Labor Market Signaling and Self-Confidence: Wage Compression and the Gender Pay Gap

Luís Santos-Pinto

Faculty of Business and Economics, University of Lausanne

This version: December 29, 2011

Abstract

I extend Spence's (1973) signaling model by assuming some workers are overconfident–they underestimate their marginal cost of acquiring education–and some are underconfident. Firms cannot observe workers' productive abilities and beliefs but know the fractions of high-ability, overconfident, and underconfident workers. I find that biased beliefs lower the wage spread and compress the wages of unbiased workers. I show that gender differences in self-confidence can contribute to the gender pay gap. If education raises productivity, men are overconfident, and women underconfident, then women will, on average, earn less than men. Finally, I show that biased beliefs can improve welfare.

I am thankful to Miguel Costa-Gomes, Robert Dur, Lorenz Goette, Hans Hvide, Bettina Klaus, David Myatt, Joel Sobel, and Michael Waldman for helpful comments and suggestions.
Faculty of Business and Economics, University of Lausanne, Internef 53, CH-1015, Lausanne, Switzerland. Ph: 41-216923658. Fax: 41-216923435. E-mail address: LuisPedro.SantosPinto@unil.ch.

1 Introduction

This paper explores the implications of worker self-confidence in the classic labor market signaling model by Spence (1973). Firms are perfectly competitive and cannot observe workers' productive abilities, which may be either high or low. Unbiased workers know their marginal cost of acquiring education. Overconfident (underconfident) workers believe that their marginal cost of acquiring education is low (high) when, in fact, it is high (low). Firms cannot observe workers' beliefs but know the fraction of high-ability, overconfident, and underconfident workers in the labor market.

I start by exploring whether workers' biased self-evaluations can lead to wage compression. Wage compression, i.e., the fact that there are lower wage differences across workers than differences in productivity, is a key feature of many labor markets and has important consequences for labor market performance. While a compressed wage structure is generally believed to reduce labor market efficiency and welfare, it has also been used to explain why firm sponsored training of employees can arise.[1]

My first result shows that biased beliefs compress wages in the sense that the wage spread with biased workers is smaller than the wage spread with rational workers. In a separating equilibrium, overconfident low-ability workers and unbiased high-ability workers choose a high education level whereas underconfident high-ability workers and unbiased low-ability workers choose a low education level. The optimal response of firms to the fact that overconfidence raises the proportion of low-ability workers in the high education group whereas underconfidence raises the proportion of high-ability workers in the low education group is to lower the wage of the high education group and raise the wage of the low education group.

I also find that the presence of biased workers in the labor market compresses the wages of unbiased workers in the sense that the wage spread across unbiased workers is less than the productivity spread. In contrast, the model predicts that the wage spread across biased workers is greater than the productivity spread.

Next, I study the impact of gender differences in self-confidence on the gender pay gap, i.e., the fact that men earn, on average, more than women.

[1] The economics literature has proposed several explanations for wage compression, ranging from labor market institutions, to incentives not to sabotage colleagues competing in a tournament, to fairness considerations in wage-setting decisions by firms. I review the relevant literature in Section 3.

I show that if education raises productivity, males are overconfident, and females underconfident, then there is a gender pay gap.[2] The intuition behind this result is as follows. Before making any educational investments, men and women are equally productive. However, if men are overconfident and women underconfident, the proportion of men among the male population who acquire a high education level will be higher than the proportion of women among the female population who acquire a high education level. If education raises productivity, then men will be, on average, more productive than women. This generates the gender pay gap.

Finally, I show that if the fraction of overconfident workers is not too high and workers are sufficiently similar in terms of productivity and cost of education, then biased beliefs can improve welfare. When all workers are unbiased, Spence's (1973) model shows that if the two groups of workers are sufficiently similar in terms of productivity and cost of education, then there exist separating equilibria with overinvestment in education by the more productive group. In this case private information about productive ability reduces welfare. However, Spence (2002) shows that it is possible to improve market efficiency with an optimal tax-subsidy schedule that consists of a rising tax on education which reduces the level of education of high-ability workers combined with a lump-sum transfer to low-ability workers so that net tax revenues are zero.

The reasons why workers' biased beliefs raise welfare are similar to those why a tax-subsidy schedule raises welfare. Overconfidence is like a "tax" on the education of unbiased high-ability workers because it lowers their wage and their education. When the fraction of overconfident workers is not too high, the education level of unbiased high-ability workers will be close to the optimal. Underconfidence is like a "subsidy" for unbiased low-ability workers because it raises their wage for a given education level. This result is consistent with the theory of the second best. According to this theory, introducing a new distortion—workers' biased beliefs—in an environment where another distortion is already present—private information about skill—, may increase welfare.

Of course, welfare does not always rise when workers have biased beliefs and are sufficiently similar in terms of productivity and cost of education. If the fraction of overconfident workers is too high, unbiased high-ability workers are "overtaxed" and end up with an education level far from the

[2] Section 4 contains a brief discussion of the large literature on the gender pay gap.

optimal. In this case there is a transfer of utility from unbiased high-ability to unbiased low-ability workers.

One policy implication of the welfare analysis is that improving workers' self-evaluations will reduce welfare when the fraction of overconfident workers is not too high and workers are sufficiently similar in terms of productivity and cost of education. In contrast, if either (i) workers are sufficiently different or (ii) a significant fraction of workers is overconfident and workers are sufficiently similar, then improving self-evaluations increases welfare.

The assumption that some workers are overconfident and others underconfident is supported by robust empirical evidence on patterns of over- and underestimation in self-evaluation of skills. Overconfidence is a staple finding in psychology and has been shown to be present in individuals' self-assessments of performance in their jobs. According to Myers (1996), a textbook in social psychology, "(...) on nearly any dimension that is both subjective and socially desirable, most people see themselves as better than average."[3]

Kruger and Dunning (1999) find that it is the poorest performers who hold the least accurate evaluations of their skills and performances, grossly overestimating how well their performances stack up against those of their peers. They observe that students performing in the bottom 25% among their peers on tests of grammar, logical reasoning, and humor tend to think that they are performing above the 60% percentile. They also find that top performers consistently underestimate how superior their performances are relative to their peers. In Kruger and Dunning (1999) studies, the top 25% tended to think that their skills lay in the 70-75% percentile, although their performances fell roughly in the 87% percentile.[4]

This paper is an additional contribution to the growing literature on the impact of behavioral biases on markets and organizations. DellaVigna and Malmendier (2004) and Gabaix and Laibson (2006) study market interac-

[3]Baker, Jensen and Murphy (1988) cite a survey of General Electric Company employees according to which 81 percent of a sample of white-collar clerical and technical workers rated their own performance as falling within the top 20 percent of their peers in similar jobs. Myers (1996) cites a study according to which, in Australia, 86 percent of people rate their job performance as above average.

[4]These patterns have been replicated among undergraduates completing a classroom exam (Dunning et al., 2003), medical students assessing their interviewing skills (Hodges, Regehr, and Martin, 2001), clerks evaluating their performance (Edwards et al., 2003), and medical laboratory technicians evaluating their on-the-job expertise (Haun et al., 2000).

tions between sophisticated firms and biased consumers. They find that in competitive markets, biased consumers may be indirectly exploited by sophisticated consumers.

Sandroni and Squintani (2007) investigate the policy implications of overconfidence in insurance markets. They find that compulsory insurance fails to make all agents better off because it is detrimental to low-risk agents. Thus, behavioral biases may weaken asymmetric information rationales for government intervention in insurance markets.

My paper is closely related to literature on the impact of overconfidence on labor market choices. Squintani (1999) studies overconfidence and on-the-job signaling. He finds that overconfident workers choose tasks that are too onerous, fail them, and, dejected by such a failure, settle down for a position inferior to their potential. Hvide (2002) shows that worker overconfidence about productivity outside the firm improves worker welfare. Santos-Pinto (2008, 2010) shows how firms can design optimal contracts to take advantage of worker overconfidence about productivity inside the firm.

Dubra (2004) and Falk, Huffman, and Sunde (2006) study the impact of self-confidence on job search. Dubra (2004) finds that if searchers are not patient, a slightly overconfident one may fare better than unbiased ones because he will search longer. Falk et al. (2006) show that unemployment duration erodes self-confidence and the willingness to continue search. This implies that falling self-confidence can be a complementary mechanism leading to negative duration dependence.

Fang and Moscarini (2005) study the implication of worker overconfidence on the firm's optimal wage-setting policies using the principal-agent approach. Wage contracts provide incentives and affect workers' confidence in their own skills by revealing private information of the firm about workers' skills. They find, using numerical examples, that overconfidence is a necessary condition for a firm to choose a non-differentiation wage policy (the most extreme form of wage compression). This happens because, when ability and effort are complements, a non-differentiation wage policy preserves worker overconfidence which in turn induces higher effort, offsetting the moral hazard inefficiency.

The paper is organized as follows. Section 2 sets-up the model. Section 3 shows that biased beliefs lower the wage spread and compress the wages of unbiased workers. Section 4 shows that gender differences in self-confidence can contribute to the gender pay gap. Section 5 analyzes the impact of biased beliefs on welfare. Section 6 explains how the paper contributes to the existing literature on labor market signaling and discusses extensions of the model. Section 7 concludes. Proofs of all results are in the Appendix.

2 The Model

For each worker there are two possible productive abilities: low, θ_L, and high, θ_H, with $0 < \theta_L < \theta_H$. Nature determines a worker's productive ability and beliefs about marginal cost of acquiring education. The worker chooses a level of education, $e \geq 0$, based on her beliefs. Two firms, 1 and 2, observe the worker's education and then simultaneously make wage offers w_1 and w_2, with $w_i \geq 0$, $i = 1, 2$. The worker accepts the highest of the two wage offers, flipping a coin in case of a tie.

The payoff of a firm that employs a worker with ability θ and education e is $\pi(w, e, \theta) = y(e, \theta) - w$, where $y(e, \theta)$ is the worker's output. The payoff of a firm that does not employ a worker is zero. High-ability workers are more productive: $y_\theta(e, \theta) > 0$. Education does not reduce productivity, i.e., $y_e(e, \theta) \geq 0$ where $y_e(e, \theta)$ is the marginal productivity of education for a worker of ability θ at education e. The marginal productivity of education is non-increasing with education: $y_{ee}(e, \theta) \leq 0$. The marginal productivity of education is non-decreasing with ability: $y_{e\theta}(e, \theta) \geq 0$.

There are four types of workers in the labor market. Unbiased high-ability workers have marginal cost of acquiring education $c_e(e, \theta_H)$ and know it. Unbiased low-ability workers have marginal cost of acquiring education $c_e(e, \theta_L)$ and know it. Overconfident low-ability workers believe their marginal cost of acquiring education is $c_e(e, \theta_H)$ when in fact it is $c_e(e, \theta_L)$. Underconfident high-ability workers believe their marginal cost of acquiring education is $c_e(e, \theta_L)$ when in fact it is $c_e(e, \theta_H)$. Let $\lambda = \Pr(\theta = \theta_H) \in (0, 1)$ be the fraction of high-ability workers, $\nu \in [0, \lambda]$ be the fraction of underconfident high-ability workers, and $\kappa \in [0, 1 - \lambda]$ be the fraction of overconfident low-ability workers. Firms cannot observe a worker's productive ability and beliefs, but know λ, κ and ν.

The utility of an employed worker is $u(w, e, \theta) = w - c(e, \theta)$, where w is the wage offer made by a firm and $c(e, \theta)$ is the cost to a worker with ability θ to obtaining education e. The utility of an unemployed worker is normalized to zero. The cost of no education is zero: $c(0, \theta) = 0$. The cost of education increases with education: $c_e(e, \theta) > 0$, where $c_e(e, \theta)$ is the marginal cost of education for a worker of ability θ at education e. The cost of education decreases with ability: $c_\theta(e,\theta) < 0$. The marginal cost of education increases with education: $c_{ee}(e,\theta) > 0$. The marginal cost of education decreases with ability: $c_{e\theta}(e,\theta) < 0$. This assumption is critical since it means that low-ability workers find signaling more costly than high-ability workers, i.e., for every e, $c_e(e,\theta_L) > c_e(e,\theta_H)$. The assumption is also known as the Spence-Mirrlees single-crossing condition since it implies that the indifference curves of low- and high-ability workers only cross once.

In a separating equilibrium, education choices are determined by workers' beliefs about their marginal cost of acquiring education: underconfident high-ability workers and unbiased low-ability workers choose a low education level, e^{LU}, whereas overconfident low-ability workers and unbiased high-ability workers choose a high education level, e^{HO}, with $e^{HO} \in [\hat{e}^B, \bar{e}^B]$, and $e^{HO} > e^{LU}$. Firms cannot distinguish between underconfident high-ability workers and unbiased low-ability workers because, at the time wage offers are made, both types of workers have the same education level: e^{LU}. Similarly, firms cannot distinguish between overconfident low-ability workers and unbiased high-ability workers because both types of workers display the same education level e^{HO}. However, firms know λ, κ and ν.

Among all workers who choose an education level e^{LU} firms know that fraction $\alpha = \frac{\nu}{1-\lambda-\kappa+\nu}$ has high ability and fraction $1-\alpha = \frac{1-\lambda-\kappa}{1-\lambda-\kappa+\nu}$ has low ability. Among all workers who choose an education level e^{HO} firms know that fraction $\beta = \frac{\kappa}{\lambda+\kappa-\nu}$ has low ability and fraction $1-\beta = \frac{\lambda-\nu}{\lambda+\kappa-\nu}$ has high ability. Hence, the firms' posterior belief that a worker has high ability after observing education level e is

$$\mu(\theta_H|e) = \begin{cases} \alpha, & \text{for } e < e^{HO} \\ 1-\beta, & \text{for } e \geq e^{HO} \end{cases}.$$

The firms' strategy is then

$$w(e) = \begin{cases} (1-\alpha)y(e,\theta_L) + \alpha y(e,\theta_H), & \text{for } e < e^{HO} \\ \beta y(e,\theta_L) + (1-\beta)y(e,\theta_H), & \text{for } e \geq e^{HO} \end{cases}. \tag{1}$$

The firms' strategy is derived from the assumption that firms make zero profits in equilibrium and that firms know λ, κ and ν. Competition between firms implies that the wage offered to each group of workers (those who choose e^{LU} and those who choose e^{HO}) must be a weighted average of the productivities of each type of worker in the group.

In a separating equilibrium the wage of overconfident low-ability workers and unbiased high-ability workers is higher than the wage of underconfident high-ability workers and unbiased low-ability workers. It follows from (1) and $e^{LU} < e^{HO}$ that this condition is satisfied if $\alpha + \beta \leq 1$, or, using the definitions of α and β,

$$(1-\lambda)\nu + \lambda\kappa \leq (1-\lambda)\lambda. \tag{2}$$

Condition (2) says that if the fractions of overconfident and underconfident workers are sufficiently small, education can serve as a signal of productive ability. When the fraction of biased workers is too high, condition (2) is violated and separating equilibria may no longer exist.[5] I assume from now on that condition (2) is satisfied.

In a separating equilibrium underconfident high-ability workers and unbiased low-ability workers do not envy overconfident low-ability workers and unbiased high-ability workers, that is

$$\begin{aligned}(1-\alpha)y(e^{LU},\theta_L) + \alpha y(e^{LU},\theta_H) - c(e^{LU},\theta_L)\\ \geq \beta y(e^{HO},\theta_L) + (1-\beta)y(e^{HO},\theta_H) - c(e^{HO},\theta_L),\end{aligned} \tag{3}$$

and overconfident low-ability workers and unbiased high-ability workers do not envy underconfident high-ability workers and unbiased low-ability workers, that is

$$\begin{aligned}\beta y(e^{HO},\theta_L) + (1-\beta)y(e^{HO},\theta_H) - c(e^{HO},\theta_H) \geq\\ (1-\alpha)y(e^{LU},\theta_L) + \alpha y(e^{LU},\theta_H) - c(e^{LU},\theta_H).\end{aligned} \tag{4}$$

Let $e^*(\sigma,\beta)$ be the solution to $max_e\left[\beta y(e,\theta_L) + (1-\beta)y(e,\theta_H) - c(e,\theta_H)\right]$, where $\sigma = (\theta_L,\theta_H)$. Let the wage and utility associated with $e^*(\sigma,\beta)$ be $w^*(\sigma,\beta) = \beta y(e^*(\sigma,\beta),\theta_L) + (1-\beta)y(e^*(\sigma,\beta),\theta_H)$ and $u^*(\sigma,\beta) = w^*(\sigma,\beta) - c(e^*(\sigma,\beta),\theta_H))$, respectively. Additionally, let $e^*(\sigma,\alpha)$ be the solution to $max_e\left[(1-\alpha)y(e,\theta_L) + \alpha y(e,\theta_H) - c(e,\theta_L)\right]$. Finally, let the wage and utility associated with $e^*(\sigma,\alpha)$ be $w^*(\sigma,\alpha) = (1-\alpha)y(e^*(\sigma,\alpha),\theta_L) + \alpha y(e^*(\sigma,\alpha),\theta_H)$ and $u^*(\sigma,\alpha) = w^*(\sigma,\alpha) - c(e^*(\sigma,\alpha),\theta_L))$.

There are two qualitatively different kinds of separating equilibria. If workers are sufficiently similar in terms of productivity and cost of acquiring

[5] There are always pooling equilibria where all types of workers choose the same education level e. The firms' posterior belief about a worker's productive ability after observing e must be the prior belief, $\mu(\theta_H|e) = \lambda$, which in turn implies that the equilibrium wage is $w = \lambda y(e,\theta_H) + (1-\lambda)y(e,\theta_L)$.

education and the fraction of biased workers is not too high, separation requires "overeducation" by the overconfident low-ability workers and unbiased high-ability workers. This happens when underconfident high-ability workers and unbiased low-ability workers prefer the wage $w^*(\sigma, \beta)$ and the education level $e^*(\sigma, \beta)$ of overconfident low-ability workers and unbiased high-ability workers, that is,

$$w^*(\sigma, \alpha) - c(e^*(\sigma, \alpha), \theta_L) < w^*(\sigma, \beta) - c(e^*(\sigma, \beta), \theta_L). \tag{5}$$

When inequality (5) is satisfied, overconfident low-ability workers and unbiased high-ability workers must choose an education level greater than $e^*(\sigma, \beta)$ to distinguish themselves from underconfident high-ability workers and unbiased low-ability workers, i.e., $e^{HO} \in [\hat{e}^{HO}, \bar{e}^{HO}]$, with $\hat{e}^{HO} > e^*(\sigma, \beta)$.[6]

If workers are sufficiently different in terms of productivity and cost of education or if the fraction of biased workers is sufficiently high, separation does not require "overeducation" by the overconfident low-ability workers and unbiased high-ability workers. In this case it is too expensive for underconfident high-ability workers and unbiased low-ability workers to acquire education $e^*(\sigma, \beta)$, even if doing so would make firms believe that they are overconfident low-ability workers or unbiased high-ability workers and cause them to pay the wage $w^*(\sigma, \beta)$, thus violating inequality (5): $w^*(\sigma, \alpha) - c(e^*(\sigma, \alpha), \theta_L) > w^*(\sigma, \beta) - c(e^*(\sigma, \beta), \theta_L)$.

3 Wage Compression

A key feature of many labor markets is the presence of wage compression across skills (see Garibaldi, 2006, pp. 21). Wage compression refers to a tendency of wages to be equalized across the skill distribution.

Campbell and Kamlani (1997) conducted a survey of 184 US firms and found that pay differentials represented about one half of the productivity differential between any two workers identical in all respects but productivity. Frank (1984a) examined wages and productivities of sales workers and university professors, and found that the more productive workers were paid

[6] Inequality (5) is satisfied if workers are sufficiently similar in terms of productivity and cost of acquiring education and the fraction of biased workers is not too high. Indeed, supposing $y(e, \theta) = \theta e$ and $c(e, \theta) = e^2/2\theta$, then $e^*(\sigma, \alpha) = \theta_L(\theta_L + \alpha\rho)$, $e^*(\sigma, \beta) = \theta_H(\theta_H - \beta\rho)$, where $\rho = \theta_H - \theta_L$, and (5) becomes $\frac{\theta_H}{\theta_L} < 2 - \frac{\theta_L(\theta_L + \alpha\rho)^2}{\theta_H(\theta_H - \beta\rho)^2}$.

less than their marginal product, while the least productive were paid more than their marginal product.

Mourre (2005) provides evidence that there is a compressed wage distribution in Europe. Wage compression mainly occurs in continental and southern countries, whilst no compression is detected in Anglo-Saxon countries and mixed evidence is found in Northern European countries. She also finds that the compression of wages is not uniform across wage levels: there is more wage compression at the lower end of the earnings distribution.

Wage compression has important consequences for labor market outcomes. Lindquist (2005) shows that even low degrees of wage compression lead to large welfare losses from costly unemployment among low-skilled workers. Acemoglu and Pischke (1999) demonstrate that wage compression may encourage employers to offer and pay for general training. By making the firm a residual claimant of productivity increases, a compressed wage structure increases the willingness of firms to finance training of their employees.

Economic theory offers two main explanations for wage compression. The first one identifies exogenous labor market frictions and institutions like mobility costs, trade-unions, efficiency wages, wage floors, or any institution which contributes to raise the reservation wage (e.g., generous unemployment benefits), as sources of wage compression. Freeman (1982) shows that unionized firms appear to have less wage dispersion than non–unionized ones.

The second type of explanations identifies endogenous causes for wage compression. Frank (1984b) shows that if workers value status, then those who put a highest value on prestige will be willing to work for a wage that is lower than their marginal product in return for having lower-level workers around who in return are paid more than their marginal product. In contrast, Lazear (1989) and Milgrom and Roberts (1990) argue that wage inequalities may give rise to rent-seeking behavior within firms when workers change their behavior with the aim of ensuring wage increases. Wage compression reduces uncooperative behavior and may be efficient. Akerlof and Yellen (1990) posit that large wage differentials between groups may be perceived as unfair and lead to reduced effort. Acemoglu and Pischke (1998) suggest asymmetric employer learning as a reason for wage compression. If productivity is only observed by a worker's current employer, then wages do not need to fully reflect productivity differences.

In this section I start by showing that biased beliefs compress wages in the sense that the wage spread with biased workers is smaller than the wage spread with rational workers. After that I show that biased beliefs compress the wages of unbiased workers, i.e., wage differences across unbiased workers are smaller than differences in productivity.

Since the wage spread is the difference between the wages of high and low education workers and these in turn depend on their education investments, I must start by characterizing the impact of biased beliefs on education investments and wages. Proposition 1 characterizes the wage and education levels of low education workers and applies to any separating equilibrium.

Proposition 1: *In a separating equilibrium: (i) the education level of underconfident high-ability workers and unbiased low-ability workers is at least the first-best education level of low-ability workers–*$e^{LU} = e^*(\sigma, \alpha) \geq e^*(\theta_L)$*–, (ii) the wage paid to underconfident high-ability workers and unbiased low-ability workers is greater than the first-best wage of low-ability workers–*$w(e^{LU}) = w^*(\sigma, \alpha) > w^*(\theta_L)$*–, and (iii) the utility of unbiased low-ability workers is greater than the first-best utility of low-ability workers–* $u(w(e^{LU}), e^{LU}, \theta_L) = u^*(\sigma, \alpha) > u^*(\theta_L)$.

The intuition behind Proposition 1 is as follows. Underconfident high-ability workers think (mistakenly) they have a high marginal cost of acquiring education and, like unbiased low-ability workers, choose a low education level. Firms observe this low education level but since they are unable to distinguish each type of worker, they pay a wage that is equal to the average product of underconfident high-ability workers and unbiased low-ability workers. This implies that, for a given education level, the marginal benefit of education is higher for underconfident high-ability workers and unbiased low-ability workers than it would be for low-ability workers if everyone were rational. Since the perceived marginal cost of education is the same, the education level of underconfident high-ability workers and unbiased low-ability workers is at least the first-best education level of low-ability workers. Thus, underconfident high-ability workers and unbiased low-ability workers are paid a higher wage than the first-best wage of low-ability workers since they have a higher average product and at least the same education. Finally, the utility of unbiased low-ability workers is higher than the first-best utility of low-ability workers because the positive direct effect of underconfidence on the wage is larger than the negative effect of shifting the education level away from the optimal one.

Proposition 2 characterizes the wage and education levels of high education workers when the model has a unique separating equilibrium.

Proposition 2: *If workers are sufficiently different in terms of productivity and cost of acquiring education or the fraction of biased workers is sufficiently high–inequality (5) is violated–, then: (i) the education level of overconfident low-ability workers and unbiased high-ability workers is at most the first-best education level of high-ability workers–*$e^{HO} = e^{*}(\sigma, \beta) \leq e^{*}(\theta_H)$*–, (ii) the wage paid to overconfident low-ability workers and unbiased high-ability workers is less than the first-best wage of high-ability workers–*$w(e^{HO}) = w^{*}(\sigma, \beta) < w^{*}(\theta_H)$*–, and (iii) the utility of unbiased high-ability workers is smaller than the first-best utility of high-ability workers–*$u(w(e^{HO}), e^{HO}, \theta_H) = u^{*}(\sigma, \beta) < u^{*}(\theta_H)$*.*

When workers are sufficiently different in terms of productivity and cost of education or the fraction of biased workers is sufficiently high, there is a unique separating equilibrium. Overconfident low-ability workers think (mistakenly) they have a low marginal cost of acquiring education and, like unbiased high-ability workers, choose a high education level. Firms observe this high education level but since they are unable to distinguish each type of worker, they pay a wage that is equal to the average product of overconfident low-ability workers and unbiased high-ability workers. This implies that, for a given education level, the marginal benefit of education is lower for overconfident low-ability workers and unbiased high-ability workers than for high-ability workers if everyone were rational. Since the perceived marginal cost of education is the same, overconfident low-ability workers and unbiased high-ability workers will, at most, acquire the first-best education level of high-ability workers. Thus, overconfident low-ability workers and unbiased high-ability workers are paid a lower wage than the first-best wage of high-ability workers since they have a lower average product and at most the same education. Finally, the utility of unbiased high-ability workers is lower than the first-best utility of high-ability workers because overconfidence has a negative direct effect on the wage and shifts the education level away from the optimal one.

I will now characterize the set of equilibria wage and education levels of high education workers when the model has a continuum of separating equilibria.

Proposition 3: *If workers are sufficiently similar in terms of productivity and cost of acquiring education and the fraction of biased workers is sufficiently small–inequality (5) is satisfied–, then: (i) the education level of overconfident low-ability workers and unbiased high-ability workers belongs* *to $[\hat{e}^{HO}, \bar{e}^{HO}]$, with $\hat{e}^{HO} < \hat{e}^{H}$ and $\bar{e}^{HO} < \bar{e}^{H}$, and (ii) the wage paid to overconfident low-ability workers and unbiased high-ability workers belongs to $[\hat{w}^{HO}, \bar{w}^{HO}]$, with $\hat{w}^{HO} < \hat{w}^{H}$ and $\bar{w}^{HO} < \bar{w}^{H}$.*

If workers are sufficiently similar in terms of productivity and cost of education and everyone is rational there is a continuum of separating equilibria where high-ability workers overinvest in education to distinguish themselves from low-ability workers.[7] In this case $e^H \in [\hat{e}^H, \bar{e}^H]$ with $\hat{e}^H > e^*(\theta_H)$. These various separating equilibria can be Pareto ranked. In all of them a high-ability worker's utility is $y(e^H, \theta_H) - c(e^H, \theta_H)$, a low-ability worker's utility is $u^*(\theta_L)$, and firms earn zero profits. However, a high-ability worker does strictly better in equilibria where she gets a lower level of education (and a lower wage) since this brings her utility closer to the complete information utility $u^*(\theta_H)$. Thus, the separating equilibrium in which the high-ability worker gets education level $\hat{e}^H$ Pareto dominates all others and is called the least cost separating equilibrium. The separating equilibrium in which the high-ability worker gets education level $\bar{e}^H$ is called the most cost separating equilibrium.

Proposition 3 shows that if inequality (5) is satisfied, then the education and wage of overconfident low-ability workers and unbiased high-ability workers in the least (most) cost separating equilibrium with biased workers is smaller than the education and wage, respectively, of high-ability workers in the least (most) cost separating equilibrium with rational workers. Thus, the set of equilibria education-wage levels of high-ability workers is higher than (in the strong set order sense) the set of equilibria education-wage levels of overconfident low-ability workers and unbiased high-ability workers.[8]

The intuition behind the result is as follows. In the least cost separating equilibrium with rational workers, low-ability workers are indifferent between getting their education-wage contract and that of high-ability workers, i.e., $u^*(\theta_L) = u(\hat{e}^H, \theta_L)$. Similarly, in the least cost separating equilibrium with biased workers, underconfident high-ability workers and unbiased low-ability workers are indifferent between getting their education-wage contract and that of overconfident low-ability workers and unbiased high-ability workers,

[7] In this case inequality (5) is satisfied when $\alpha = \beta = 0$, that is, $w^*(\theta_L) - c(e^*(\theta_L), \theta_L) < w^*(\theta_H) - c(e^*(\theta_H, \theta_L)$.

[8] A set $M \subseteq \mathbf{R}$ is as high as another set $N \subseteq \mathbf{R}$ (in the strong set order), written $M \succsim_S N$, if for every $x \in M$ and $y \in N$, $y \geq x$ implies both $x \in M \cap N$ and $y \in M \cap N$. A set M is higher than N, written $M \succ_S N$ if M is as high as N but N is not as high as M.

i.e., $u^*(\sigma, \alpha) = u(\hat{e}^{HO}, \theta_L)$.

We know from Proposition 1 part (iii) that the utility of underconfident high-ability workers and unbiased low-ability workers is higher than the first-best utility of a low-ability worker, i.e., $u^*(\sigma, \alpha) > u^*(\theta_L)$. This implies that, in the least-cost separating equilibrium with biased workers, the utility of underconfident high-ability workers and unbiased low-ability workers of the education-wage contract $(\hat{e}^{HO}, \hat{w}^{HO})$ is higher than the utility of $(\hat{e}^H, \hat{w}^H)$, i.e., $u(\hat{e}^{HO}, \theta_L) > u(\hat{e}^H, \theta_L)$. But this can only be true if $\hat{e}^{HO}$ is lower than $\hat{e}^H$, i.e., overconfident low-ability workers and unbiased high-ability workers choose a lower education level than high-ability workers would if everyone were rational.

The wage of overconfident low-ability workers and unbiased high-ability workers in the least cost separating equilibrium with biased workers is smaller than the wage of high-ability workers in the least cost separating equilibrium with rational workers because the education and productive ability of overconfident low-ability workers and unbiased high-ability workers are smaller than those of high-ability workers.

We are now ready to summarize the impact of workers' biased beliefs on the wage spread.

Corollary 1:
(i) If workers are sufficiently different in terms of productivity and cost of education–inequality (5) is violated when $\alpha = \beta = 0$–, then the equilibrium wage spread with biased workers is smaller than that with rational workers, i.e., $\triangle w^*(\alpha, \beta) < \triangle w^*$;
(ii) If workers are sufficiently similar in terms of productivity and cost of education and the fraction of biased workers is not too high–inequality (5) is satisfied–, then the wage spread in the least (most) cost separating equilibrium with biased workers is smaller than the wage spread in the least (most) cost separating equilibrium with rational workers, i.e., $\triangle\hat{w}^B < \triangle\hat{w}^R$ ($\triangle\bar{w}^B < \triangle\bar{w}^R$).

The intuition behind Corollary 1 is as follows. The presence of underconfident high-ability workers in the low-education group implies that the average product of a worker in that group is higher than the product of a low-ability worker. So, for the same education level, firms must pay a higher wage to workers in the low-education group than they would to low-ability workers if everyone were rational. This implies that the marginal benefit of education is higher for underconfident high-ability workers and unbiased low-ability workers than for low-ability workers if everyone were rational. Given that the perceived marginal cost of education is the same, underconfident high-ability workers and unbiased low-ability workers will acquire at least the first-best education level of low-ability workers. Since underconfident high-ability workers and unbiased low-ability workers have a higher average product and have acquired at least the education level that low-ability workers would acquire if everyone were rational, they must be paid a higher wage.

If workers are sufficiently different in terms of productivity and cost of education, the model has a unique separating equilibrium. In this equilibrium, workers in the high-education group do not need to overinvest in education. The presence of overconfident low-ability workers in the high-education group implies that the average product of a worker in that group is lower than the product of a high-ability worker when everyone is rational. So, for the same education level, firms must pay a lower wage to workers in the high-education group than they would to high-ability workers if everyone were rational. This implies that the marginal benefit of education is lower for overconfident low-ability workers and unbiased high-ability workers than it would be for high-ability workers if everyone were rational. Given that the perceived marginal cost of education is the same, overconfident low-ability workers and unbiased high-ability workers will acquire at most the first-best education level of high-ability workers. Since overconfident low-ability workers and unbiased high-ability workers have a lower average product and have acquired at most the education level high-ability workers would acquire if everyone were rational, they must be paid a lower wage. Hence, when workers are sufficiently different in terms of productivity and cost of education, the equilibrium wage spread with biased workers is smaller than the one with rational workers.

If workers are sufficiently similar in terms of productivity and cost of education and the fraction of biased workers is not too high, the model has a continuum of separating equilibria where workers in the high-education group overinvest in education to distinguish themselves from those in the low-education group. In the least cost separating equilibrium, the incentive compatibility condition of underconfident high-ability workers and unbiased low-ability workers is binding. This can only happen if overconfident low-ability workers and unbiased high-ability workers acquire less education than high-ability workers would if everyone were rational. The lower education and lower average product imply that overconfident low-ability workers and unbiased high-ability workers are paid a lower wage than high-ability workers when everyone is rational. Hence, the wage spread in the least cost separating equilibrium with biased workers is smaller than that in the least cost separating equilibrium with rational workers.[9]

Corollary 1 shows that biased beliefs compress wages in the sense that the wage spread with biased workers is smaller than the wage spread with rational workers. When all workers are rational ($\kappa = \nu = 0$) workers' wages are equal to their productivity, i.e., $w^H = y(e^H, \theta_H)$ and $w^L = y(e^L, \theta_L)$. In the model with biased workers ($\kappa, \nu > 0$) we have four groups of workers with different productivities. Unbiased high-ability workers have productivity $y(e^{HO}, \theta_H)$, overconfident low-ability workers $y(e^{HO}, \theta_L)$, underconfident high-ability workers $y(e^{LU}, \theta_H)$, and unbiased low-ability workers $y(e^{LU}, \theta_L)$. The wage spread across high and low paid workers (or equivalently, high and low education workers) just equals the difference in average productivities across these groups.

The wage spread with biased workers is smaller than that with rational workers because the existence of overconfident low-ability workers lowers the average productivity of the high-education group and the existence of underconfident high-ability workers raises the average productivity of the low-education group. This is a weak form of wage compression because it says nothing about differences in wages across workers relative to differences in productivity.

The existence of workers with biased beliefs compresses the wages of unbiased workers relative to their productivity. This happens because the wage of unbiased high-ability workers is smaller than productivity, $w^{HO} = \beta y(e^{HO}, \theta_L) + (1-\beta) y(e^{HO}, \theta_H) < y(e^{HO}, \theta_H)$, whereas the wage of unbiased low-ability workers is greater than productivity, $w^{LU} = (1-\alpha) y(e^{LU}, \theta_L) + \alpha y(e^{LU}, \theta_H) > y(e^{LU}, \theta_L)$. Hence, the wage spread across unbiased workers is less than the productivity spread: $w^{HO} - w^{LU} < y(e^{HO}, \theta_H) - y(e^{LU}, \theta_L)$.

The model also predicts that the wage spread across biased workers, i.e., the difference in productivities of overconfident low-ability and underconfident high-ability workers, is greater than the productivity spread: $w^{HO} - w^{LU} > y(e^{HO}, \theta_L) - y(e^{LU}, \theta_H)$. This happens because the wage of overconfident low-ability workers is greater than productivity, $w^{HO} > y(e^{HO}, \theta_L)$,

[9] In the most cost separating equilibrium the intuition is similar with the difference that it is the incentive compatibility constraint of overconfident low-ability workers and unbiased high-ability workers that binds.

whereas the wage of underconfident high-ability workers is smaller than productivity, $w^{LU} < y(e^{LU}, \theta_H)$. This is one sense in which wage compression does not hold in this model.

4 Gender Pay Gap

Empirical evidence shows that, on average, women are paid less than men. This is known as the gender pay gap.[10] The gender pay gap may be statistically decomposed into two components: one due to gender differences in measured characteristics, and the other due to "unexplained" and potentially due to discrimination.

Various explanations have been offered to justify the existence of a gender pay gap. The gender pay gap may be due to differences in human capital of men and women (see Mincer and Polachek, 1974). This is consistent with empirical evidence which shows that, on average, men acquire more college education than women and men have more full-time labor market experience than women. According to Eckel and Grossman (2003) the gap might be due to gender differences in risk aversion. Experimental studies have shown that women are less willing than men to take risks or to enter a competitive environment such as a tournament (see Niederle and Vesterlund, 2007).

Labor market discrimination may also affect women's wages. Becker (1971) shows how a preference for men over women, either on the part of employers, employees or costumers, can lead to women being paid less than men. Rothschild and Stiglitz (1982), explain discrimination as arising due to differences in the noise of productivity signals across gender. If output depends upon matching a worker's quality type with a job and women have noisier signals than men, then their quality is more difficult to ascertain and they should be paid a lower wage.

In this section I show that gender differences in self-confidence can contribute to the gender pay gap. I do not pretend to capture the "full picture" about the gender pay gap, but rather show the possibility of a link between

[10] In the US, the gender pay gap is measured as the ratio of female to male median earnings among full-time, year-round workers. According to the Bureau of Labor Statistics (2010), women who worked full time in wage and salary jobs had median weekly earnings of \$657 in 2009. This represented 80 percent of men's median weekly earnings (\$819). At EU level, the gender pay gap is defined as the relative difference in the average gross hourly earnings of women and men within the economy as a whole. Eurostat (2011) found a gender pay gap of 17.5 percent on average in the 27 EU Member States in 2009.

psychological differences between men and women and labor market outcomes.[11]

Gender differences in self-confidence have been extensively documented in a variety of settings. Numerous psychology studies purport to show that men are more (over-)confident than women (see references in Barber and Odean, 2001). For example, testing for the perception of competence on various tasks, Beyer (1990) finds that men tend either to be accurate or to over-estimate their ability, whereas women tend to be either accurate or to under-estimate their ability.

Paglin and Rufolo (1990) report that the propensity of women to choose less mathematical college majors can account for the entire gender wage differential among college graduates. Correll (2001) finds that males are more likely to perceive that they are good at math than are those females with equal math grades and test scores. She also finds that self-assessments of task competence influence career-relevant decisions, even when controlling for commonly accepted measures of ability. For males and females, the higher they rate their mathematical competence, the greater the odds that they will continue on the path leading to careers in the quantitative professions.

Bengtsson, Persson and Willenhag (2005) use the structure of the Economics I exam at Stockholm University to look for gender differences in self-confidence. By answering an extra, optional question, the students can aim for a higher mark. They find that there are striking differences between male and female students in terms of choices and outcomes. They find a clear gender difference in that male students are more inclined than female students to take this opportunity. They also find that female students are slightly better at passing the exam, but male students are much better at getting the highest grade.

To study the impact of gender differences in self-confidence on the gender pay gap I assume that labor supply is composed of males and females. Let λ denote the proportion of high-ability workers in the male and female populations. Thus, before any educational investments are made, men and women are equally productive.

[11] Waldman (1994) provides an evolutionary explanation for gender differences in self-confidence. He considers an environment where individuals compete in wealth accumulation, utility depends on wealth and disutility from effort, and males can overestimate or underestimate their own abilities. He finds that if there is sexual inheritance of the traits disutility from effort and perception of ability, then males exhibiting both disutility from effort and overestimation of abilities can be an evolutionary stable strategy.

Among the low-ability males, proportion κ_m is overconfident and, among the high-ability males, proportion ν_m is underconfident. Among the low-ability females, proportion κ_f is overconfident and, among the high-ability females, proportion ν_f is underconfident. Firms know λ, κ_m, ν_m, κ_f and ν_f. The mean wage paid to males is equal to

$$w_m = (1 - \lambda - \kappa_m + \nu_m) w_m^{LU} + (\lambda + \kappa_m - \nu_m) w_m^{HO}. \tag{6}$$

and mean wage paid to females to

$$w_f = (1 - \lambda - \kappa_f + \nu_f) w_f^{LU} + (\lambda + \kappa_f - \nu_f) w_f^{HO}. \tag{7}$$

If all men and women have correct beliefs there is no gender pay gap since $\nu_m = \kappa_f = \kappa_m = \nu_f = 0$ implies $w_m = w_f$.

To analyze the impact of gender differences in self-confidence on the gender pay gap I assume $y(e, \theta) = e + \theta$ and $c(e, \theta) = e^2/2\theta$. In this case education and productive ability are independent since $y_{e\theta} = 0$. I also assume some males are overconfident–$\kappa_m \in (0, 1 - \lambda]$–and no male is underconfident–$\nu_m = 0$. Finally, I assume some females are underconfident–$\nu_f \in (0, \lambda]$–and no female is overconfident–$\kappa_f = 0$.

Proposition 4: *Let $y(e, \theta) = e + \theta$, $c(e, \theta) = e^2/2\theta$, $\kappa_m \in (0, 1 - \lambda]$, $\nu_f \in (0, \lambda]$, and $\nu_m = \kappa_f = 0$.*
(i) If workers are sufficiently different in terms of productivity and cost of education, i.e., $\frac{\theta_H}{\theta_L} > 3 - 2\min\{\alpha_f, \beta_m\}$, then $w_f^* < w^* < w_m^*$;
(ii) If workers are sufficiently similar in terms of productivity and cost of education, i.e., $\frac{\theta_H}{\theta_L} < 3 - 2\max\{\alpha_f, \beta_m\}$, and $\lambda \leq \frac{2}{3}$, then $\hat{w}_f \leq \hat{w} < \hat{w}_m$;
(iii) If $\frac{\theta_H}{\theta_L} < 3 - 2\max\{\alpha_f, \beta_m\}$, $\lambda > \frac{2}{3}$, and male overconfident is high relative to female underconfidence, i.e., $\frac{\kappa_m}{\nu_f} > \frac{3\lambda - 2}{1 - \lambda}$, then $\hat{w} < \hat{w}_f < \hat{w}_m$;
(iv) If $\frac{\theta_H}{\theta_L} < 3 - 2\max\{\alpha_f, \beta_m\}$, $\lambda > \frac{2}{3}$, and male overconfidence is low relative to female underconfidence, i.e., $\frac{\kappa_m}{\nu_f} < \frac{3\lambda - 2}{1 - \lambda}$, then $\hat{w} < \hat{w}_m < \hat{w}_f$.

Part (i) of Proposition 4 characterizes the impact of gender differences in self-confidence on the gender pay gap when workers are sufficiently different in terms of productivity and cost of education. It shows that if some males are overconfident and some females are underconfident then males will, on average, earn more than females. The intuition behind this result is as follows.

Male overconfidence has two effects on the mean wage of men. First, overconfident low-ability men believe they have a low marginal cost of education and so acquire a high education level. Since education increases productivity, the higher education level of overconfident low-ability men implies that firms pay them a higher wage than the wage firms would pay to low-ability workers if everyone were rational. Second, unbiased high-ability men acquire less education than they would if everyone were rational because firms cannot distinguish between them and overconfident low-ability men and therefore pay them their average productivity. So, unbiased high-ability men are paid a lower wage than would be paid to high-ability workers if everyone were rational. The increase in the wage of overconfident low-ability men is of first-order whereas the decrease in the wage of unbiased high-ability men is of second-order and therefore the mean wage of males is greater than it would be if everyone were rational.

Female underconfidence has two effects on the mean wage of women. First, underconfident high-ability women believe they have high marginal cost of education and so acquire a low education level. Since education increases productivity, the lower education level of underconfident high-ability women implies that firms pay them a lower wage than the wage firms would pay to high-ability workers if everyone were rational. Second, unbiased low-ability women acquire more education than they would if everyone were rational because firms cannot distinguish between them and underconfident high-ability women and therefore pay them their average productivity. Hence, unbiased low-ability women are paid a higher wage than would be paid to low-ability workers if everyone were rational. The decrease in the wage of underconfident high-ability women is of first-order whereas the increase in the wage of unbiased low-ability women is of second-order and therefore the mean wage of females is lower than it would be if everyone were rational.

Parts (ii)-(iv) of Proposition 4 refer to the case when workers are sufficiently similar in terms of productivity and cost of education.[12] Part (ii) shows that if some males are overconfident, some females underconfident, and the fraction of high-ability workers is at most 2/3, then males will, on average, earn more than females. Part (iii) shows that if $\lambda > 2/3$ and male overconfidence is high compared to female underconfidence, then males will earn, on average, more than females. Finally, part (iv) shows that if $\lambda > 2/3$

[12] Cho and Kreps' (1987) intuitive criterion selects the least cost separating equilibrium as the unique prediction of Spence's signaling model. Therefore, when workers are sufficiently similar in terms of productivity and cost of education, I focus on the impact of gender differences in self-confidence on the least cost separating equilibrium mean wages paid to males and to females, $\hat{w}_m$ and $\hat{w}_f$, respectively.

and male overconfidence is low compared to female underconfidence, then females will earn, on average, more than males.

As before, male overconfidence raises the wage of overconfident low-ability males in relation to the wage low-ability workers would get if everyone were rational and lowers the wage of unbiased high-ability males in relation to the wage high-ability workers would get if everyone were rational. The first effect is of first-order whereas the second effect is of second-order and therefore the mean wage of males is greater than it would be if everyone were rational. Also, female underconfidence lowers the wage of underconfident high-ability females in relation to the wage high-ability workers would get if everyone were rational and raises the wage of unbiased low-ability females in relation to the wage low-ability workers would get if everyone were rational. However, when workers are sufficiently similar in terms of productivity and cost of education, female underconfidence has an additional effect on female wages. In the least cost separating equilibrium, unbiased low-ability females must be indifferent between acquiring their low education level and the high education level acquired by unbiased high-ability women. Since female underconfidence raises the wage of unbiased low-ability females in relation to the wage low-ability workers would get if everyone were rational, unbiased high-ability women must overinvest more in education than high-ability workers would if everyone were rational. The higher (over)investment in education increases the productivity of unbiased high-ability females and, therefore, their wage.

When the fraction of high-ability workers is at most $2/3$, the mean wage of women is at most the mean wage workers would get if everyone were rational because the decrease in the wage of underconfident high-ability females dominates the increase in the wages of unbiased low- and high-ability females. When $\lambda > 2/3$ and male overconfidence is high compared to female underconfidence, the mean wage of males raises by more than the mean wage of females relative to the mean wage workers would get if everyone were rational. Hence, males will, on average, earn more than females. When $\lambda > 2/3$ and male overconfidence is low compared to female underconfidence, the mean wage of males raises by less than the mean wage of females relative to the mean wage workers would get if everyone were rational. Hence, females will, on average, earn more than males.

My next result characterizes the impact of gender differences on the gender pay gap when $y(e, \theta) = e\theta$ and $c(e, \theta) = e^2/2\theta$. In this case education and productive ability are complements since $y_{e\theta} > 0$.

Proposition 5: *Let* $y(e,\theta) = e\theta$, $c(e,\theta) = e^2/2\theta$, $\kappa_m \in (0, 1-\lambda]$, $\nu_f \in (0, \lambda]$, $\nu_m = \kappa_f = 0$, $\mu_f = \frac{(\theta_L+\alpha_f\rho)^2}{\theta_H^2}$, $\mu_m = \frac{\theta_L^2}{(\theta_H-\beta_m\rho)^2}$, $\phi = \frac{\theta_H^2-\theta_L^2+\theta_L\theta_H}{\theta_H^2-\theta_L^2-\theta_L\theta_H}$, *and* $\Phi = \frac{\rho\sqrt{\theta_H^2-\theta_L^2}+\theta_H^2+\theta_L\theta_H+\frac{\lambda}{1-\lambda}\theta_L^2}{\rho\sqrt{\theta_H^2-\theta_L^2}-(\theta_L^2+\theta_L\theta_H-\theta_H^2)}$.

(i) If $2 - \frac{\theta_L}{\theta_H}\min\{\mu_f,\mu_m\} < \frac{\theta_H}{\theta_L} \leq \frac{1+\sqrt{5}}{2}$, *then* $\hat{w}_f < \hat{w} \leq \hat{w}_m$*;*

(ii) If $\frac{\theta_H}{\theta_L} > \max\left\{2 - \frac{\theta_L}{\theta_H}\min\{\mu_f,\mu_m\}, \frac{1+\sqrt{5}}{2}\right\}$ *and* $\frac{\kappa_m}{\nu_f} < \phi$, *then* $w_f^* < w_m^* < w^*$*;*

(iii) If $\frac{\theta_H}{\theta_L} > \max\left\{2 - \frac{\theta_L}{\theta_H}\min\{\mu_f,\mu_m\}, \frac{1+\sqrt{5}}{2}\right\}$ *and* $\frac{\kappa_m}{\nu_f} > \phi$, *then* $w_m^* < w_f^* < w^*$*;*

(iv) If $\frac{\theta_H}{\theta_L} < \min\left\{\sqrt{2}, 2 - \frac{\theta_L}{\theta_H}\max\{\mu_f,\mu_m\}\right\}$, *then* $\hat{w}_f \leq \hat{w} < \hat{w}_m$*;*

(v) If $\sqrt{2} < \frac{\theta_H}{\theta_L} < 2 - \frac{\theta_L}{\theta_H}\max\{\mu_f,\mu_m\}$ *and* $\frac{\kappa_m}{\nu_f} < \Phi$, *then* $\hat{w}_f < \hat{w}_m < \hat{w}$*;*

(vi) If $\sqrt{2} < \frac{\theta_H}{\theta_L} < 2 - \frac{\theta_L}{\theta_H}\max\{\mu_f,\mu_m\}$ *and* $\frac{\kappa_m}{\nu_f} > \Phi$, *then* $\hat{w}_m < \hat{w}_f < \hat{w}$.

Parts (i)-(iii) of Proposition 5 characterize the impact of gender differences in self-confidence on the gender pay gap when workers are sufficiently different in terms of productivity and cost of education. Parts (iv)-(vi) refer to the case when workers are sufficiently similar in terms of productivity and cost of education.

Proposition 5 shows that the main qualitative findings in Proposition 4 also apply when education and productive ability are complements: for most parameter configurations, gender differences in self-confidence imply that males will, on average, earn more than females.

The main qualitative difference between the two propositions is that part (iii) of Proposition 5 shows that if workers are sufficiently different in terms of productivity and cost of education, and male overconfidence is high relative to female underconfidence, then females might earn more, on average, than males.[13]

This model only explains a gender pay gap due to measured characteristics: differences in educational investments of males and females. The assumption that education raises productivity, $y_e > 0$, is critical to this result. If education has no direct effect on productivity, $y_e = 0$, then biased beliefs do not lead males' educational investments to differ from those of females

[13] To see this let $\phi(r) = \frac{r^2-1+r}{r^2-1-r}$ where $r = \frac{\theta_H}{\theta_L}$. The inequality $\frac{\kappa_m}{\nu_f} > \phi(r)$ is only satisfied when the proportion of overconfident males is high relative to that of underconfident females since $\phi(r) > 1$ for all $r > 1$ and $\phi(2) = 5$.

and there is no gender pay gap.

Gender differentials in educational investments are well documented. In the US women have recently surpassed men in terms of completing secondary and post-secondary education.[14] Gender differences in field of study composition are larger (more unequal) than those for college attainment. While women are now more likely to complete a college degree than men, the distribution of college majors among college graduates remains unequal with women about 2/3 as likely as men to major in science, mathematics, engineering, business or economics. In contrast, women are much more likely than men to major in humanities, social sciences and teaching (see Gemici and Wiswall, 2011). My model shows how gender differences in self-confidence can contribute to gender differences in educational investments across fields of study which in turn lead to the gender pay gap.

5 Welfare Impact of Biased Beliefs

In this section I characterize the impact of biased beliefs on welfare. To do that I compare welfare levels with biased and rational workers. In both cases firms make zero profits so welfare is equal to the weighted average of the utilities of each group of workers.

To evaluate the utility of a biased worker I take the perspective of an outside observer who knows the worker's actual marginal cost of acquiring education.[15] Hence, welfare with biased workers is

$$\begin{aligned} W^B = &(\lambda-\nu)u(w(e^{HO}),e^{HO},\theta_H)+\nu u(w(e^{LU}),e^{LU},\theta_H) \\ &+\kappa u(w(e^{HO}),e^{HO},\theta_L)+(1-\lambda-\kappa)u(w(e^{LU}),e^{LU},\theta_L). \end{aligned} \tag{8}$$

When all workers are unbiased, Spence's (1973) model shows that if the two groups of workers are sufficiently different, then there is a unique separating equilibrium where investments in education are the efficient ones and

[14] In 2007, the Current Population Survey (2007) estimated that 18,423,000 males ages eighteen and over held a bachelor's degree, while 20,501,000 females ages eighteen and over held one. Fewer males held a master's degree, as well: 6,472,000 males had received one and 7,283,000 females had. However, more men held professional and doctoral degrees than women. 2,033,000 males held professional degrees and 1,079,000 females did and 1,678,000 males had received a doctoral degree, while 817,000 females had.

[15] This is the "hardest" test. An alternative would be to measure the utility of biased workers according to their perceived utility function.

the outcome is as if there was perfect information in the market place. Thus, the existence of workers with biased beliefs acts as a distortion which lowers welfare when workers are sufficiently different.

Let us then focus on the interesting case where workers are sufficiently similar in terms of productivity and cost of education. One of the main results of Spence's (1973) model is that in this case private information about ability reduces welfare. This happens because high-ability workers must overinvest in education (by comparison with the complete information education level) to distinguish themselves from low-ability workers.

My first welfare result shows that if workers are sufficiently similar in terms of productivity and cost of education and all biased workers are overconfident, then biased beliefs reduce welfare.

Proposition 6: *If all biased workers are overconfident and workers are sufficiently similar in terms of productivity and cost of acquiring education–inequality (5) is satisfied when* $\alpha = 0$ *and* $\beta > 0$*–, then* $\hat{W}^R > \hat{W}^B$*, i.e., welfare in the least cost separating equilibrium with rational workers is higher than in the least cost separating equilibrium with overconfident workers.*

This result follows from the fact that the existence of overconfident low-ability workers and the absence of underconfident high-ability workers imply that the utility of unbiased high-ability workers is smaller than the utility of high-ability workers would be if everyone were rational.

The existence of overconfident low-ability workers has two effects on the utility of unbiased high-ability workers. On the one hand, the fact that unbiased high-ability workers are pooled with overconfident low-ability workers implies that the wage of unbiased high-ability workers is lower than that of high-ability in the rational model. The lower wage has a first-order adverse impact on the utility of unbiased high-ability workers. On the other hand, for a fixed education level, the lower wage relaxes the incentive compatibility condition of unbiased low-ability workers. Hence, (over)investment in education by unbiased high-ability workers will be smaller than would be that by high-ability workers if everyone were rational. The lower (over)investment in education has a second-order favorable impact on the utility of unbiased high-ability workers. Hence, unbiased high-ability workers are worse off than high-ability workers when everyone is rational.

Can biased beliefs improve welfare when workers are sufficiently similar in terms of productivity and cost of education? It turns out that the existence of underconfident high-ability workers is a necessary (but not sufficient) condition for biased beliefs to improve welfare. First, the presence of underconfident high-ability workers implies that unbiased low-ability workers attain a higher utility than low-ability workers when everyone is rational–Proposition 2 part (iii). Second, if some high-ability workers are underconfident, then unbiased high-ability workers do not need to overinvest in education as much as high-ability workers when everyone is rational. Hence, the existence of underconfident low-ability workers improves the utility of unbiased low-ability workers and might also improve the utility of unbiased high-ability workers.

To fully characterize the impact of biased beliefs on welfare when workers are sufficiently similar in terms of productivity and cost of education we also need to take into account the utility of biased workers. Generally, we cannot determine how the ex-post utilities of underconfident high-ability workers and overconfident low-ability workers compare to the utilities of high- and low-ability workers, respectively.

To answer this question I specialize the model by assuming $y(e, \theta) = \theta$ and $c(e, \theta) = e^{\gamma}/\theta$, with $\gamma > 1$. In this case education has no impact on productivity like in Spence (1973). My last result provides bounds on the fractions of overconfident low-ability and underconfident high-ability workers under which biased beliefs improve welfare in the specialized model.

Proposition 7: *Let $y(e, \theta) = \theta$ and $c(e, \theta) = e^{\gamma}/\theta$, with $\gamma > 1$. If*

$$\kappa < \frac{(1-\lambda)\theta_L}{(1-\lambda)\theta_L + \lambda\theta_H}, \tag{9}$$

and

$$\lambda(1-\kappa)\frac{\rho}{\theta_H} \leq \nu < \lambda\left(1 - \frac{\kappa}{1-\lambda}\right), \tag{10}$$

then $\hat{W}^B > \hat{W}^R$.

Condition (9) provides an upper bound for the fraction of overconfident low-ability workers and condition (10) provides lower and upper bounds for the fraction of underconfident high-ability workers. When these conditions are satisfied, welfare in the least cost separating equilibrium with biased workers is higher than welfare in the least cost separating equilibrium with rational workers.

The existence of underconfident high-ability workers leads to a first-order increase in the wage of unbiased low-ability workers but only a second-order increase in their cost of education. Thus, unbiased low-ability workers do better than low-ability workers would if everyone were rational. The existence of biased workers lowers the wage and the cost of education of unbiased high-ability workers. If the fraction of overconfident low-ability workers is not too high and the fraction of underconfident high-ability workers is moderate, the fall in cost of education of unbiased high-ability workers is higher than the fall in the wage. Thus, unbiased high-ability workers do better than high-ability workers would if everyone were rational.

6 Discussion

In this section I explain how my analysis contributes to the theoretical and empirical literature on the role of education as a labor market signal. I also discuss some extensions and limitations of the analysis.

In Spence's (1973) basic signaling model education does not raise productivity, workers are perfectly informed about their abilities, all information about a worker's type is collected prior to the entry in the labor market, and there are only two types of workers and one sector. This model has been extended in several directions.

Spence (1974) considers a two-sector model with different returns to ability in the two sectors where education may be productive. Riley (1979a) shows how the model can be extended to the case in which worker types are continuous.

Weiss (1983), Riley (1979b), and Hvide (2003) relax the assumption that workers are perfectly informed about their abilities. Weiss (1983) considers a setting where workers have superior, but imperfect, information about their own abilities and must take a pass-fail test upon their completion of schooling. Riley (1979b) shows that if a worker's career lasts more than a single period and firms learn about true ability, then the importance of the initial education signal as a factor determining compensation should decrease. Hvide (2003) extends Spence (1974) by assuming that workers have imperfect information about their own ability and performance contracts are an alternative sorting mechanism to schooling. He finds that workers with intermediate ability acquire education, while the most able skip it.

In the previous models all information about a worker's type is collected prior to the entry in the labor market. Alós-Ferrer and Prat (2008) extend Spence (1973) by assuming that workers are also able to reveal their ability after the educational signaling stage. They find that the outcomes of the signaling game are affected by the learning process: new equilibria emerge and these can survive the Intuitive Criterion.

My paper extends Spence's (1973) by assuming that some workers have biased beliefs about their marginal cost of acquiring education and abilities (since marginal costs of schooling and productive abilities are negatively related). I show that biased beliefs compress wages and that gender differences in self-confidence can contribute to the gender pay gap.

Many economists have conducted empirical work aimed at exploring the signaling role of education. It has proved difficult, however, to distinguish between human capital and signaling explanations of the observed relationship between education and earnings.[16] One exception is Tyler, Murnane and Willett (2000). Using differential state General Educational Development (GED) passing standards as an identification strategy, Tyler et al. (2000) find that the signaling value of the GED increased the 1995 earnings of young white dropouts on the margin of passing the exams by 10 to 19 percent.

Signaling can be a significant factor in promotion and termination decisions.[17] DeVaro and Waldman (forthcoming) find empirical support for signaling being important for understanding the differences between promotion practices of workers with bachelors and masters degrees, while the evidence concerning the importance of signaling for workers with high school and Ph.D. diplomas is mixed. Gibbons and Katz (1991) note that with asymmetric information, termination is a negative signal of worker quality when the terminating firm is a viable operation. This is not true if the entire firm is closing down. They then look at data from the Displaced Worker Supplements to the Current Population Survey and find strong supporting evidence.

The literature on statistical discrimination and employer learning assumes that worker productivity is hard to observe by firms.[18] This assumption gen-

[16] See Riley (2001), Lange and Topel (2006) and DeVaro and Waldman (forthcoming).

[17] In Waldman (1984) firms compete for labor using the job assignment of a competitor's employee as a signal of his ability. Because an employer does not want to signal the true capacity of a good employee to competitors, employees might not be assigned tasks which maximize their contribution to the firm's profits.

[18] There are two ways of thinking about learning in labor markets. Harris and Holmstrom (1982), Farber and Gibbons (1996), Gibbons and Waldman (1999), and Altonji and Pierret (2001), assume that learning is symmetric, that is, any information revealed about a worker's ability during a worker's career is public knowledge. In contrast, Waldman (1984), Greenwald (1986), and Gibbons and Katz (1991), assume that learning is asymmetric, that is, information about a worker's ability is revealed only to the current employer, while other firms only observe promotion and firing decisions.

erates testable predictions for the interaction between schooling and ability with experience in earning regressions. Farber and Gibbons (1996) find that unmeasured skill becomes increasingly important to wage setting over time, as employers' learn about ability. Altonji and Pierret (2001) find that employers initially use education as a proxy of unobserved ability, and rely on it less as ability becomes known. Lange (2007) estimates that the speed with which employers learn is rapid and argues that signaling can only explain a small fraction of the private gains from schooling.

In a Spence (1973) separating equilibrium there is a one-to-one correspondence between ability and education level, so there is nothing to learn after individuals start working. This is problematic because, as we have seen, the empirical evidence suggests that there is substantial learning after employment starts. My paper provides a potential explanation for this empirical finding which is based on the existence of workers with mistaken beliefs.

The model could be extended by assuming that, after hiring decisions are made and workers are paid their initial wages, firms observe workers' output and readjust wages before the start of an additional production period. In this case the wages of unbiased high-ability workers and underconfident high-ability workers would rise whereas the wages of unbiased low-ability workers and overconfident low-ability workers would fall. This implies that wage compression would fall with worker age. This prediction might not be supported by empirical evidence. For example, Frazis and Loewenstein (2006), using data from the Employment Opportunity Pilot Project, find that only 32 percent of differences in starting productivity (measured by subjective ratings of employers) are reflected in differences in starting wages and that productivity growth of 10 percent results in wage growth of only 2.6 percent.

Another possible extension is to relax the assumption that the utility of a worker who stays at home is zero. For a high enough reservation utility, workers with a high marginal cost of acquiring education might rather stay at home than get educated. The existence of underconfident high-ability workers raises the number of workers who choose to stay at home and can lead to significant welfare losses.

Finally, the model could be extended to include job assignment in the spirit of Gibbons and Waldman (1999, 2006). Suppose that each firm has two different jobs, 1 and 2, and that the efficient assignment for a high productivity worker is job 2, while the efficient assignment for a low productivity worker is job 1. Firms assign workers to jobs on the basis of their education level since workers' productivities are unobservable. In a separating equilibrium, high education workers would be assigned to job 2 and low education workers to job 1. Introducing biased beliefs in this set-up would lead to welfare losses due to job misassignment since firms would assign overconfident workers to job 2 and underconfident workers to job 1.

7 Conclusion

This paper extends Spence's (1973) model by assuming that a fraction of workers in the labor market do not accurately evaluate their marginal cost of acquiring education. More precisely, some workers are overconfident and some underconfident. In addition, I assume that firms know the fractions of high-ability, overconfident and underconfident workers.

I find that the wage spread with biased workers is smaller than that with rational workers. I also find that biased beliefs compress the wages of unbiased workers relative to their productivities. I show that gender differences in self-assessments can contribute to the gender pay gap. If men are overconfident, women underconfident, and education raises productivity, then men will, on average, earn more than women.

Finally, I show that if workers are sufficiently similar in terms of productivity and cost of education, then biased beliefs improve welfare provided that the fraction of overconfident low-ability workers is not too high and the fraction of underconfident high-ability workers is moderate.

The model also provides a potential explanation for the empirical finding that there is substantial learning about workers' productivity after employment starts. At the start of the employment relationship firms cannot distinguish between overconfident low-ability workers and unbiased high-ability workers and pay both types of workers the average productivity associated with their high educational investment. Similarly, firms cannot distinguish between underconfident high-ability workers and unbiased low-ability workers and pay both types of workers the average productivity associated with their low educational investment. As time passes, firms observe workers' output and are able to learn the true skill of each worker which, together with their educational investments, determine their true productivity.

8 Appendix

Proof of Proposition 1:
(i) The education level of underconfident high-ability workers and unbiased low-ability workers, e^{LU}, is the solution to $\max_e (1-\alpha)y(e,\theta_L)+\alpha y(e,\theta_H)-c(e,\theta_L)$. Thus, e^{LU} is implicitly defined by the first-order condition: $(1-\alpha)y_e(e,\theta_L)+\alpha y_e(e,\theta_H)-c_e(e,\theta_L)=0$. The second-order condition is satisfied since $y_{ee}\leq 0$, $c_{ee}>0$, and $\alpha>0$ imply $(1-\alpha)y_{ee}(e,\theta_L)+\alpha y_{ee}(e,\theta_L)-c_{ee}(e,\theta_L)<0$. From the implicit definition of e^{LU} and $y_{e\theta}\geq 0$ we have

$$\frac{\partial e^{LU}}{\partial \alpha}=-\frac{y_e(e,\theta_H)-y_e(e,\theta_L)}{(1-\alpha)y_{ee}(e,\theta_L)+\alpha y_{ee}(e,\theta_L)-c_{ee}(e,\theta_L)}\geq 0. \tag{11}$$

It follows from (11) that $e^{LU}\geq e^*(\theta_L)=\arg\max_e y(e,\theta_L)-c(e,\theta_L)$.
(ii) The wage of underconfident high-ability workers and unbiased low-ability workers is $w(e^{LU})=(1-\alpha)y(e^{LU},\theta_L)+\alpha y(e^{LU},\theta_H)$. The first-best wage of low-ability workers is $w^*(\theta_L)=y(e^*(\theta_L),\theta_L)$. From (i) we know that $e^{LU}\geq e^*(\theta_L)$ which implies $y(e^{LU},\theta_L)\geq y(e^*(\theta_L),\theta_L)$. This, together with $\alpha>0$ and the definitions of $w(e^{LU})$ and $w^*(\theta_L)$, implies $w(e^{LU})>w^*(\theta_L)$.
(iii) The utility of education level e^{LU} for an unbiased low-ability worker is $u(w(e^{LU}),e^{LU},\theta_L)=(1-\alpha)y(e^{LU},\theta_L)+\alpha y(e^{LU},\theta_H)-c(e^{LU},\theta_L)$. We have $\frac{\partial u}{\partial\alpha}=\frac{\partial u}{\partial e}\frac{de^{LU}}{d\alpha}+\frac{\partial u}{\partial\alpha}$. The first term is zero–from the Envelope Theorem–and the sign of the second term is positive since $y_\theta>0$. Hence, $u(w(e^{LU}),e^{LU},\theta_L)>u^*(\theta_L)=y(e^*(\theta_L),\theta_L)-c(e^*(\theta_L),\theta_L)$. *Q.E.D.*

Proof of Proposition 2:
(i) If inequality (5) is violated, the education level of overconfident low-ability workers and unbiased high-ability workers, e^{HO}, is the solution to $\max_e \beta y(e,\theta_L)+(1-\beta)y(e,\theta_H)-c(e,\theta_H)$. Thus, e^{HO} is implicitly defined by the first-order condition $\beta y_e(e,\theta_L)+(1-\beta)y_e(e,\theta_H)-c_e(e,\theta_H)=0$. The second-order condition is satisfied since $y_{ee}\leq 0$, $c_{ee}>0$, and $\beta>0$ imply $\beta y_{ee}(e,\theta_L)+(1-\beta)y_{ee}(e,\theta_H)-c_{ee}(e,\theta_H)<0$. From the implicit definition of e^{HO} and $y_{e\theta}\geq 0$ we have

$$\frac{\partial e^{HO}}{\partial \alpha}=-\frac{-[y_e(e,\theta_H)-y_e(e,\theta_L)]}{\beta y_{ee}(e,\theta_L)+(1-\beta)y_{ee}(e,\theta_H)-c_{ee}(e,\theta_H)}\leq 0. \tag{12}$$

It follows from (12) that $e^{HO}\leq e^*(\theta_H)=\arg\max_e y(e,\theta_H)-c(e,\theta_H)$.
(ii) The wage of overconfident low-ability workers and unbiased high-ability workers is $w(e^{HO})=\beta y(e^{HO},\theta_L)+(1-\beta)y(e^{HO},\theta_H)$. The first-best wage

of high-ability workers is $w^*(\theta_H) = y(e^*(\theta_H), \theta_H)$. From (i) we know that $e^{HO} \leq e^*(\theta_H)$ which implies $y(e^{HO}, \theta_H) \leq y(e^*(\theta_H), \theta_H)$. This, together with $\beta > 0$ and the definitions of $w(e^{HO})$ and $w^*(\theta_H)$, implies $w(e^{HO}) < w^*(\theta_H)$.
(iii) The utility of education level e^{HO} for an unbiased high-ability worker is $u(w(e^{HO}), e^{HO}, \theta_H) = \beta y(e^{HO}, \theta_L) + (1-\beta) y(e^{HO}, \theta_H) - c(e^{HO}, \theta_H)$. We have $\frac{\partial u}{\partial \beta} = \frac{\partial u}{\partial e}\frac{de^{HO}}{d\beta} + \frac{\partial u}{\partial \beta}$. The first term is zero–from the Envelope Theorem–and the sign of the second term is negative since $y_\theta > 0$. Hence, $u(w(e^{HO}), e^{HO}, \theta_H) < u^*(\theta_H) = y(e^*(\theta_H), \theta_H) - c(e^*(\theta_H), \theta_H)$. *Q.E.D.*

Proof of Proposition 3:

(i) In the least cost separating equilibrium with biased workers inequality (3) is binding. Denote the lowest e^{HO} that satisfies inequality (3) by $\hat{e}^{HO}$. In a separating equilibrium with rational workers (i.e., $\alpha = \beta = 0$), inequality (3) becomes

$$y(e^*(\theta_L), \theta_L) - c(e^*(\theta_L), \theta_L) \geq y(e^H, \theta_H) - c(e^H, \theta_L). \qquad (13)$$

In the least cost separating equilibrium with rational workers inequality (13) is binding. Denote the lowest e^H that satisfies inequality (13) by $\hat{e}^H$. Can it be that $\hat{e}^{HO} \geq \hat{e}^H$? No, since in that case inequality (3) would not bind because

$$\begin{aligned}
&(1-\alpha) y(e^{LU}, \theta_L) + \alpha y(e^{LU}, \theta_H) - c(e^{LU}, \theta_L) > y(e^*(\theta_L), \theta_L) - c(e^*(\theta_L), \theta_L) \\
&= y(\hat{e}^H, \theta_H) - c(\hat{e}^H, \theta_H) - [c(\hat{e}^H, \theta_L) - c(\hat{e}^H, \theta_H)] > y(\hat{e}^{HO}, \theta_H) \\
&- c(\hat{e}^{HO}, \theta_H) - \beta \left[y(\hat{e}^{HO}, \theta_H) - y(\hat{e}^{HO}, \theta_L) \right] - [c(\hat{e}^{HO}, \theta_L) - c(\hat{e}^{HO}, \theta_H)],
\end{aligned}$$

where the first inequality follows from Proposition 1 part (iii), the equality follows from the definition of $\hat{e}^H$, and the last inequality follows from: (1) $\hat{e}^{HO} \geq \hat{e}^H > e^*(\theta_H)$ implies $y(\hat{e}^H, \theta_H) - c(\hat{e}^H, \theta_H) \geq y(\hat{e}^{HO}, \theta_H) - c(\hat{e}^{HO}, \theta_H)$, (2) $y_\theta(e, \theta) > 0$ implies $y(\hat{e}^{HO}, \theta_H) > y(\hat{e}^{HO}, \theta_L)$, and (3) $\hat{e}^{HO} \geq \hat{e}^H$, $c_\theta(e, \theta) < 0$, and $c_{e\theta}(e, \theta) < 0$ imply $c(\hat{e}^{HO}, \theta_L) - c(\hat{e}^{HO}, \theta_H) \geq c(\hat{e}^H, \theta_L) - c(\hat{e}^H, \theta_H)$. Hence, it must be that $\hat{e}^{HO} < \hat{e}^H$.
In the most cost separating equilibrium with biased workers inequality (4) is binding. Denote the highest e^{HO} that satisfies inequality (4) by $\bar{e}^{HO}$. In a separating equilibrium with rational workers (i.e., $\alpha = \beta = 0$) inequality (4) reduces to

$$y(e^H, \theta_H) - c(e^H, \theta_H) \geq y(e^*(\theta_L), \theta_L) - c(e^*(\theta_L), \theta_H). \qquad (14)$$

In the most cost separating equilibrium with rational workers inequality (14) is binding. Denote the highest e^H that satisfies inequality (14) by $\bar{e}^H$. Can it be that $\bar{e}^{HO} \geq \bar{e}^H$? No, since in that case inequality (4) is violated because

$$\begin{aligned}
&y(e^{LU}, \theta_L) - c(e^{LU}, \theta_L) + \alpha \left[y(e^{LU}, \theta_H) - y(e^{LU}, \theta_L) \right] \\
&+ \left[c(e^{LU}, \theta_L) - c(e^{LU}, \theta_H) \right] > y(e^*(\theta_L), \theta_L) - c(e^*(\theta_L), \theta_L) \\
&+ \left[c(e^*(\theta_L), \theta_L) - c(e^*(\theta_L), \theta_H) \right] = y(\bar{e}^H, \theta_H) - c(\bar{e}^H, \theta_H) \\
&> y(\bar{e}^{HO}, \theta_H) - c(\bar{e}^{HO}, \theta_H) - \beta \left[y(\bar{e}^{HO}, \theta_H) - y(\bar{e}^{HO}, \theta_L) \right],
\end{aligned}$$

where the first inequality follows from Proposition 1 part (iii) and $c(e^{LU}, \theta_L) - c(e^{LU}, \theta_H) \geq c(e^*(\theta_L), \theta_L) - c(e^*(\theta_L), \theta_H)$, the equality follows from the definition of $\bar{e}^H$, and the last inequality follows from: (1) $\bar{e}^{HO} \geq \bar{e}^H > e^*(\theta_H)$ implies $y(\bar{e}^H, \theta_H) - c(\bar{e}^H, \theta_H) \geq y(\bar{e}^{HO}, \theta_H) - c(\bar{e}^{HO}, \theta_H)$, and (2) $y_\theta(e, \theta) > 0$ implies $y(\bar{e}^{HO}, \theta_H) > y(\bar{e}^{HO}, \theta_L)$. Hence, it must be that $\bar{e}^{HO} < \bar{e}^H$.
(ii) If $\hat{e}^{HO} < \hat{e}^H$, then $w(\hat{e}^{HO}) = \beta y(\hat{e}^{HO}, \theta_L) + (1-\beta)\, y(\hat{e}^{HO}, \theta_H) < y(\hat{e}^H, \theta_H) = w(\hat{e}^H)$. If $\bar{e}^{HO} < \bar{e}^H$, then $w(\bar{e}^{HO}) = \beta y(\bar{e}^{HO}, \theta_L) + (1-\beta) y(\bar{e}^{HO}, \theta_H) < y(\bar{e}^H, \theta_H) = w(\bar{e}^H)$. *Q.E.D.*

Proof of Proposition 4:
When $y(e, \theta) = e + \theta$, the firms' strategy for females is

$$w_f(e) = \begin{cases} (\theta_L + \alpha_f \rho) + e, & \text{for } e < e_f^{HO} \\ \theta_H + e, & \text{for } e \geq e_f^{HO} \end{cases},$$

and that for males is

$$w_m(e) = \begin{cases} \theta_L + e, & \text{for } e < e_m^{HO} \\ (\theta_H - \beta_m \rho) + e, & \text{for } e \geq e_m^{HO} \end{cases},$$

where $\alpha_f = \nu_f/(1 - \lambda + \nu_f)$, $\beta_m = \kappa_m/(\lambda + \kappa_m)$, and $\rho = \theta_H - \theta_L$.
(i) When $\frac{\theta_H}{\theta_L} > 3 - 2\beta_m$ there is a unique separating equilibrium for the male population. Unbiased low-ability males acquire education $e_m^{LU} = \theta_L$ and receive wage $w_m^{LU} = 2\theta_L$. Overconfident low-ability males and unbiased high-ability males acquire education $e_m^{HO} = \theta_H$ and receive wage $w_m^{HO} = 2\theta_H - \beta_m \rho$. Hence, from (6), the mean wage paid to males is equal to

$$\begin{aligned}
w_m^* &= (1 - \lambda - \kappa_m) 2\theta_L + (\lambda + \kappa_m)(2\theta_H - \beta_m \rho) \\
&= 2(1-\lambda)\theta_L + 2\lambda\theta_H + \kappa_m (2 - \beta_m)\rho - \lambda \beta_m \rho \\
&= w^* + \kappa_m \frac{2\lambda + \kappa_m}{\lambda + \kappa_m} \rho - \lambda \frac{\kappa_m}{\lambda + \kappa_m} \rho \\
&= w^* + \kappa_m \rho, \qquad (15)
\end{aligned}$$

and so $w^* < w_m^*$.

When $\frac{\theta_H}{\theta_L} > 3 - 2\alpha$ there is a unique separating equilibrium for the female population. Underconfident high-ability females and unbiased low-ability females acquire education $e_f^{LU} = \theta_L$ and receive wage $w_f^{LU} = 2\theta_L + \alpha_f \rho$. Unbiased high-ability females acquire education $e_f^{HO} = \theta_H$ and receive wage $w_f^{HO} = 2\theta_H$. Hence, from (7), the mean wage paid to females is equal to

$$\begin{aligned} w_f^* &= (1 - \lambda + \nu_f)(2\theta_L + \alpha_f \rho) + (\lambda - \nu_f) 2\theta_H \\ &= 2(1 - \lambda)\theta_L + 2\lambda\theta_H - \nu_f(2 - \alpha_f)\rho + (1 - \lambda)\alpha_f \rho \\ &= w^* - \nu_f \frac{2(1 - \lambda) + \nu_f}{1 - \lambda + \nu_f} \rho + (1 - \lambda) \frac{\nu_f}{1 - \lambda + \nu_f} \rho \\ &= w^* - \nu_f \rho, \end{aligned} \tag{16}$$

and so $w_f^* < w^*$.

When $\frac{\theta_H}{\theta_L} > 3 - 2\min\{\alpha_f, \beta_m\}$, we have $w_m^* - w_f^* = (\kappa_m + \nu_f)\rho > 0$, and ${w_f}^* < w^* < w_m^*$.

(ii)-(iv) When $\frac{\theta_H}{\theta_L} < 3 - 2\beta_m$ there is a continuum of separating equilibria for the male population. In the least cost separating equilibrium, unbiased low-ability males acquire education $e_m^{LU} = \theta_L$ and receive wage $w_m^{LU} = 2\theta_L$. The incentive compatibility condition of unbiased low-ability males is binding and so the education level of overconfident low-ability males and unbiased high-ability males is given by

$$2\theta_L - \frac{\theta_L^2}{2\theta_L} = (\theta_H - \beta_m \rho) + e - \frac{e^2}{2\theta_L}. \tag{17}$$

Solving (17) for e we obtain

$$\hat{e}_m^{HO} = \theta_L + \sqrt{2\theta_L(1 - \beta_m)\rho}.$$

So, overconfident low-ability males and unbiased high-ability males receive wage

$$\hat{w}_m^{HO} = (\theta_H - \beta_m \rho) + \theta_L + \sqrt{2\theta_L(1 - \beta_m)\rho}.$$

Hence, from (6), the mean wage paid to males is equal to

$$\hat{w}_m = (1 - \lambda - \kappa_m) 2\theta_L + (\lambda + \kappa_m) \left[(\theta_H - \beta_m \rho) + \theta_L + \sqrt{2\theta_L(1 - \beta_m)\rho} \right]. \tag{18}$$

Setting $\kappa_m = \beta_m = 0$ in (18) we obtain the least cost separating wage when everyone is rational

$$\hat{w} = (1-\lambda)2\theta_L + \lambda\left(\theta_H + \theta_L + \sqrt{2\theta_L\rho}\right). \tag{19}$$

Using (19) we can rewrite (18) as

$$\begin{aligned}\hat{w}_m = \hat{w} + \kappa_m &\left[(1-\beta_m)\rho + \sqrt{2\theta_L(1-\beta_m)\rho}\right] \\ &- \lambda\left[\beta_m\rho + \sqrt{2\theta_L\rho}\left(1-\sqrt{1-\beta_m}\right)\right].\end{aligned}$$

A first-order Taylor series expansion of $\hat{w}_m$ around $\kappa_m = 0$ gives us

$$\hat{w}_m \approx \hat{w}_m|_{\kappa_m=0} + \left.\frac{\partial \hat{w}_m}{\partial \kappa_m}\right|_{\kappa_m=0} \kappa_m = \hat{w} + \frac{1}{2}\sqrt{2\theta_L\rho}\kappa_m, \tag{20}$$

and so $\hat{w}_m > \hat{w}$.

When $\frac{\theta_H}{\theta_L} < 3-2\alpha_f$ there is a continuum of separating equilibria for the female population. In the least cost separating equilibrium, underconfident high-ability females and unbiased low-ability females acquire education $e_f^{LU} = \theta_L$ and receive wage $w_f^{LU} = 2\theta_L + \alpha_f\rho$. The incentive compatibility condition of underconfident high-ability females and unbiased low-ability females is binding and so the education level of unbiased high-ability females is given by

$$2\theta_L + \alpha_f\rho - \frac{\theta_L^2}{2\theta_L} = \theta_H + e - \frac{e^2}{2\theta_L}. \tag{21}$$

Solving (21) for e we obtain

$$\hat{e}_f^{HO} = \theta_L + \sqrt{2\theta_L(1+\alpha_f)\rho}.$$

So, unbiased high-ability females receive wage

$$\hat{w}_f^{HO} = \theta_H + \theta_L + \sqrt{2\theta_L(1+\alpha_f)\rho}.$$

Hence, from (7), the mean wage paid to females is equal to

$$\hat{w}_f = (1-\lambda+\nu_f)(2\theta_L+\alpha_f\rho) + (\lambda-\nu_f)\left(\theta_H + \theta_L + \sqrt{2\theta_L(1+\alpha_f)\rho}\right). \tag{22}$$

Using (19) we can rewrite (22) as

$$\hat{w}_f = \hat{w} - \nu_f \left[(1-\alpha_f)\rho + \sqrt{2\theta_L(1+\alpha_f)\rho} \right. \\ \left. + (1-\lambda)\alpha_f\rho + \lambda\sqrt{2\theta_L\rho}\left(\sqrt{1+\alpha_f}-1\right) \right].$$

A first-order Taylor series expansion of $\hat{w}_f$ around $\nu_f = 0$ gives us

$$\hat{w}_f \approx \hat{w}_f|_{\nu_f=0} + \left.\frac{\partial \hat{w}_f}{\partial \nu_f}\right|_{\nu_f=0} \nu_f = \hat{w} + \frac{3\lambda-2}{2(1-\lambda)}\sqrt{2\theta_L\rho}\nu_f. \tag{23}$$

The expansion shows that if $\lambda \leq 2/3$, then $\hat{w}_f \leq \hat{w}$. However, if $\lambda > 2/3$, then $\hat{w}_f > \hat{w}$.

When $\frac{\theta_H}{\theta_L} < 3 - 2\max\{\alpha_f, \beta_m\}$, the first-order Taylor series expansions (20) and (23) imply that

$$\hat{w}_m - \hat{w}_f \approx \sqrt{2\theta_L\rho}\left(\frac{1}{2}\kappa_m - \frac{3\lambda-2}{2(1-\lambda)}\nu_f\right). \tag{24}$$

From (20), (23), and (24) we see that $\lambda \leq 2/3$ implies $\hat{w}_f \leq \hat{w} < \hat{w}_m$, which proves (ii). We also see that $\lambda > 2/3$ and $\frac{\kappa_m}{\nu_f} > \frac{3\lambda-2}{1-\lambda}$ imply $\hat{w} < \hat{w}_f < \hat{w}_m$, which proves (iii). Finally, we see that $\lambda > 2/3$ and $\frac{\kappa_m}{\nu_f} < \frac{3\lambda-2}{1-\lambda}$ imply $\hat{w} < \hat{w}_m < \hat{w}_f$, which proves (iv). *Q.E.D.*

Proof of Proposition 5:

When $y(e,\theta) = e\theta$, the firms' strategy for females is

$$w_f(e) = \begin{cases} (\theta_L + \alpha_f\rho)e, & \text{for } e < e_f^{HO} \\ \theta_H e, & \text{for } e \geq e_f^{HO} \end{cases},$$

and that for males is

$$w_m(e) = \begin{cases} \theta_L e, & \text{for } e < e_m^{HO} \\ (\theta_H - \beta_m\rho)e, & \text{for } e \geq e_m^{HO} \end{cases}.$$

(i)-(iii) When $\frac{\theta_H}{\theta_L} > 2 - \frac{\theta_L^3}{\theta_H(\theta_H-\beta_m\rho)^2}$ there is a unique separating equilibrium for the male population. Unbiased low-ability males acquire education $e_m^{LU} = \theta_L^2$ and receive wage $w_m^{LU} = \theta_L^3$. Overconfident low-ability males and unbiased high-ability males acquire education $e_m^{HO} = (\theta_H - \beta_m\rho)\theta_H$ and receive wage

$w_m^{HO} = (\theta_H - \beta_m\rho)^2\theta_H$. Hence, from (6), the mean wage paid to males is equal to

$$w_m^* = (1-\lambda-\kappa_m)\theta_L^3 + (\lambda+\kappa_m)(\theta_H-\beta_m\rho)^2\theta_H.$$

A first-order Taylor series expansion of w_m^* around $\kappa_m = 0$ (it is not necessary to use a second-order expansion since the second order term in the expansion is very small) gives us

$$w_m^* \approx w_m^*|_{\kappa_m=0} + \left.\frac{\partial w_m^*}{\partial \kappa_m}\right|_{\kappa_m=0} \kappa_m = w^* - \rho(\theta_H^2 - \theta_L^2 - \theta_L\theta_H)\kappa_m. \tag{25}$$

The expansion shows that if $\theta_H^2 - \theta_L^2 - \theta_L\theta_H \leq 0$, which is equivalent to $\frac{\theta_H}{\theta_L} \leq \frac{1+\sqrt{5}}{2}$, then $w_m^* \geq w^*$. However, if $\frac{\theta_H}{\theta_L} > \frac{1+\sqrt{5}}{2}$, then $w_m^* < w^*$.
When $\frac{\theta_H}{\theta_L} > 2 - \frac{\theta_L(\theta_L+\alpha_f\rho)^2}{\theta_H^3}$ there is a unique separating equilibrium for the female population. Underconfident high-ability females and unbiased low-ability females acquire education $e_f^{LU} = (\theta_L+\alpha_f\rho)\theta_L$ and receive wage $w_f^{LU} = (\theta_L+\alpha_f\rho)^2\theta_L$. Unbiased high-ability females acquire education $e_f^{HO} = \theta_H^2$ and receive wage $w_f^{HO} = \theta_H^3$. Hence, from (7), the mean wage paid to females is equal to

$$w_f^* = (1-\lambda+\nu_f)(\theta_L+\alpha_f\rho)^2\theta_L + (\lambda-\nu_f)\theta_H^3.$$

A first-order Taylor series expansion of w_f^* around $\nu_f = 0$ gives us

$$w_f^* \approx w_f^*|_{\nu_f=0} + \left.\frac{\partial w_f^*}{\partial \nu_f}\right|_{\nu_f=0} \nu_f = w^* - \rho(\theta_H^2 - \theta_L^2 + \theta_L\theta_H)\nu_f, \tag{26}$$

and so $wf^* < w^*$.
When $\frac{\theta_H}{\theta_L} > 2 - \frac{\theta_L}{\theta_H}\min\left\{\frac{(\theta_L+\alpha_f\rho)^2}{\theta_H^2}, \frac{\theta_L^2}{(\theta_H-\beta_m\rho)^2}\right\}$, the first-order Taylor series expansions (25) and (26) imply that

$$w_m^* - w_f^* \approx -\rho(\theta_H^2 - \theta_L^2 - \theta_L\theta_H)\kappa_m + \rho(\theta_H^2 - \theta_L^2 + \theta_L\theta_H)\nu_f. \tag{27}$$

It follows from (25), (26) and (27) that $\frac{\theta_H}{\theta_L} \leq \frac{1+\sqrt{5}}{2}$ implies $w_f^* < w^* \leq w_m^*$, which proves (i). Additionally, we have $\frac{\theta_H}{\theta_L} > \frac{1+\sqrt{5}}{2}$ and $\frac{\kappa_m}{\nu_f} < \phi$ imply $w_f^* < w_m^* < w^*$, which proves (ii). Finally, we have $\frac{\theta_H}{\theta_L} > \frac{1+\sqrt{5}}{2}$ and $\frac{\kappa_m}{\nu_f} > \phi$ imply $w_m^* < w_f^* < w^*$, which proves (iii).
(iv)-(vi) When $\frac{\theta_H}{\theta_L} < 2 - \frac{\theta_L^3}{\theta_H(\theta_H-\beta_m\rho)^2}$ there is a continuum of separating equilibria for the male population. In the least cost separating equilibrium,

unbiased low-ability males acquire education $e_m^{LU} = \theta_L^2$ and receive wage $w_m^{LU} = \theta_L^3$. The incentive compatibility condition of unbiased low-ability males is binding and so the education level of overconfident low-ability males and unbiased high-ability males is given by

$$\theta_L^3 - \frac{\theta_L^4}{2\theta_L} = (\theta_H - \beta_m\rho)e - \frac{e^2}{2\theta_L}. \tag{28}$$

Solving (28) for e we obtain

$$\hat{e}_m^{HO} = \theta_L\left[\theta_H - \beta_m\rho + \sqrt{(\theta_H - \beta_m\rho)^2 - \theta_L^2}\right].$$

So, overconfident low-ability males and unbiased high-ability males receive wage

$$\hat{w}_m^{HO} = (\theta_H - \beta_m\rho)\theta_L\left[\theta_H - \beta_m\rho + \sqrt{(\theta_H - \beta_m\rho)^2 - \theta_L^2}\right].$$

Hence, from (6), the mean wage paid to males is equal to

$$\hat{w}_m = (1-\lambda-\kappa_m)\theta_L^3 + (\lambda+\kappa_m)(\theta_H - \beta_m\rho)\theta_L\left[\theta_H - \beta_m\rho + \sqrt{(\theta_H - \beta_m\rho)^2 - \theta_L^2}\right].$$

A first-order Taylor series expansion of $\hat{w}_m$ around $\kappa_m = 0$ gives us

$$\hat{w}_m \approx \hat{w}_m|_{\kappa_m=0} + \left.\frac{\partial \hat{w}_m}{\partial \kappa_m}\right|_{\kappa_m=0}\kappa_m = \hat{w} + \rho\theta_L\left(-\rho + \frac{\theta_L^2 + \theta_L\theta_H - \theta_H^2}{\sqrt{\theta_H^2 - \theta_L^2}}\right)\kappa_m. \tag{29}$$

The expansion shows that if $\theta_L^2 + \theta_L\theta_H - \theta_H^2 > \rho\sqrt{\theta_H^2 - \theta_L^2}$ or $\frac{\theta_H}{\theta_L} < \sqrt{2}$, then $\hat{w} < \hat{w}_m$.

When $\frac{\theta_H}{\theta_L} < 2 - \frac{\theta_L(\theta_L + \alpha_f\rho)^2}{\theta_H^3}$ there is a continuum of separating equilibria for the female population. In the least cost separating equilibrium, underconfident high-ability females and unbiased low-ability females acquire education $e_f^{LU} = (\theta_L + \alpha_f\rho)\theta_L$ and receive wage $w_f^{LU} = (\theta_L + \alpha_f\rho)^2\theta_L$. The incentive compatibility condition of underconfident high-ability females and unbiased low-ability females is binding and so the education level of unbiased high-ability females is given by

$$(\theta_L + \alpha_f\rho)^2\theta_L - \frac{(\theta_L + \alpha_f\rho)^2\theta_L^2}{2\theta_L} = \theta_H e - \frac{e^2}{2\theta_L}. \tag{30}$$

Solving (30) for e we obtain

$$\hat{e}_f^{HO} = \theta_L \left[\theta_H + \sqrt{\theta_H^2 - (\theta_L + \alpha_f \rho)^2}\right].$$

So, unbiased high-ability females receive wage

$$\hat{w}_f^{HO} = \theta_H \theta_L \left[\theta_H + \sqrt{\theta_H^2 - (\theta_L + \alpha_f \rho)^2}\right],$$

Hence, from (7), the mean wage paid to females is equal to

$$\hat{w}_f = (1 - \lambda + \nu_f)(\theta_L + \alpha_f \rho)^2 \theta_L + (\lambda - \nu_f)\theta_H \theta_L \left[\theta_H + \sqrt{\theta_H^2 - (\theta_L + \alpha_f \rho)^2}\right].$$

A first-order Taylor series expansion of $\hat{w}_f$ around $\nu_f = 0$ gives us

$$\hat{w}_f \approx \hat{w}_f|_{\nu_f=0} + \left.\frac{\partial \hat{w}_f}{\partial \nu_f}\right|_{\nu_f=0} \nu_f = \hat{w} - \rho\theta_L \left(\rho + \frac{\theta_H^2 + \theta_L\theta_H + \frac{\lambda}{1-\lambda}\theta_L^2}{\sqrt{\theta_H^2 - \theta_L^2}}\right)\nu_f, \quad (31)$$

and so $\hat{w}_f < \hat{w}$.

When $\frac{\theta_H}{\theta_L} < 2 - \frac{\theta_L}{\theta_H} \max\left\{\frac{(\theta_L+\alpha_f\rho)^2}{\theta_H^2}, \frac{\theta_L^2}{(\theta_H-\beta_m\rho)^2}\right\}$, the first-order Taylor series expansions (29) and (31) imply that

$$\begin{aligned}\hat{w}_m - \hat{w}_f \approx\ & \rho\theta_L \left(-\rho + \frac{\theta_L^2 + \theta_L\theta_H - \theta_H^2}{\sqrt{\theta_H^2 - \theta_L^2}}\right)\kappa_m \\ & + \rho\theta_L \left(\rho + \frac{\theta_H^2 + \theta_L\theta_H + \frac{\lambda}{1-\lambda}\theta_L^2}{\sqrt{\theta_H^2 - \theta_L^2}}\right)\nu_f. \quad (32)\end{aligned}$$

From (29), (31), and (32) we see that $\frac{\theta_H}{\theta_L} \leq \sqrt{2}$ implies $\hat{w}_f \leq \hat{w} < \hat{w}_m$, which proves (iv). When $\frac{\theta_H}{\theta_L} > \sqrt{2}$ and $\frac{\kappa_m}{\nu_f} < \Phi$ we have $\hat{w}_f < \hat{w}_m < \hat{w}$, which proves (v). Finally, when $\frac{\theta_H}{\theta_L} > \sqrt{2}$ and $\frac{\kappa_m}{\nu_f} > \Phi$ we have $\hat{w}_m < \hat{w}_f < \hat{w}$, which proves (vi) *Q.E.D.*

Proof of Proposition 6:
We have

$$\begin{aligned}\hat{W}^B - \hat{W}^R = \lambda & \left[u(w(\hat{e}^{HO}), \hat{e}^{HO}, \theta_H) - u(w(\hat{e}^H), \hat{e}^H, \theta_H)\right] \\ & + (1-\lambda)\left[u(w(e^{LU}), e^{LU}, \theta_L) - u^*(\theta_L)\right] \\ & - \nu[u(w(\hat{e}^{HO}), \hat{e}^{HO}, \theta_H) - u(w(e^{LU}), e^{LU}, \theta_H)]. \quad (33)\end{aligned}$$

If $\nu = \alpha = 0$, then $u(w(e^{LU}), e^{LU}, \theta_L) = u^*(\theta_L)$. Hence, (33) reduces to $\hat{W}^B - \hat{W}^R = \lambda \left[u(w(\hat{e}^{HO}), \hat{e}^{HO}, \theta_H) - u(w(\hat{e}^H), \hat{e}^H, \theta_H)\right]$. In the least cost separating equilibrium with overconfident workers, the incentive compatibility condition of unbiased low-ability workers is binding: $u^*(\theta_L) = \beta y(\hat{e}^{HO}, \theta_L) + (1-\beta) y(\hat{e}^{HO}, \theta_H) - c(\hat{e}^{HO}, \theta_L)$. In the least cost separating equilibrium with rational workers, the incentive compatibility condition of low-ability workers is binding: $u^*(\theta_L) = y(\hat{e}^H, \theta_H) - c(\hat{e}^H, \theta_L)$. Hence, we have $\beta y(\hat{e}^{HO}, \theta_L) + (1-\beta) y(\hat{e}^{HO}, \theta_H) - c(\hat{e}^{HO}, \theta_L) = y(\hat{e}^H, \theta_H) - c(\hat{e}^H, \theta_L)$. This is equivalent to

$$\begin{aligned} u(w(\hat{e}^{HO}), \hat{e}^{HO}, \theta_H) = u(w(\hat{e}^H), \hat{e}^H, \theta_H) &- \left[c(\hat{e}^H, \theta_L) - c(\hat{e}^H, \theta_H)\right] \\ &+ \left[c(\hat{e}^{HO}, \theta_L) - c(\hat{e}^{HO}, \theta_H)\right]. \end{aligned}$$

The assumption $c_{e\theta} < 0$ and $\hat{e}^H > \hat{e}^{HO}$ implies $c(\hat{e}^{HO}, \theta_L) - c(\hat{e}^{HO}, \theta_H) < c(\hat{e}^H, \theta_L) - c(\hat{e}^H, \theta_H)$. Thus, $u(w(\hat{e}^{HO}), \hat{e}^{HO}, \theta_H) < u(w(\hat{e}^H), \hat{e}^H, \theta_H)$ which implies $\hat{W}^B < \hat{W}^R$ *Q.E.D.*

Proof of Proposition 7:

When $y(e, \theta) = \theta$, the firms' strategy is

$$w(e) = \begin{cases} \theta_L + \alpha\rho, & \text{for } e < e^{HO} \\ \theta_H - \beta\rho, & \text{for } e \geq e^{HO} \end{cases}.$$

Underconfident high-ability workers and unbiased low-ability workers choose zero education since education is not productive. So, $e^{LU} = e^*(\theta_L) = 0$. Underconfident high-ability workers and unbiased low-ability workers do not envy overconfident low-ability workers and unbiased high-ability workers, i.e., $\theta_L + \alpha\rho \geq \theta_H - \beta\rho - (e^{HO})^\gamma / \theta_L$. The education level of overconfident low-ability workers and unbiased high-ability workers in the least cost separating equilibrium, $\hat{e}^{HO}$, satisfies their incentive compatibility condition as an equality. Thus, $\hat{e}^{HO} = [\rho\theta_L[1 - (\alpha + \beta)]]^{\frac{1}{\gamma}}$. Welfare in the least cost separating equilibrium with biased workers is

$$\hat{W}^B = (\lambda - \nu)\left(\rho + \frac{\theta_L^2}{\theta_H} + \alpha\frac{\rho\theta_L}{\theta_H} - \beta\frac{\rho^2}{\theta_H}\right) + (1 - \lambda + \nu)(\theta_L + \alpha\rho). \quad (34)$$

Welfare in the least cost separating equilibrium with rational workers (i.e., $\alpha = \beta = \kappa = \nu = 0$) is

$$\hat{W}^R = \lambda\left(\rho + \frac{\theta_L^2}{\theta_H}\right) + (1 - \lambda)\theta_L. \quad (35)$$

It follows from (34), (35) that welfare with biased workers is higher than welfare with rational workers if $\alpha > [\nu + (\alpha + \beta)(\lambda - \nu)] \frac{\rho}{\theta_H}$. Substituting α and β and simplifying terms we obtain

$$\nu > \left[\lambda(1 - \kappa) - \frac{(\lambda - \nu)^2}{\lambda + \kappa - \nu} \right] \frac{\rho}{\theta_H}. \tag{36}$$

Since $\lambda = \arg\min_{v \in [0,\lambda]} \frac{(\lambda - \nu)^2}{\lambda + \kappa - \nu}$ we have that $\nu > \lambda(1 - \kappa) \frac{\rho}{\theta_H}$ implies (36). A necessary condition for a separating equilibrium to exist is that the wage of underconfident high-ability workers and unbiased low-ability workers is less than the wage of overconfident low-ability workers and unbiased high-ability workers, i.e., (2) must be satisfied. Solving (2) with respect to ν we obtain

$$\nu \leq \lambda \left(1 - \frac{\kappa}{1 - \lambda} \right). \tag{37}$$

For (36) and (37) to provide a lower and an upper bound for ν, respectively, it must be the case that

$$\lambda(1 - \kappa) \frac{\rho}{\theta_H} < \lambda \left(1 - \frac{\kappa}{1 - \lambda} \right).$$

Solving this inequality with respect to κ we obtain (9). Thus, (9) and (10) imply $\hat{W}^R < \hat{W}^B$. *Q.E.D.*

9 References

Acemoglu, Daron, and Jörn-Steffen Pischke. 1998. Why do firms train: Theory and evidence. *Quarterly Journal of Economics* 113:79-119.

Acemoglu, Daron, and Jörn-Steffen Pischke. 1999. The structure of wages and investment in general training. *Journal of Political Economy* 107:539-572.

Akerlof, George, and Janet Yellen. 1990. The fair-wage effort hypothesis and unemployment. *Quarterly Journal of Economics* 102, no. 2:255-283.

Alós-Ferrer, Carlos, and Julien Prat. 2008. Job market signaling and employer learning. Unpublished manuscript, Department of Economics, Universitat Konstanz.

Altonji, Joseph, and Charles Pierret. 2001. Employer learning and statistical discrimination. *Quarterly Journal of Economics* 116:313-350.

Baker, George, Michael Jensen, and Kevin Murphy. 1988. Compensation and incentives: Practice vs. theory. *Journal of Finance* XLIII, no. 3:593-616.

Barber, Brad, and Terrance Odean. 2001. Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics* 116, no. 1:261–292.

Becker, Gary. 1971. *The economics of discrimination.* Chicago: University of Chicago Press.

Bengtsson, Claes, Mats Persson, and Peter Willenhag. 2005. Gender and overconfidence. *Economics Letters* 86:199–203.

Beyer, Sylvia. 1990. Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology* 59: 960–970.

Campbell, Carl, and Kunal Kamlani. 1997. The reasons for wage rigidity: Evidence from a survey of firms. *Quarterly Journal of Economics* 112:759-89.

Cho, In-Koo, and David Kreps. 1987. Signaling games and stable equilibria. *Quarterly Journal of Economics* 102:179-221.

Correll, Shelley. 2001. Gender and the career choice process: The role of biased self-assessments. *American Journal of Sociology* 106:1691-1730.

Della Vigna, Stefano, and Ulrike Malmendier. 2004. Contract design and self-control: Theory and evidence. *Quarterly Journal of Economics* 119, no. 2:353-402.

DeVaro, Jed, and Michael Waldman. Forthcoming. The signaling role of promotions: Further theory and empirical evidence. *Journal of Labor Economics.*

Dubra, Juan. 2004. Optimism and overconfidence in search. *Review of Economic Dynamics* 7:198-218.

David Dunning, Kerri Johnson, Joyce Ehrlinger, and Justin Kruger. 2003. Why people fail to recognize their own competence. *Current Directions in Psychological Science* 12:83-87.

Eckel, Catherine, and Philip Grossman. 2003. Sex and risk: Experimental evidence. In *Handbook of Experimental Economics.* North-Holland/Elsevier Press.

Edwards, Rodney, Kenneth Kellner, Christopher Sistron, and Elizabeth Magyari. 2003. Medical students self-assessment of performance on an obstetrics and gynecology clerkship. *American Journal of Obstetrics and Gynecology* 188:1078-1082.

European Commission. Eurostat. 2011. *Labour market statistics.* Luxembourg: Publications Office of the European Union.

Falk, Armin, David Huffman, and Uwe Sunde. 2006. Do I have what it takes? Equilibrium search with type uncertainty and non-participation. IZA Discussion Paper no. 2531.

Fang, Hanming, and Giuseppe Moscarini. 2005. Morale hazard. *Journal of Monetary Economics* 52: 749-777.

Farber, Henry, and Robert Gibbons. 1996. Learning and wage dynamics. *Quarterly Journal of Economics* 111:1007-1047.

Frank, Robert. 1984a. Are workers paid their marginal products? *American Economic Review* 74, no. 4:549-571.

———— . 1984b. Interdependent preferences and the competitive wage structure. *RAND Journal of Economics* 15, no. 4:510-520.

Frazis, Harley, and Mark Loewenstein. 2006. Wage compression and the division of returns to productivity growth: Evidence from EOPP. Working Paper 398, U.S. Bureau of Labor Statistics.

Freeman, Richard. 1982. Union wage practices and wage dispersion within establishments. *Industrial and Labor Relations Review* 36, no. 1:3-21.

Gabaix, Xavier, and David Laibson. 2006. Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics* 121, no. 2:505-540.

Garibaldi, Pietro. 2006. *Personnel economics in imperfect labor markets.* Oxford University Press.

Gemici, Ahu, and Matthew Wiswall. 2011. Evolution of gender differences in post-secondary human capital investments: College majors. Unpublished manuscript, Department of Economics, New York University.

Gibbons, Robert, and Lawrence Katz. 1991. Layoffs and lemons. *Journal of Labor Economics* 9:351-380.

Gibbons, Robert, and Michael Waldman. 1999. A theory of wage and promotion dynamics inside firms. *Quarterly Journal of Economics* 114:1321-1358.

———. 2006. Enriching a theory of wage and promotion dynamics inside firms. *Journal of Labor Economics* 24, no. 1:59-106.

Greenwald, Bruce. 1986. Adverse selection in the labour market. *Review of Economic Studies* 53:325-347.

Harris, Milton, and Bengt Holmstrom. 1982. A theory of wage dynamics. *Review of Economic Studies* 49:315-333.

Haun, Daniel, Andrea Zeringue, Argie Leach, and Angela Foley. 2000. Assessing the competence of specimen-processing personnel. *Laboratory Medicine* 31:633-637.

Hodges, Brian, Glenn Regehr, and Dawn Martin. 2001. Difficulties in recognizing one's own incompetence: novice physicians who are unskilled and unaware of it. *Academic Medicine* 76:S87-S89.

Hvide, Hans. 2002. Pragmatic beliefs and overconfidence. *Journal of Economic Behavior and Organization* 48, no. 1:15–28.

Hvide, Hans. 2003. Education and the allocation of talent. *Journal of Labor Economics* 21, no. 4:945-976.

Kruger, Justin, and David Dunning. 1999. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology* 77:1121-1134.

Lange, Fabian. 2007. The speed of employer learning. *Journal of Labor Economics* 25, no. 1:1-35.

Lazear, Edward. 1989. Pay equality and industrial politics. *Journal of Political Economy* 97, no. 3:561-580.

Lindquist, Matthew. 2005. The welfare costs of union wage compression. *European Economic Review* 49:639-658.

Milgrom, Paul, and John Roberts. 1990. The efficiency of equity in organizational decision processes. *American Economic Review* 80, no. 2:154-159.

Mincer, Jacob, and Solomon Polachek. 1974. Family investments in human capital: Earnings of women. *Journal of Political Economy* 82:76–108.

Mourre, Gilles. 2005. Wage compression and employment in Europe: First evidence from the structure of earnings survey 2002. European Commission, Directorate-General for Economic and Financial Affairs, Economic Papers, Brussels.

Myers, David. 1996. *Social psychology.* New York: McGraw-Hill.

Niederle, Muriel, and Lise Vesterlund. 2007. Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics* 122:1067–1101.

Paglin, Morton, and Anthony Rufolo. 1990. Heterogeneous human capital, occupational choice, and male-female earnings differences. *Journal of Labor Economics* 8:123–144.

Riley, John. 1979a. Informational equilibrium. *Econometrica* 47:331-359.

———— . 1979b. Testing the educational screening hypothesis. *Journal of Political Economy* 87:S227-S252.

———— . 2001. Silver signals, twenty-five years of screening and signaling. *Journal of Economic Literature* 39:432-478.

Rothschild, Michael, and Joseph Stiglitz. 1982. A model of employment outcomes illustrating the effect of the structure of information on the level of distribution of income. *Economics Letters* 10:231–236.

Sandroni, Alvaro, and Francesco Squintani. 2007. Overconfidence, insurance, and paternalism. *American Economic Review* 97, no. 5:1994-2004.

Santos-Pinto, Luís. 2008. Positive self-image and incentives in organizations. *Economic Journal* 118:1315-1332

———— . 2010. Positive self-Image in tournaments. *International Economic Review* 51, no. 2:475-496

Spence, Michael. 1973. Job market signaling. *Quarterly Journal of Economics* 87, no. 3:355-374.

———— . 1974. Competitive and optimal responses to signaling: An analysis of efficiency and distribution. *Journal of Economic Theory* 8:296-332.

———— . 2002. Signaling in retrospect and the informational structure of markets. *American Economic Review* 92, no. 3:434-459.

Squintani, Francesco. 1999. On-the-job-signaling and self-confidence. Unpublished manuscript, Department of Economics, Northwestern University.

Tyler, John, Richard Murnane, and John Willett. 2000. Estimating the labor market signaling value of the GED. *Quarterly Journal of Economics* 115, no. 2:431-468.

U.S. Census Bureau. Current Population Survey. 2007. *Annual social and economic supplement.* Washington, DC: U.S. Government Printing Office.

U.S. Department of Labor. Bureau of Labor Statistics. 2010. *Women in the labor force: A databook.* Washington, DC: U.S. Government Printing Office.

Waldman, Michael. 1984. Job assignments, signalling, and efficiency. *Rand Journal of Economics* 15: 255-267.

———— . 1994. Systematic errors and the theory of natural selection, *American Economic Review* 84:482-497.

Weiss, Andrew. 1983. A sorting-cum-learning model of education. *Journal of Political Economy* 91, no. 3:420-442.