80)
Contribution of explanatory variables to					
B1. Differences in characteristics:$(\Delta \overline{X}'\hat{\gamma})$					
Non-cognitive traits:	**0.80**	**0.91**		**0.14**	**0.96**
	(0.24)	(0.21)		(0.11)	(0.21)
Self-esteem	0.14	0.18		0.22	0.21
	(0.06)	(0.02)		(0.09)	(0.05)
External locus	-0.46	-0.18		-0.08	-0.07
of control	(0.06)	(0.09)		(0.09)	(0.05)
Money/work	0.65	0.82			0.71
very important	(0.13)	(0.12)			(0.12)
People/family	0.47	0.09			0.11
very important	(0.13)	(0.21)			(0.13)
Human capital & cognitive skills :		**0.25**	**-2.24**	**-2.22**	**-2.17**
		(0.01)	(0.23)	(0.23)	(0.23)
HS math score		0.25	0.13	0.12	0.13
		(0.01)	(0.01)	(0.01)	(0.01)
education			-2.71	-2.68	-2.61
			(0.21)	(0.21)	(0.21)
experience, tenure			0.33	0.34	0.31
			(0.08)	(0.08)	(0.08)
B2. Advantage of men: $\overline{X}_m(\widehat{\beta}_m - \widehat{\gamma})$					
Non-cognitive traits:	-4.89	-8.97		-4.87	-7.24
Human capital & cognitive skills :			-5.36	-5.48	-5.42
B3. Disadvantage of women: $-\overline{X}_f(\widehat{\beta}_f - \widehat{\gamma})$					
Non-cognitive traits:	3.04	9.39		5.75	9.86
Human capital & cognitive skills :			8.59	8.75	8.67
B4. Sum unexplained: $\overline{X}_m(\widehat{\beta}_m - \widehat{\gamma}) - \overline{X}_f(\widehat{\beta}_f - \widehat{\gamma})$					
Non-cognitive traits:	-1.85	0.42		0.89	2.62
Human capital & cognitive skills :			3.22	3.27	3.24

Note: Standard errors are in parentheses. The bold face numbers sum up the corresponding differences in characteristics. The decompositions are compatible with the regression in Table 3c, with each column corresponding to specification from that table. However, the actual regressions used in the computations follow the specifications of equations (9), (10) and (14), and are thus impervious to the left-category problem.

Table A1 – Questions on Self-Esteem (SE) and External Locus of Control (LC) and Associated Variables

	SE	LC	SE	LC	SE	LC
		NLS72			**NELS88**	
Year (Age)	1974(20)		1979(25)		1992(16)	
How do you feel about each of the following statement?						
a. I feel good about my self	SQ132A		FT195A		F2S66A	
b. In my life, good luck is more important than hard work for success		SQ132B		FT195B		F2S66B
c. I feel I am a person of work, the equal to other people	SQ132C		FT195C		F2S66C	
d. I am able to do things as well as other people	SQ132D		FT195D		F2S66D	
e. Every time I try to get ahead, something or somebody stops me		SQ132E		FT195E		F2S66E
f. My plans hardly ever work out, so planning only makes me unhappy		SQ132F		FT195F		F2S66F
h. On the whole, I am satisfied with myself	SQ132H		FT195H		F2S66H	
k. When I make plans, I am almost certain I can make them work				FT195K		F2S66K
l. I feel I do not have much to be proud of			FT195L		F2S66L	

Note: Respondents were asked to circle one of "Agree strongly", "Agree", "Disagree" and "Diasagree strongly" as answers. Agreement was coded as 1 and diasgreement was coded as 0, except for question k were the coding was reversed. In 1992, part of the sample's answers were recorded by the variable F2D57.

Table A2 – Questions on Important Values and Associated Variables

	Money	People	Money	People	Money	People	Money	People	Money	People
NLS72										
Year (Age)	1972 (18)		1973 (19)		1974 (20)		1976 (22)		1979 (25)	
How important is each of the following to you in selecting a job or career?										
a. Making a lot of money	BQ24A									
c. Opportunities to be helpful to others or useful to society		BQ24C								
h. The chance to be a leader	BQ24H									
i. Opportunities to work with people rather than things		BQ24I								
How important is each of the following to you in your life?										
a. Being successful in work	BQ20A		FQ20A		SQ148A		TQ151A		FT197A	
c. Having lots of money	BQ20C		FQ20C		SQ148C		TQ151C		FT197C	
g. Being able to give children better opportunities		BQ20G		FQ20G		SQ148G		TQ151G		FT197G
h. Living close to parents and relatives		BQ20H		FQ20H		SQ148H		TQ151H		FT197H
NELS88										
Year (Age)	1990 (14)		1992 (16)		1994 (18)					
How important is each of the following to you in your life?										
a. Being successful in work	F1S46A		F2S40A		SUCSLWRK					
c. Having lots of money	F1S46C		F2S40C		LOTSMONY					
f. Helping other people in my community		F1S46F		F2S40G						
g. Being able to give children better opportunities		F1S46G		F2S40H		CHLDOPTY				
How do you think other students see you?										
g. As part of the leading crowd	F1S67G									

Notes: Respondents were asked to circle one of "Not important", "Somewhat important" and "Very important" as answers to the above questions. Responses were coded as one if the answer was "very important", zero otherwise. Valid responses to question F1S67G were "Strongly Agree", "Agree", "Disagree" and "Diasagree strongly" as answers. "Strongly agree" was deemed similar to "Very important" and coded as 1, zero otherwise.

TABLE A3 – DECOMPOSITION OF THE GENDER GAP USING DIFFERENT REFERENCE WAGE STRUCTURES - 1986 NLS72

Difference: 22.96 log points	Explained $\Delta\overline{\mathbf{X}}'\widehat{\beta}_b$	Unexplained: Advantage of Men $\overline{X}_m(\widehat{\beta}_m - \widehat{\beta_b})$	Unexplained: Disadvantage of Women $-\overline{X}_f(\widehat{\beta}_f - \widehat{\beta_b})$	Sum or $\overline{X}_b\Delta\beta$
a. Male Coefficients	1.48	—	—	21.48
$\widehat{\beta}_b = \widehat{\beta}_m$	(1.20)			(1.97)
b. Female Coefficients	6.75	—	—	16.20
$\widehat{\beta}_b = \widehat{\beta}_f$	(1.02)			(1.86)
c. Pooled Coefficients	8.69	6.49	7.77	14.27
$\widehat{\beta}_b = \widehat{\beta}_p$	(0.65)	(1.27)	(1.48)	(1.95)
d. Cotton Coefficients	3.89	7.39	11.68	19.07
$\widehat{\beta}_b = w\widehat{\beta}_m + (1-w)\widehat{\beta}_f$	(1.13)	(1.60)	(1.85)	(2.45)
e. Regression-Compatible	5.70	8.62	8.63	17.25
$\widehat{\beta}_b = \widehat{\gamma}$	(0.71)	(0.85)	(0.85)	(1.20)

Note: Standard errors in parentheses. In the Cotton (1988) decomposition, the weight w is the proportion of males in the sample.