

Home Production, Market Production and the Gender Wage Gap: Incentives and Expectations*

Stefania Albanesi[†], Claudia Olivetti[‡]

August 2005 (First draft: April 2005)

Abstract

This paper explores the hypothesis that gender wage differentials arise from the interaction between the intra-household allocation of labor and the contractual relation between firms and workers in the presence of private information on workers' labor market attachment.

In our model, households efficiently choose the contribution of each spouse to home production. Workers with high home hours are less attached to market work. Individual home hours and effort applied to market work are private information. Firms offer incentive compatible labor contracts that are constrained-efficient. Optimal contracts imply earnings that are inversely related to home hours. If firms believe women to be less attached to market work than men, they will offer women contracts with lower earnings and lower hours even in the absence of gender differences in productivity. If firms believe that labor market attachment is the same across genders, they will offer the same contract to male and female workers. Spouses' optimal allocation of home hours will respond to firms' beliefs, thus generating the potential for statistical discrimination by gender.

It is the incentive problem in the labor market that gives rise to statistical discrimination in our model. The central role for informational problems as an important determinant of gender differences in labor market outcomes is motivated by the large variation in gender earnings differentials across industries and occupations observed in the data. We document these differences using Census data and we relate them to the severity of incentive problem across industries and occupations.

*We wish to thank Mark Bills, Peter Gottschalk, Kevin Lang, Marjorie McElroy, Stephanie Schmitt-Grohe, B. Ravikumar, Pedro Teles, Gianluca Violante, and seminar participants at Duke University, the Federal Reserve Bank of Minneapolis, NBER SI, SITE, ESSIM, Northwestern University, the University of Pittsburgh, University of Rochester, Queen's University and SUNY Albany for very helpful comments.

[†]Columbia University. Corresponding author. Address: 1022 International Affairs Building, 420 West 118th Street, New York, NY 10027. Tel. 212-854-4478. E-mail: sa2310@columbia.edu.

[‡]Boston University

1 Introduction

The purpose of this paper is to study the emergence and persistence of gender differences in wages and in the division of labor within the household. We explore the hypothesis that gender wage differentials arise from the interaction between the intra-household allocation of labor and the contractual relation between firms and workers in the presence of private information on workers' labor market attachment.¹

Our theoretical analysis is based on a model in which the intra-household allocation is efficient. We model the households according to Chiappori's (1988, 1997) "collective labor supply" model.² In particular, households are assumed to optimally choose the contribution of each spouse to home production. Workers with high home hours are less attached to market work. Individual home hours and effort applied to market work are private information. Firms, in an extension of the framework developed by Holmstrom and Milgrom (1991), face both adverse selection and moral hazard in contracting with workers. They offer incentive compatible labor contracts that are constrained-efficient. Optimal contracts imply earnings that are inversely related to home hours.

If firms believe women to be less attached to market work than men, they will offer women labor contracts with lower earnings and lower hours even in the absence of gender differences in productivity. This implies that it is efficient for wives to allocate more time to home production. Hence, women will be less attached to market work and firms' expectations will be confirmed. If firms believe that labor market attachment is the same across genders, they will offer the same labor contracts to male and female workers. As a consequence, spouses will face the same earning opportunities, and the efficient intra-household allocation of labor will not be related to gender. Hence, absent ex-ante differences in productivity for male and female workers, there are two types of equilibria. One in which there is a systematic gender differential in earnings, accompanied by a gendered intra-household division of labor; and one in which earnings and home hours are not systematically related to gender. It is statistical discrimination that determines gender differentials in the first type of equilibria. Given that firms use gender as a screening device in this type of equilibrium, discrimination actually reduces the incentive problem for firms, eliminating adverse selection. This property of the model implies that equilibria with gender discrimination will be hard to break, which is consistent with the persistence of the gender earnings gap and the intra-household division of labor observed in the data.³

If, instead, we allow for higher relative productivity of women in home production, female

¹There is a vast empirical literature that studies the gender wage gap and its evolution over time. Altonji and Blank (1999) provide an extensive review of this literature for the post-war period. Goldin (1990) presents an extensive historical analysis for the United States.

²This paradigm does not focus on a particular model of spousal interaction, rather it merely restricts household decisions to be Pareto efficient. This framework is consistent with a variety of "household bargaining" models, as in McElroy and Horney (1981) and Manser and Brown (1980).

³O'Neill (2003) shows that there is still a 10% differential in female and male wages in the U.S. in 2000 that remains unexplained by gender differences in schooling, actual experience and job characteristics. Moreover, PSID data for the period 1976-2001 show that husbands' home hours are roughly one third of wives' and that this difference is stable over time.

workers will be less attached to market work if the relative productivity difference is large enough. In this case, the interaction between the optimal intra-household allocation process and the incentive problem in the labor market serves as an amplification mechanism, so that the gender differential in earnings is larger than the difference in relative productivity.

Our model bridges three literatures: the literature on the sexual division of labor in the Beckerian tradition; the one on statistical discrimination, as in Coate and Loury (1993) and Lundberg and Starz (1983); and finally the literature on incentive contracts and job design, as in Holmstrom and Milgrom (1991). As argued by Becker (1985), what distinguishes gender and racial discrimination is that the feedback on the optimal intra-household division of labor generates a larger impact on earnings. The centerpiece of our model is to identify the source of statistical discrimination with the incentive problem in the interaction between firms and workers, given that labor market attachment, which is inversely related to hours devoted to home production, is private information. The central role of informational problems for gender differences in labor market outcomes is motivated by the large variation in gender earnings differentials across industries and occupations that we find in the data.

We use a variety of data sources to support three important predictions of our model. First, according to our model, gender earning differentials should be higher in industries/occupations in which the incentive problem is more severe. Moreover, this pattern should be stronger for married workers than single workers. We use Census data for year 2000 to show that, for the sample of married workers, gender differentials in earnings are greatest in high management and sales occupations. We argue that incentive problems are most stringent in these occupations. High level managers have a wide range of responsibilities, hence the uncertainty associated with their performance, given their effort should be greater than for low level management. Similarly, sales volumes depend to a large degree on variables that are not directly related to sales personnel's effort. These considerations are less important for production workers. By contrast, this pattern is not observed for the sample of single workers, for which the gender earnings ratio is smallest in high management occupations in most industries. Second, we should observe a negative correlation between the female/male ratio of home hours and the female/male earnings ratio across industries/occupations. This correlation should be stronger for married workers than for single workers. Since the Census does not include information on the hours devoted to home production activities, we use PSID data from the late 1990s to document this fact. We find a negative and significant correlation between the two ratios for married workers. The correlation is halved for single workers. Finally, our model also predict a negative correlation between relative home hours of *husbands* and *wives* and their relative earnings, as would be the case for other efficient models of intra-household allocation. We use the PSID to show that this is true in the data.

Our model highlights the importance of incentives and differences in the pay structure in determining gender differences in earnings. This prediction also resonates with current debates on gender discrimination in personnel policy. For example, in June 2004 a federal judge ruled in favor of class-action status for the Dukes vs Wal-Mart gender discrimination lawsuit. The ruling was based on extensive evidence presented by the plaintiffs, Drogin (2003), showing that women working at Wal-Mart stores face pay disparities in most job categories, and take longer

to enter management positions.⁴ Finally, it is also interesting to note how expectations of a gender wage gap characterize both male and female workers. As documented by Babcock and Laschever (2003): “Women report salary expectations between 3 and 32 percent lower than those of men for the same jobs; men expect to earn 13 percent more than women during their first year of full-time work and 32 percent more at their career peaks.”

Our paper is organized as follows. Section 2 presents the model and discusses the results of numerical simulations. Section 3 reports evidence supporting the model’s predictions. Finally, Section 4 concludes.

2 The Model

The economy is populated by a continuum of agents, ex ante identical except for gender. The population is equally divided by gender, all agents are married and belong to a household. Households are made up of two agents of different gender. Individual utility is defined over market consumption, leisure and consumption of a public home good, which is household specific. The public home good is produced using time of each spouse and the market good. The economy is also populated by a continuum of identical firms. Firms hire workers to produce consumption goods. In this section, we describe the household problem, the firms’ problem and present our definition of equilibrium for this economy.

2.1 Households

Each individual has a utility function:

$$U(c_i, h_i, n_i, e_i) + \theta \log(G), \quad (1)$$

where c_i is individual consumption of the market good, h_i denotes home hours and n_i denotes market hours, and e_i denotes effort applied to market work for $i = f, m$, with f denoting females and m males. We adopt the following CARA specification:

$$U(c, h, n, e) = -\exp(-\sigma [c - v(h, n, e)]),$$

where the coefficient of relative risk aversion σ is strictly greater than 0 and $v(\cdot)$ denotes the cost of market and home work. We assume that v is increasing in all its arguments, twice continuously differentiable and that:

$$v_{hm} > 0, v_{he} > 0. \quad (2)$$

Hence, the marginal cost of market hours and effort is increasing in home hours⁵, and the marginal cost of effort is decreasing in market hours.

⁴Discrimination lawsuits based on analogous complaints were filed by a team of women brokers at Merrill Lynch and by women researchers working at Rand corporation during the summer of 2004. See The New York Times, August 22, 2004 and The New York Times, September 5, 2004, respectively.

⁵For a discussion of this, see Becker (1985).

The production function for the home public good is $G = g(h_f, h_m, k)$, where k is the amount of market good used in home production. We assume that g is increasing in each argument and concave, and restrict attention to specifications in which h_f and h_m are substitutes in the production of the public home good.

Each household is endowed with initial wealth. As in Chiappori (1988, 1997), households decide on consumption of the home public good, on the optimal production of the home good, and on how to share household wealth. Each spouse chooses market hours, effort and private consumption. All decisions occur simultaneously.⁶

Households and individuals take as given the price of the market good and the labor contracts offered by firms, \mathcal{C} . Here, \mathcal{C} is a mapping that specifies earnings w , hours n , and effort e , as a function of gender and home hours: $\mathcal{C} = \{w_i, n_i, e_i\}(h)$ for $i = f, m$ where w_i denotes total earnings. We will describe the derivation of this mapping in sections 2.2 and 2.3. We represent the solution of the spouses' individual choice problem as a value function $V_i(s_i, h_i; \mathcal{C})$.⁷ Then, the households' problem is to choose G, k, h_i and s_i to maximize:

$$\sum_i \lambda_i V_i(s_i, h_i; \mathcal{C}) + \theta \log(G), \quad (\text{H})$$

subject to (3) and $\sum_i s_i + k = a + \Pi$. Here, a is household wealth, s_i is the share of household wealth distributed to each spouse, Π denotes dividend income, and λ_i , for $i = f, m$, represents the weight of each spouse in household decisions.

2.1.1 Solution to the Household Problem

The household choice of G, k and h_f, h_m can further be simplified by first analyzing the intra-household allocation of home hours for a given k and G . This amounts to a cost minimization problem and is independent of the weights λ_i . The opportunity cost of home hours for a spouse is her labor earning potential, which depends on labor contracts. The substitutability of spousal hours in the production of the public home good implies that marginal differences in market earnings will give rise to an asymmetric intra-household allocation of home hours, with the spouse with lower earning potential in market work devoting more time to home production. It is important to note that we interpret the intra-household allocation of home hours as a long term arrangement of the spouses, that may be costly to reverse in the short run. This is consistent with an interpretation of our model as a description of a long-run spousal role allocation within the household.

We adopt the following functional form for g :

$$g(h_f, h_m, k) = \begin{cases} (h_f + h_m)^\delta k^{1-\delta} & \text{for } h_f, h_m > 0 \text{ and } \max\{h_f, h_m\} / \min\{h_f, h_m\} \leq \xi, \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

⁶This way of modelling the household is consistent with a broad class of efficient bargaining models. See Bergstrom (1997) for a review.

⁷The household model does not contemplate a participation decision by workers. This is consistent with the fact that a participation constraint is imposed on the firm problem. The household model can be easily extended to allow for a participation decision.

with $\delta \in [0, 1]$ and $\xi > 1$. This specification captures the externality associated with each spouse's home hours, given the public nature of the home good. At the same time, it imposes that production of the home good be joint, in the sense that both spouses need to provide a minimum contribution⁸.

The technology for home production implies that the optimal allocation of home hours does not depend on the weights λ_i and is given by:

$$\begin{cases} h_f = \xi h_m \text{ if } w_f(h) < w_m(h), \\ h_m = \xi h_f \text{ if } w_f(h) > w_m(h) \\ \text{prob}(h_f = \xi h_m) = 0.5 \text{ if } w_f(h) = w_m(h). \end{cases} \quad (4)$$

When $w_f(h) = w_m(h)$, households are indifferent over the allocation of home hours across spouses. We assume that they randomize. By substituting (3) and (4) into the objective function, the household optimization problem amounts to the choice of h_f , k , and s_i subject to the household budget constraint. The solution to this problem gives rise to the policy functions: $s_i(a, \mathcal{C})$, $h_i(a, \mathcal{C})$, $k(a, \mathcal{C})$ and $G(a, \mathcal{C})$.

We assume that θ is the same for all households. Then, in any equilibrium, there will be two values of home hours in the population, $\{h_L, h_H\}$ with $h_H = \xi h_L$. Workers with high home hours face a higher marginal cost of market work than workers with low home hours. Given this, we interpret workers with high home hours as having *low labor market attachment*.

2.2 Firms and Labor Contracts

Each firm hires one worker. Output is related to hours and effort exerted by the worker, according to:

$$y = f(n, e) + \omega,$$

Here, the function $f(n, e)$ denotes expected output, where f is strictly increasing in both hours, n , and effort, e , twice continuously differentiable and weakly concave. We assume $f_{ne} \geq 0$, so that hours and effort are substitutes in production, and we restrict $e \in \{0, 1\}$. The random variable ω is distributed normally with zero mean and variance $\Sigma > 0$.

Labor contracts, $\{w, n, e\}(h)$, are constrained-efficient. Hours, effort and earnings are chosen to maximize the surplus from the employment relationship subject to feasibility and participation constraint and incentive compatibility constraints. We impose free entry in the firm sector, which requires firm profits to be equal to zero and pins down the level of labor earnings. The properties of labor contracts depend on the informational assumptions. *We assume that workers' effort and home hours are not observed by firms, while output and market hours are observable.* Hence, firms face both moral hazard and adverse selection when contracting with workers⁹. To elucidate the role of these incentive problems in the determination

⁸This specification for g is a special case of a CES. We adopt it for analytical simplicity. However, any functional form for g that allows for some degree of substitutability in spousal home hours would deliver similar predictions.

⁹The CARA preference specification implies that home hours is the only intrinsically relevant worker characteristics for firms.

of earnings, we first derive the properties of constrained-efficient labor contracts when there is no private information and when only home hours are observable.

If home hours and effort are observable, firms choose n and e for an individual with home hours equal to h to solve:

$$\max_{n \in N(h), e \in \{0,1\}} f(n, e) - v(h, n, e). \quad (\text{Problem 1})$$

Here, $N(h)$ is the set of feasible values of n , which may depend on h . The first order conditions for n is:

$$f_n(n, e) - v_n(h, n, e) = 0,$$

which implies that n is decreasing in h for given e , by (2). Optimality also implies that $e = 1$ only if:

$$\frac{f_e(n, 1)}{f_n(n, 1)} \geq \frac{v_e(h, n, 1)}{v_n(h, n, 1)}. \quad (5)$$

Condition (5) states that the marginal rate of substitution between hours and effort in production is greater than their marginal rate of substitution in the workers' utility. We assume that it is satisfied for all values of n and h at \bar{e} , so that the highest effort level will be implemented for all agents absent private information.

If firms observe home hours but effort is not observable, firms face a moral hazard problem. The constrained-efficient labor contract will be of the form $\{w(y), n, e\}(h)$, where the dependence of earnings on output, y , may allow to implement a positive level of effort¹⁰. The need to provide incentives by making earnings depend on output, y , implies that earnings are stochastic. This introduces an inefficiency and reduces the surplus from the employment relationship, given that workers are risk averse. Under CARA utility, without loss of generality, we restrict attention to earnings functions of the form: $w(y) = \bar{w} + \tilde{w}y$. We refer to \bar{w} and $\tilde{w}y$ as the salary and the incentive component of earnings, respectively.

The constrained-efficient labor contract for an individual with home hours h solves the problem:

$$\max_{\bar{w}, \tilde{w}, n \in N(h), e \in \{0,1\}} f(n, e) - v(h, n, e) - \sigma \Sigma (\tilde{w})^2 / 2 \quad (\text{Problem 2})$$

subject to

$$E \left[\frac{u(n, 1)}{u(n, 0)} \right] \geq 1. \quad (6)$$

Here, $f(n, e) - v(h, n, e) - \sigma \Sigma (\tilde{w})^2 / 2$ is the surplus from the employment relationship, (6) is the incentive compatibility constraint and all expectations are taken with respect to the random variable ω .

Given the CARA assumption on preferences, we can write:

$$E \left[\frac{u(n, 1)}{u(n, 0)} \right] = - \exp \{ -\sigma [\tilde{w} [f(n, 1) - f(n, 0)] - v(h, n, 1) + v(h, n, 0)] \}.$$

¹⁰This can be interpreted as a multi-task moral hazard problem, as in Holmstrom and Milgrom (1991). The fact that hours are observable corresponds to there being a perfectly informative signal on hours.

Then, the incentive compatibility constraint simplifies to:

$$\tilde{w} [f(n, e) - f(n, 0)] - v(h, n, e) + v(h, n, 0) \geq 0. \quad (7)$$

We first consider the case in which $e = 1$ is implemented. The first order necessary conditions are:

$$0 = f_n(n, 1) - v_n(h, n, 1) + \mu \{ \tilde{w} [f_n(n, 1) - f_n(n, 0)] - v_n(h, n, 1) + v_n(h, n, 0) \}, \quad (8)$$

$$\mu [f(n, 1) - f(n, 0)] - \sigma \Sigma \tilde{w} \leq 0, \quad (9)$$

with equality at $\tilde{w} > 0$,

$$[\tilde{w} [f(n, 1) - f(n, 0)] - v(h, n, 1) + v(h, n, 0)] \mu = 0, \quad \mu \geq 0,$$

μ is the multiplier on (7).

Condition (8) implies that the optimal value of n is decreasing in h for given e , from (2). Implementing $e = 1$ requires $\tilde{w} > 0$ ¹¹. Since $\Sigma\sigma > 0$, the incentive compatibility constraint will be binding from (9), and:

$$\tilde{w} = \frac{v(h, n, 1) - v(h, n, 0)}{f(n, 1) - f(n, 0)} > 0. \quad (10)$$

The zero profit condition implies that:

$$\bar{w} = f(n, e) - \tilde{w} f(n, e). \quad (11)$$

Then, the fraction of earnings accounted for by incentive pay is equal to \tilde{w} .

Total earnings are decreasing in home hours, since the surplus from the employment relation is also decreasing in home hours. The remaining properties of the earnings function in the optimal contract when high effort is implemented depend on the production function and on the utility cost of working. We summarize them in the following proposition¹².

Proposition 1 *If the optimal contract implements $e = 1$, then:*

- i) if $f_{ne} \geq 0$ and $v_{ne} \leq 0$, \tilde{w} is increasing in h .*
- ii) if $f_{ne} < 0$ and $v_{ne} > 0$, \tilde{w} could be increasing or decreasing in h .*
- iii) for $h_L < h_H$, $\tilde{w}_L f(n_L, 1) > \tilde{w}_H f(n_H, 1)$ if $h_H - h_L$ is sufficiently large.*

If low effort is implemented, the incentive compatibility constraint will not be binding and the resulting value of n will be same as with full information for the case in which $e = 0$. The proof of this proposition is reported in appendix A.

¹¹If n and e are perfect substitutes for the worker, the firm cannot influence and agent's choice of e . We exclude this case.

¹²These predictions would arise from a large class of models with hidden effort, and they are not specific to this example.

The gains from implementing high effort depend on the production technology. The choice of the effort to be implemented in Problem 2 depends on the utility cost of implementing high effort for given h . The multiplier on the incentive compatibility constraint, μ , can be interpreted as a measure of the cost of implementing high effort and as an indicator for the severity of the incentive problem:

$$\mu = \frac{\sigma \Sigma \tilde{w}}{f(n, 1) - f(n, 0)}. \quad (12)$$

If incentive pay, \tilde{w} , is increasing in h , the multiplier μ will be higher for workers with high home hours, which corresponds to a more severe incentive problem for these workers. As shown in Proposition 1, the relation between \tilde{w} and h depends on the degree of substitutability between n and e both in the production function and in the cost of working.

The multiplier also rises with Σ , the variance of output given effort and market hours, and with the coefficient of risk aversion, σ . A higher value of Σ implies that output is a less informative signal of workers' effort. This increases the severity of the incentive problem, especially when \tilde{w} is high. Similarly, the cost of implementing high effort is increasing in the workers' risk aversion.

If both home hours and effort are unobserved, firms face both adverse selection and moral hazard¹³. This introduces additional constraints on optimal contracts, stemming from the incentive compatibility condition on home hours. We refer to these as the adverse selection incentive compatibility constraints. As in the case where home hours are observed, firms will post two contracts, designed respectively for workers with low and high home hours. The adverse selection incentive compatibility constraint implies that workers will self-select the contract appropriate to their level of home hours. However, a binding adverse selection incentive compatibility constraint for one type of worker exacerbates the incentive problem associated with hidden effort. This reduces the surplus generated from the employment relation and influences the level of effort that can be implemented for both types of workers, as well as the level of hours and earnings specified by the optimal contract.

The pattern of binding adverse selection incentive compatibility constraints depends on the properties of the functions f and v . We assume that $f(n, 1) - \sigma \Sigma > 0$ for all $n \in N(h_j)$, $j = L, H$ and we impose $\tilde{w}_j \leq 1$ ¹⁴. This ensures that a worker's utility is increasing in \tilde{w}_j for each type of worker. If $f_{ne} \geq 0$ and $v_{ne} \leq 0$, \tilde{w} is increasing in h from the incentive compatibility constraint for effort. Under these conditions, the adverse selection incentive

¹³MacLeod and Malcomson (1988) study a dynamic model of labor contracts with moral hazard and adverse selection.

¹⁴For the case with moral hazard and adverse selection, we interpret that the salary component of earnings is disbursed at the beginning of the period, while the incentive pay component is disbursed at the end of the period after uncertainty is realized.

The condition $\tilde{w}_j < 1$ for $j = H, L$ does not ensure that the firms make non-negative profits ex ante. However, it does imply that the contract is renegotiation proof. That is the firm does not have an incentive to not pay the worker when output is realized, since ex post profits are also guaranteed to be positive, net of the salary which has already been disbursed.

We do not impose this constraint directly on the problem. We solve a relaxed contracting problem without this constraint and verify ex post that it is not violated.

compatibility constraint for workers with high home hours will not be binding. We focus on this case here.

If π_j is the fraction of workers with home hours equal to h_j for $j = L, H$, constrained-efficient contracts solve the problem¹⁵:

$$\max_{\{T_j, n_j \in N(h_j), e_j \in \{0,1\}\}_{j=L,H}} \sum_{j=L,H} \pi_j \left[f(n_j, e_j) - \sigma \Sigma \frac{(\tilde{w}_j)^2}{2} - T_j - v(h_j, n_j, e_j) \right] \quad (\text{Problem 3})$$

s.t.

$$\tilde{w}_j [f(n_j, 1) - f(n_j, 0)] - v(h_j, n_j, 1) + v(h_j, n_j, 0) \geq 0, \text{ for } j = L, H, \quad (13)$$

$$T_L + \tilde{w}_L f(n_L, e_L) - \sigma \Sigma \frac{(\tilde{w}_L)^2}{2} - v(h_L, n_L, e_L) \geq -\sigma \Sigma \frac{(\tilde{w}_H)^2}{2} + \max_{e=0,1} \{ \tilde{w}_H f(n_H, e) - v(h_L, n_H, e) \}. \quad (14)$$

The variable T_j denotes the informational rent for a type j worker, which stems from adverse selection and also reduces the surplus from the employment relationship. For this problem, $T_L = \bar{w}_L - \bar{w}_H$ and $T_H = 0$ by construction. The variable \bar{w}_H will be pinned down by the zero profit condition.

The moral hazard incentive compatibility constraints (13) are the same as in Problem 2. The adverse selection incentive compatibility constraint, (14), is simplified from:

$$E \left[\frac{u_L(n_L, e_L)}{\max_{e \in \{0,1\}} u_L(n_H, e)} \right] \geq 1,$$

where $u_j(n_k, e_k) = -\exp\{-\sigma [\tilde{w}_k f(n_k, e_k) - v(h_j, n_k, e_k)]\}$ for $j, k = H, L$ using the CARA preference specification. The formulation of this constraint reflects the fact that a worker with low home hours may not find it optimal to choose the same level of effort as a worker with high home hours when offered the same contract.

The first order necessary conditions for this problem assuming $e_L = e_H = 1$ are:

$$\begin{aligned} 0 &= \pi_L (f_{n,L} - v_n(h_L, n_L, e_L)) \\ &+ \mu_L \{ \tilde{w}_L (f_n(n_L, 1) - f_n(n_L, 0)) + v_n(h_L, n_L, 0) - v_n(h_L, n_L, 1) \} \\ &+ \phi_L \{ [\tilde{w}_L f_{n,L} - v_n(h_L, n_L, e_L)] \}, \end{aligned} \quad (15)$$

$$\begin{aligned} 0 &= \pi_H (f_{n,H} - v_n(h_H, n_H, e_H)) \\ &+ \mu_H \{ \tilde{w}_H (f_n(n_H, 1) - f_n(n_H, 0)) + v_n(h_H, n_H, 0) - v_n(h_H, n_H, 1) \} \\ &- \phi_L \{ [\tilde{w}_H f_{n,H} - v_n(h_L, n_H, e_H)] \}, \end{aligned} \quad (16)$$

$$-\pi_L + \phi_L \leq 0, \quad (17)$$

with equality at $\bar{w}_L > 0$,

¹⁵A version of the Revelation Principle holds for this problem, as shown in Laffont and Martimort (2002).

$$0 \geq \mu_L [f(n_L, 1) - f(n_L, 0)] - \pi_L \sigma \Sigma \tilde{w}_L - \phi_L (-f(n_L, e_L) + \sigma \Sigma \tilde{w}_L), \quad (18)$$

with equality at $\tilde{w}_L > 0$,

$$0 \geq \mu_H [f(n_H, 1) - f(n_H, 0)] - \pi_H \sigma \Sigma \tilde{w}_H + \phi_L (-f(n_H, e_H) + \sigma \Sigma \tilde{w}_H), \quad (19)$$

with equality at $\tilde{w}_H > 0$, where μ_j and ϕ_j are the multipliers on the effort and screening incentive compatibility constraints.

The moral hazard incentive compatibility constraints imply:

$$\tilde{w}_j = \frac{v(h_i, n_j, 1) - v(h_i, n_j, 0)}{f(n_j, 1) - f(n_j, 0)},$$

for $j = H, L$.

By (18) and (19), jointly with $f(n_j, 1) > \sigma \Sigma \tilde{w}_j$, a binding adverse selection incentive compatibility constraint for H workers increase the value of μ_H , relative to the case when home hours are observable, while the opposite is true for L workers. Using (17):

$$\begin{aligned} \mu_L &= \frac{\pi_L \sigma \Sigma \tilde{w}_L - \pi_L (f(n_L, e_L) - \sigma \Sigma \tilde{w}_L)}{f(n_L, 1) - f(n_L, 0)}, \\ \mu_H &= \frac{\pi_H \sigma \Sigma \tilde{w}_H + \pi_L (f(n_H, e_H) - \sigma \Sigma \tilde{w}_H)}{f(n_H, 1) - f(n_H, 0)} \end{aligned}$$

Hence, the moral hazard problem is exacerbated for workers with high home hours.

The additional incentive constraint gives rise to a loss of efficiency relative to the case where home hours are known. For parameter values such that high effort is implemented for workers with high and low home hours when home hours are known, if the solution to Problem 3 also features $e_L = e_H = 1$, then n_H will be lower than in the solution to Problem 2. However, depending on parameter values, it is also possible that high effort cannot be implemented for both types of workers in Problem 3.

2.3 Equilibrium

Based on the previous analysis, we allow firms to offer contracts of the form: $\{w_i(y), n_i, e_i\}(h)$ for $i = f, m$, with $w_i(y) = \bar{w}_i + \tilde{w}_i y$. Since gender is observable, contracts are allowed to depend on gender. However, given that the contract space is unrestricted, firms will find it optimal to offer different contracts to female and male workers *only if* they believe the distribution of home hours to be different across genders.

The assumptions on the home production technology and on household preference over θ imply that the population is equally split across agents with home hours equal to h_L and $h_H = \xi h_L$. However, given that individuals of different genders are ex ante identical, the distribution of home hours across genders is not pinned down by technology. Instead, it depends on firms' self-fulfilling beliefs about this distribution. We denote with $\pi(h_j|i)$ for $j = H, L$ and $i = f, m$, firms' belief about the distribution of home hours in the population.

Hence, we can describe the set of available labor contracts as a mapping $\mathcal{C}(\pi(h|f), \pi(h|m))$ from firms' beliefs to a function $\{w_i(y), n_i, e_i\}(h)$ for $i = f, m$.

We say that there is *gender discrimination* when firms believe that the distribution of home hours is different in the population of female and male workers. We now formally define the equilibrium and then illustrate the properties of equilibria with and without gender discrimination.

Definition 1 *An equilibrium is given by a set of firm beliefs on the distribution on labor market attachment, $\pi(h_j|i)$ for $i = f, m$ and $j = H, L$, labor contracts $\mathcal{C}(\pi(h|f), \pi(h|m)) = \{w_i, n_i, e_i\}(h)$ for $i = f, m$, and policy functions for the household problem $\{G, k, h_f, h_m, s_f, s_m\}$, so that:*

- i) Labor contracts solve the firms' problem given beliefs $\pi(h_j|i)$;*
- ii) Household policy functions solve the household problem given labor contracts $\{w_i, n_i, e_i\}(h)$ for $i = f, m$;*
- iii) The resulting distribution of labor market attachment in the population is consistent with firms' beliefs.*

In an *equilibrium with gender discrimination*, beliefs about the distribution of home hours are degenerate. In particular, in an equilibrium with discrimination of female workers, $\pi(h_H|f) = 1$ and $\pi(h_L|m) = 1$. In such an equilibrium, gender acts as a screening device, and the only incentive problem is moral hazard with respect to the worker's hidden effort. Firms determine labor contracts by solving Problem 2, with $h_f = h_H$ and $h_m = h_L$. They will offer male workers the contract corresponding to h_L and the one corresponding to h_H to female workers. Given that earning potential from market work is lower for women, households will find it optimal to set $h_f = \xi h_m$, confirming firms' expectations.

In an *equilibrium without gender discrimination*, firms believe that the distribution of home hours is the same for male and female workers, which implies: $\pi(h_H|f) = 0.5$ and $\pi(h_L|m) = 0.5$. Thus, gender does not reveal information about a worker's home hours. Labor contracts will solve Problem 3. The same selection of labor contracts will be offered to female and male workers. Households will be indifferent over which spouse should be assigned high home hours and they will randomize over this choice according to (4). This results in a symmetric distribution of home hours for male and female workers, thus confirming firms' expectations¹⁶.

2.3.1 Equilibrium with Ex-ante Differences Across Genders

In our model, the three types of equilibria are equally likely, given that there are no ex ante differences across genders. However, the prevailing gender role distinction in most societies is

¹⁶Francois (1998) presents a model in which equilibria with gender wage differentials are self-fulfilling. His result relies on exogenously given job heterogeneity. One class of jobs operate under an efficiency wage setting while a second class of jobs operate under piece rate wage setting. Earnings are higher in the efficiency wage jobs. In an equilibrium with female wage discrimination, the first class of jobs is assigned to men, the second to women. The female wage differential stems from job segregation. If all workers were to operate under the same job, the gender wage gap would be reversed. Hence, this model cannot account for gender differentials within the same occupational categories.

one in which men specialize in market production and women in home production. Typically, gender differences in labor market outcomes, such as the earnings gap, have been ascribed to this division of labor, which is seen as the result of biological differences- women’s ability to bear children. In this section, we explore this argument in the context of our model.

We assume that female and male workers are equally productive in market work, but female workers are more productive in home work. Specifically, we posit that the production function for the public home good is given by:

$$G = k^\delta [h_m + (1 + \varepsilon) h_f]^{1-\delta},$$

where $\varepsilon > 0$ and $h_m, h_f \in \{h_L, h_H\}$. A strictly positive value of ε corresponds to higher relative productivity in market vs. home work of male workers with respect to female workers. Women’s higher relative productivity in home production is related to their ability to bear children. The parameter ε can be interpreted as tied to the decreased relative market productivity of women during and after pregnancy. Alternatively, if children are viewed as a component of the public home good, ε captures women’s higher relative contribution to the nourishment of children via breast feeding. Hence, lower values of ε can be interpreted as corresponding to technological advances, such as medical improvements reducing the physical stress associated with pregnancy and the introduction of baby formula.

The following result holds.

Proposition 2 *In the economy with homogeneous households, if θ and a are such that $h_f \neq h_m$, there exists $\bar{\varepsilon}(h_L, h_H) > 0$, such that:*

- i) for $\varepsilon < \bar{\varepsilon}(h_L, h_H)$, there are three possible equilibrium distribution of beliefs $\pi(h_H|f) = 1$ and $\pi(h_L|m) = 1, \pi(h_H|m) = 1$ and $\pi(h_L|f) = 1$, and $\pi(h_H|f) = 0.5$ and $\pi(h_L|m) = 0.5$;*
- ii) for $\varepsilon \geq \bar{\varepsilon}(h_L, h_H)$, there exists one equilibrium distribution of beliefs, with $\pi(h_H|f) = 1, \pi(h_L|m) = 1$.*

Proof. TBA ■

If ε is greater than $\bar{\varepsilon}(h_H, h_L)$, no self-fulfilling equilibria are possible. In this case, the incentive problem in the labor market provides an amplification mechanism for the relative productivity differences across genders. Hence, the equilibrium gender differences in earnings will be greater than the relative productivity difference. The threshold $\bar{\varepsilon}(h_H, h_L)$ depends on the home production technology, specifically, on the ratio h_H/h_L . Technological changes that impact the complementarity between spouses’ hours in the production of the public home good change this threshold and the region of self-fulfilling equilibria.

Based on this result, we can interpret the prevailing pattern of gender specialization in the context of our model in the following way. Initially, high values of ε due to poor medical technology imply that the only possible equilibrium is one in which women are mostly devoted to home production and men specialize in market work. Subsequent improvements in medical technology reduce the value of ε , thus making ungendered equilibria possible. However, the self-fulfilling nature of equilibria with gender discrimination for low ε , coupled with the gendered initial conditions, may have contributed to the persistence in gender differences in

labor market outcomes and household roles, despite the lack of significant differences in relative productivities.

2.4 Numerical Experiments

We illustrate the properties of efficient labor contracts and the resulting implications for the gender earnings differentials numerically for equilibria with and without gender discrimination. We restrict attention to the following functional forms. The disutility from labor is:

$$v(h, n, e) = -\gamma \log(T - h - \eta(n, e)), \quad (20)$$

where η represents the time cost of market work. We assume:

$$\eta(n, e) = [\nu(n+1)^\rho + (1-\nu)(e+1)^\rho]^{1/\rho} - 1, \quad (21)$$

with $\rho, \nu \in (0, 1)$, which implies $\eta_{ne} > 0$ and $v_{ne} > 0$. The production function is:

$$f(n, e) = [(1-\phi)(1+\tau e)^z + \phi(1+n)^z]^{1/z} - 1, \quad (22)$$

with $\tau > 0$ and $\phi, z \in (0, 1)$. The functional forms for η and f are extremely flexible and allow us to explore the properties of optimal contracts as a function of the degree of substitutability between effort and hours in each function.

Our benchmark parameterization is based on empirical evidence where possible and is reported in Table 2. The value of δ is set following Benhabib, Rogerson and Wright (1991). The parameter ξ is chosen to be consistent with the ratio of home hours for married women and men, which ranges from 4.3 in 1976 to 2.8 in 1993 based on PSID data. We set ξ equal to 3, which is the average value of this statistic for the 1990's. The values of γ and T are set so that the ratio of market to home hours matches data from the PSID. The average of this ratio for the 1990's is equal to 1.25 for married female workers and to 6.04 for married male workers. The parameters ρ, ν, z, ϕ, τ and Σ , which pertain to the utility cost of working and the firm's production technology, cannot be calibrated based on aggregate data. We restrict attention to the range $\phi, \nu > 0.5$ and $z, \rho > 0.5$. We set τ so that it is optimal to implement high effort for both types of workers, given the other parameters. The value assigned to the parameter Σ implies that the standard deviation of output conditional on effort and hours is approximately equal to 9% for males and 10% for females.

σ	γ	T	ν	ρ	ϕ	τ	Σ	z	δ	ξ	θ
1.1	0.3	1	0.68	0.8	0.68	2.5	1	0.75	0.68	3	1

We study the basic properties of our model in equilibria with and without female discrimination¹⁷. Table 3 reports summary information for the equilibrium with female discrimination:

¹⁷Equilibria with male discrimination, which are also possible, are the same as equilibria with female discrimination, with all outcomes relabelled by gender.

	w	\tilde{w}	\bar{w}	n	e	Σ/y	h	y	n/h
Male	3.32	0.89	0.33	0.66	1	0.09	0.1	3.3216	6.6
Female	2.88	0.91	0.24	0.37	1	0.10	0.3	2.8808	1.23

The female/male earnings ratio is equal to 0.89. We use a simple Oaxaca-Blinder decomposition of female earnings, at the male parameters, to evaluate the component of the earnings gap not explained by differences in market hours. The decomposition is based on the earnings schedule, hours and output generated by the model. We find that 13% of the earnings gap is not explained by differences in hours worked on the market across genders. These values are consistent with the corresponding data statistics, reported by O’Neill (2000). Using CPS data, she shows that the gender earnings ratio is equal to 0.80, and that there a 10% unexplained gender differential in earnings after controlling for observable differences in human capital and job characteristics.

The properties of the equilibrium without gender discrimination for our benchmark parameterization are displayed in table 4:

	w	\tilde{w}	n	e	n/h
h_L	2.14	0	0.80	0	8
h_H	1.74	0	0.55	0	1.83

In this case, low effort is implemented for both types of workers, despite the fact that high effort is optimal for both males and females in the equilibrium with gender discrimination. Total earnings are lower for both types of workers relative to the equilibrium with discrimination. The ratio of earnings for workers with high home hours relative to workers with low home hours is equal to 0.81- lower than in the equilibrium with gender discrimination.

We now focus on the equilibrium with female discrimination in which female workers have high home hours and the only incentive problem is moral hazard due to hidden effort. A firm may decide to implement low or high effort for a worker depending on the cost of implementing high effort. Such a cost, which is larger for female workers, depends on the degree of substitutability of n and e in utility and in production and on the workers’ relative cost of effort, as discussed in section 2.2.

To investigate this, we vary each relevant parameter in isolation, maintaining the other parameters at their benchmark values. Figure 1 displays how the fraction of incentive pay in earnings for females and males and the earnings ratio vary with ρ , which corresponds to the degree of complementarity between output and hours in the utility cost of working. The dashed and dash-dotted lines represent the fraction of incentive pay in earnings, which is equal to \tilde{w} , for females and males, respectively. The solid line represents the earnings ratio. All numbers are expressed in percentages. The vertical line corresponds to our benchmark parameterization. As ρ increases, the fraction of incentive pay increases for both male and female workers, but

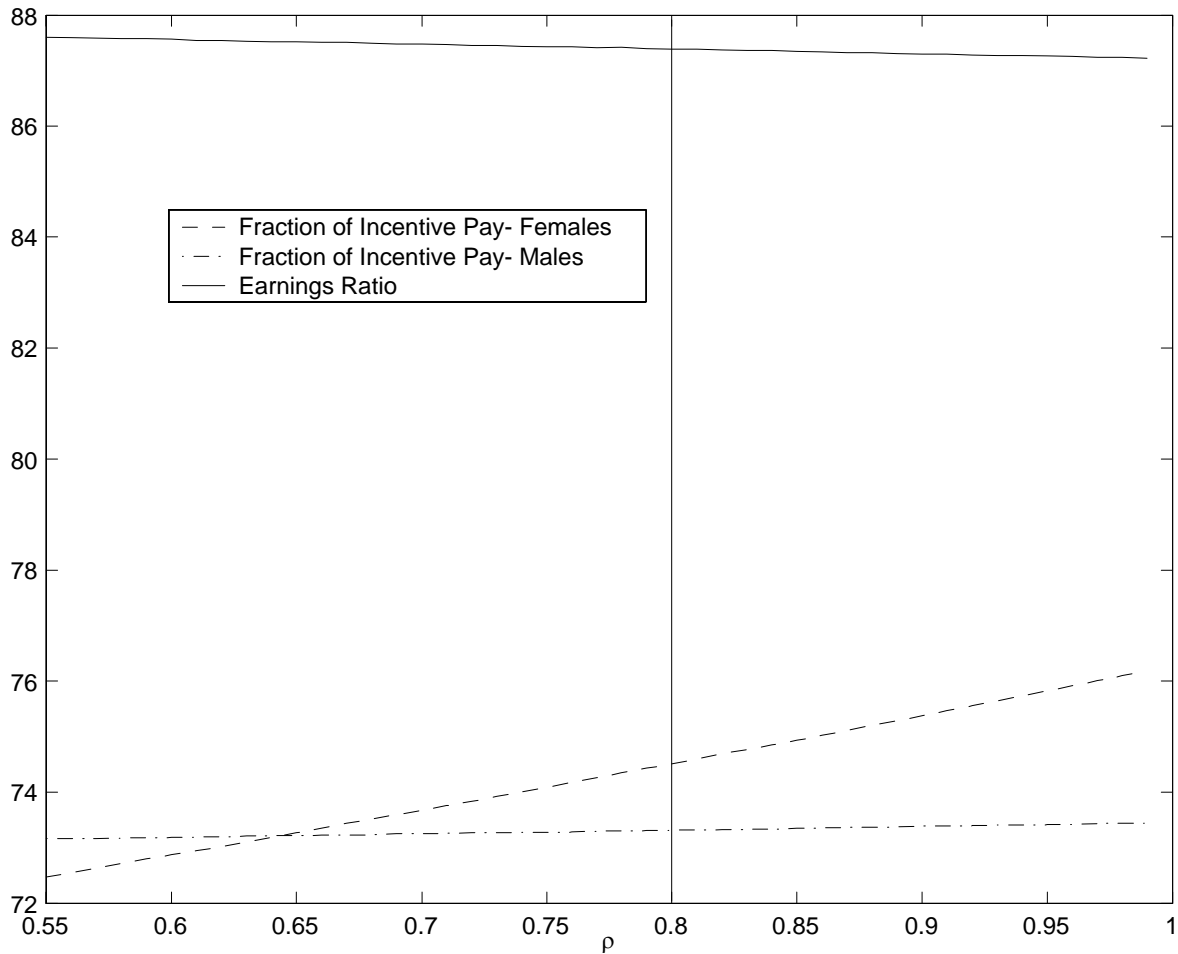


Figure 1:

more so for females. Higher values of ρ correspond to higher substitutability between effort and hours in utility. This exacerbates the incentive problem, more so for women who have a higher marginal cost of effort. For high enough values of ρ , the fraction of incentive pay is higher for women than men.

Figure 2 plots the fraction of incentive pay in earnings for females and males and the earnings ratio as a function of z . Lower values of z correspond to lower substitutability between effort and hours in production. This reduces the marginal product of effort for firms at lower values of hours. It follows that lower values of z increase the cost of implementing high effort more for female workers, since they work fewer hours. As long as high effort is being implemented for male and female workers, the fraction of incentive pay in earnings is very similar for both genders, though higher for women. For sufficiently low values of z , however, firms will find it optimal to implement low effort for female workers, while high effort is still implemented for males. This determines a large drop in the earnings ratio.

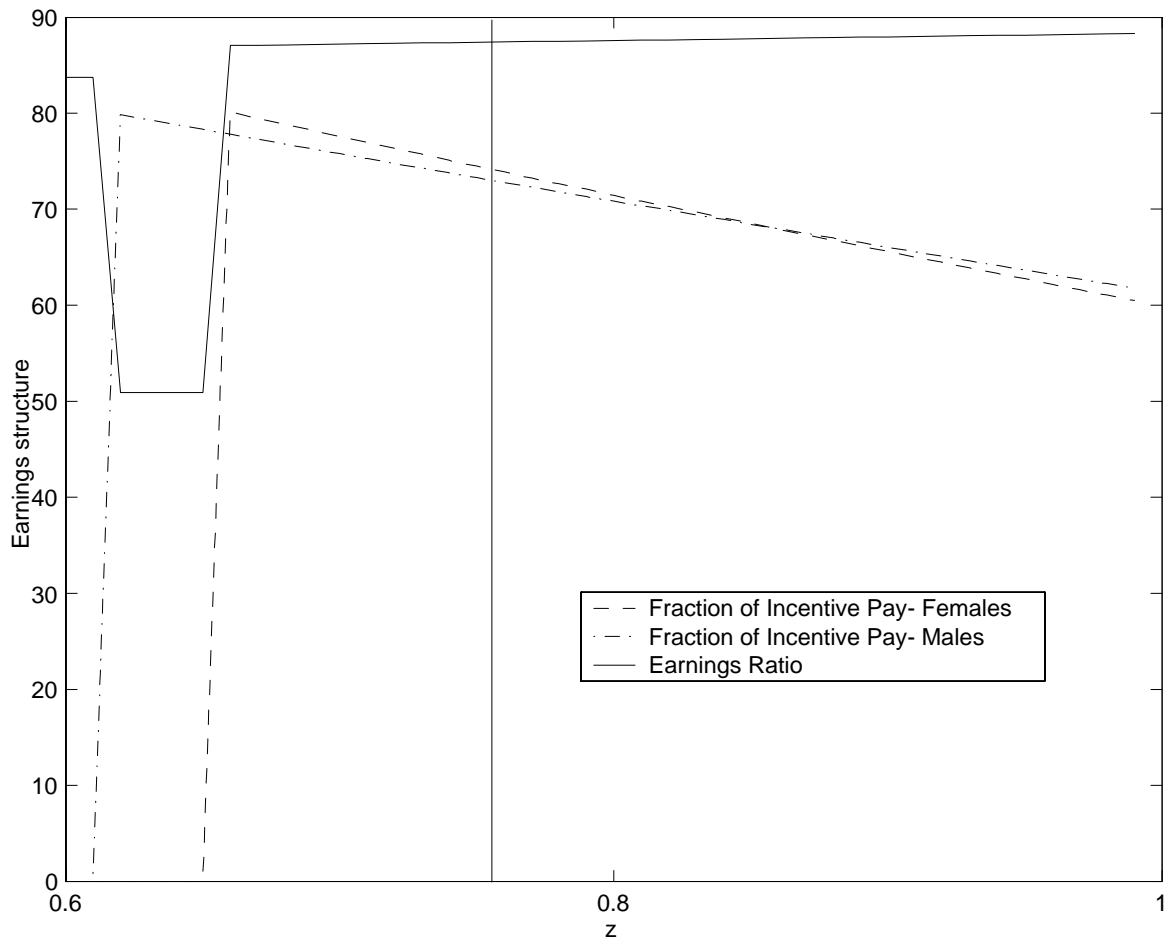


Figure 2:

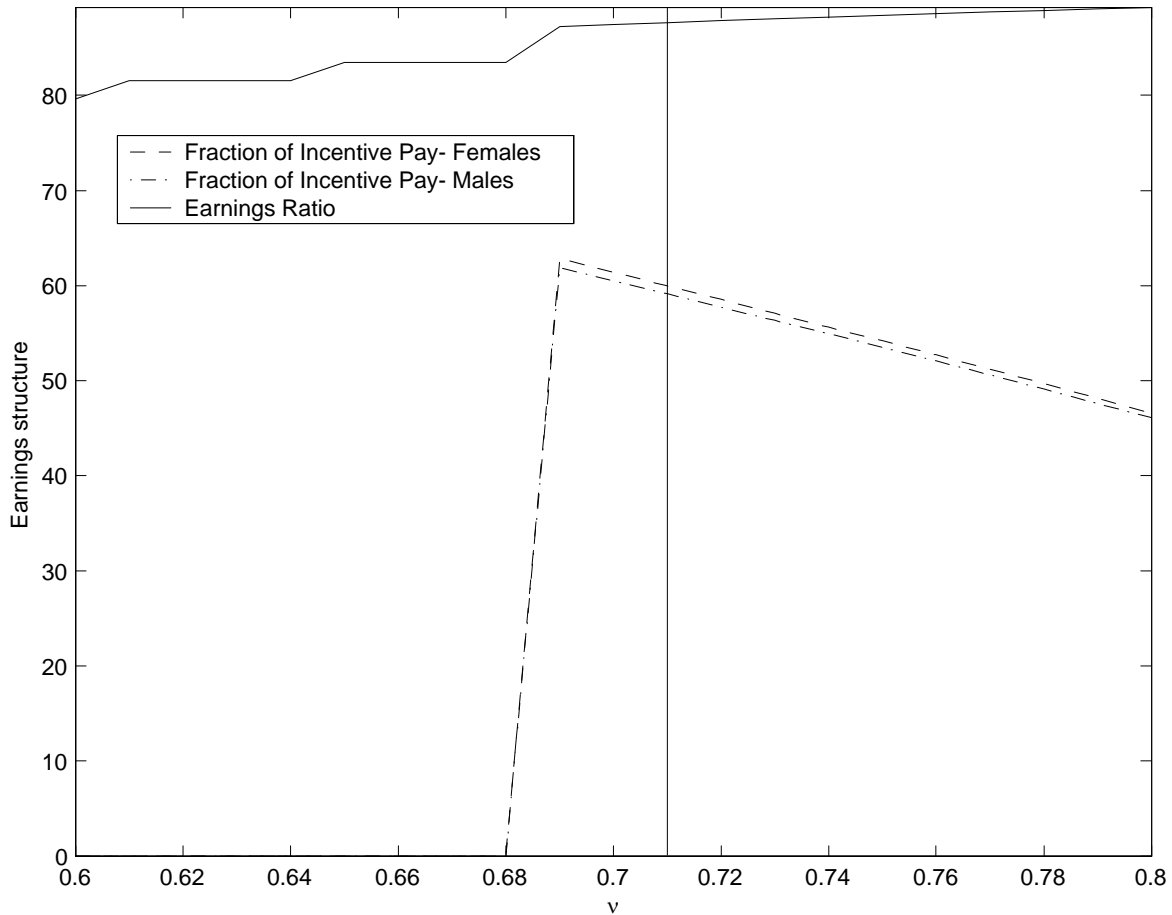


Figure 3:

Figure 3 plots the fraction of incentive pay in earnings for females and males and the earnings ratio as a function of ν , which determines the relative cost of effort for workers. As ν declines, the relative cost of effort increases, and this exacerbates the incentive problem. The interpretation is similar to the one for Figures 1 and 2.

Our findings for ν are consistent with historical evidence on the gender wage gap and on the structure of earnings for female and male workers in manufacturing and in the clerical sector in Goldin (1986, 1990). Goldin argues that clerical sector work was perceived as less onerous, involving less physical fatigue and discomfort than manufacturing work. She documents that in manufacturing, piece-rate compensation was more prevalent for women. Specifically, 47% of female operatives and 13% of males in the same positions are paid by the piece in manufacturing in 1890. This accords with equilibria in our model in which male and female workers both exert high effort. Goldin argues that in the clerical sector, “career tracks” emerged as a standard motivational device, replacing the piece-rate compensation schemes prevalent in manufacturing. She maintains that in the clerical sector: “Firms often used sex as a signal of shorter expected job tenure. ... By segregating workers by sex into job ladders (and some

dead-end positions), firms may have been better able to use the effort-inducing and ability-revealing mechanisms of the wage structure.” If we adopt a job ladder interpretation, labor contracts implementing low effort can be thought of as "dead end jobs", while labor contracts that implement high effort can be thought of as positions that allow for career growth. Then, our results on the effect of a higher relative cost of effort on the earnings structure are consistent with the differences in earnings structure by gender across these two sectors. Moreover, these findings are in line with the transition to the clerical sector in the period 1930/1940, when, as discussed in Goldin (1990), there was also a significant rise in the gender wage gap.

Finally, we explore the properties of the model as a function of the parameter Σ , which corresponds to the variance of output for given hours and effort. An increase in this parameter makes it harder to infer effort from observed output. Figure 4 displays the fraction of incentive pay in earnings for females and males and the earnings ratio for different values of Σ . We vary Σ from values approximately equal to 10% of output to values approximately equal to 60% of output for males. As Σ increases, the surplus from the employment relationship declines, given that workers are risk averse. This can be seen from the multiplier on the incentive compatibility constraint, derived in section 2.2, which is proportional to Σ . It follows that the returns from implementing high effort for firms decline. This effect is greater for women, who have a higher marginal cost of effort. Then, there will be a critical value of Σ , for which it becomes too costly for firms to implement high effort for females, but not for males. As discussed for Figure 1-3, as long as high effort is being implemented for both, the fraction of incentive pay in earnings is similar across genders, though higher for women and the earnings ratio roughly constant. The earnings ratio drops significantly, when low effort is implemented for females but not for males.

3 Connecting the Model with the Evidence

There are three distinct predictions of our framework that we need to document. First, the gender earnings differentials should be higher in industries/occupations in which the incentive problem is more severe. Moreover, this pattern should be stronger for married workers than for single workers. We use Census 2000 data to study gender earnings differentials by marital status across industries and across four broad occupational categories. We find that for married individuals gender differentials are greatest in high management and sales occupations across all industries. We argue that these occupations are likely to be the ones in which incentive problems are most stringent. For the sample of single workers we do not observe the same pattern of gender differentials across industries and occupations.

Second, according to our model, we should observe a negative correlation between the female/male ratio of home hours and the female/male earnings ratio across industries/occupations. This correlation should be stronger for married workers than for single workers. Since the Census does not include information on the hours devoted to home production activities, we use PSID data from the late 1990s to document this fact. We find a negative and significant correlation between the two ratios for married workers. The correlation is halved for single workers. Finally, our model also predicts a negative correlation between relative home hours of

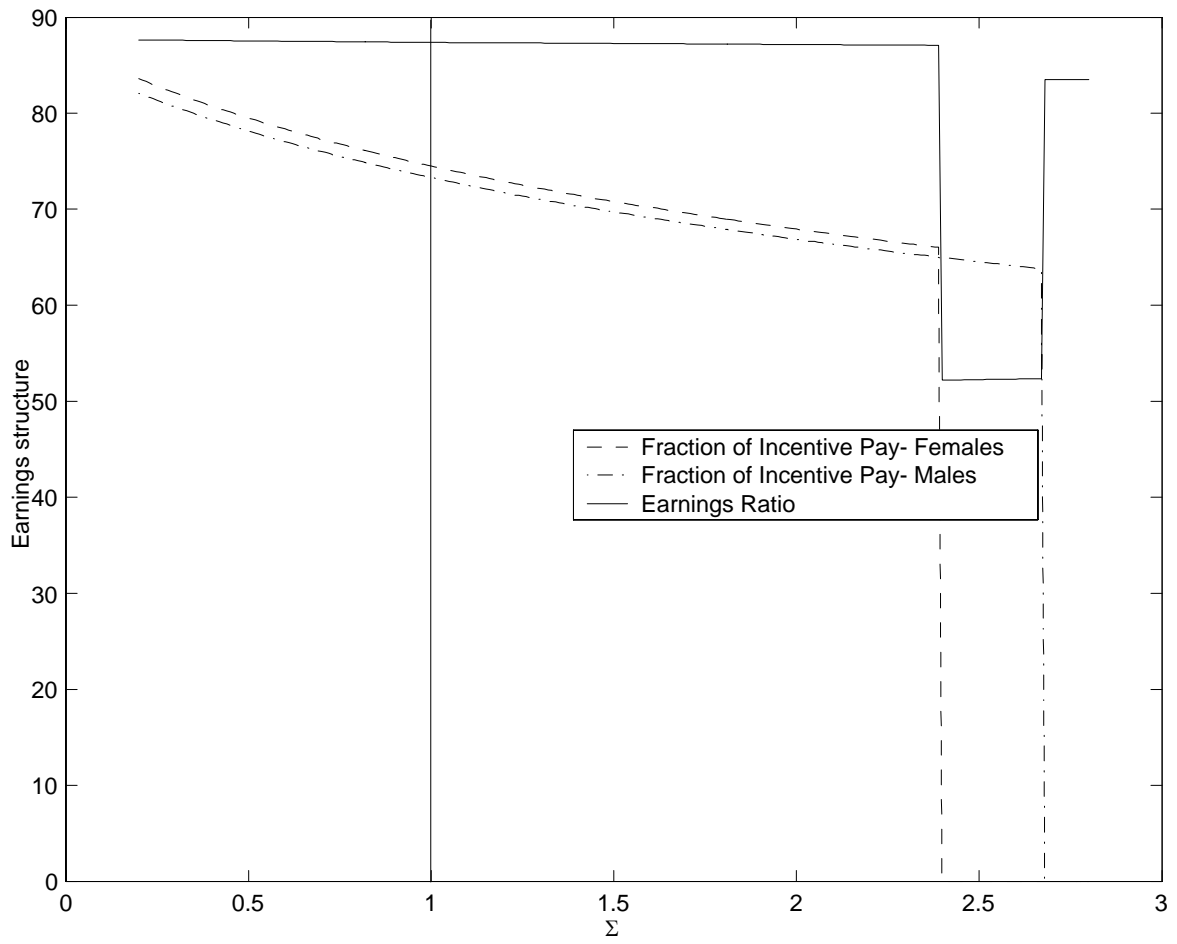


Figure 4:

husbands and *wives* and their relative earnings, as would be the case for other efficient models of intra-household allocation. We use the PSID to show that there is indeed a negative and significant correlation between the husband/wife ratio of home hours and the husband/wife ratio of total labor income in the data.

3.1 Evidence from the Census

In this section, we use Census data to document differences in gender earnings differentials across all industries for four broad occupational categories: higher level management, lower level management, sales and production.¹⁸ We consider this occupational classification based on the notion that incentive problems are more stringent in sales and high management relative to production and low management. We link the severity of the incentive problem to the degree of uncertainty over the workers' effort for given observable measures of performance, which correspond, respectively, to the parameter Σ and output in our model¹⁹. For management occupations, the uncertainty associated with managers' effort given observable performance measures should be related to the range of responsibilities held. Hence, we expect it to be greater for high level managers. For sales occupations, sales volumes are typically used as a benchmark measure of performance. Yet, these depend to a large degree on variables that are not directly related to a sales personnel's effort and hours of work and may be uncertain²⁰. These considerations are less important for production workers.

Following the Census classification we consider 19 industries: Agriculture; Forestry Fishing and Hunting, Mining, Utilities, Construction, Manufacturing, Wholesale Trade, Retail Trade, Transportation and Warehousing, Information, Finance and Insurance, Real Estate and Rental/Leasing, Professional, Scientific and Technical Services, Administrative and Support and Waste Management and Remediation Services, Educational Services, Health Care and Social Assistance, Arts, Entertainment and Recreation, Accommodation and Food Services, Other Services (except Public Administration), and Public Administration. Our sample includes all individuals between 15 and 64 years of age, who are not in school, do not reside on a farm or live in group quarters. We also exclude the armed forces and restrict attention to those individuals who work at least 50 weeks in the previous year and who usually work at least 30 hours per week.

For each industry and for each of the four broad occupational categories we compute the gender gap in earnings by running median regressions that control for a gender dummy as well

¹⁸We use Census 2000. The industry variable, INDNAICS, reports the type of establishment in which a person worked in terms of the good or service produced. Industries are coded according to the North American Industrial Classification System developed in 1997. We use the variable OCCSOC for occupation. OCCSOC classifies occupations according to the 1998 Standard Occupational Classification (SOC) system. The Census also provide an aggregation of all the occupations in 23 broader categories that include the four categories considered in the analysis. The definition of production occupations also includes construction and extraction workers.

¹⁹See sections 2.2 and 2.4.

²⁰For example, sales workers are typically assigned to specific territories or products. Hence, sales volumes will fluctuate with shocks to local demand. See Catalyst (1995) for a description of the sales occupation, especially in relation to gender.

as for human capital variables - age and its square term and education. Our measure of the median gender gap in earnings is the coefficient of the gender dummy in the regression. For the education variable, we group individuals according to four broad educational categories: less than high school, high school completed, some college and college completed. We construct four education dummies based on this categorization. The first dummy is equal to one if an individual has completed less than twelve years of schooling and is equal to zero otherwise. The second dummy variable is equal to one if he or she has completed twelve years of schooling, and is equal to zero otherwise. The third dummy variable equals one if the individual has completed between twelve and fifteen years of schooling and it is equal to zero otherwise. Finally, the fourth dummy variable is equal to one if an individual has completed at least sixteen years of education and it equals zero otherwise. The omitted dummy variable corresponds to individuals who completed less than twelve years of schooling. The dependent variable is the log of annual earnings. We use total labor earnings in our analysis because this is the data counterpart of the measure of total labor compensation that we consider in our model. However, one could argue that this is not the appropriate measure of labor compensation when making gender comparisons, since women tend to work less hours on the labor market than men do. Hence, our analysis could be confounding gender differences in market hours and in hourly compensation. This concern is attenuated by the fact that we only consider individuals that usually work at least 30 hours per week and who were employed for at least 50 weeks in the previous year. This sample selection criterion reduces considerably the variation in the number of market hours within and between gender groups. Further, we have also performed our analysis using the log of hourly wages as a dependent variable. Our findings in this case are consistent with the ones reported in this section.

Table 5 reports the results of this analysis. The first column in the table reports the female/male ratio of median earnings for full-time year-round workers across all industries for the four broad occupational categories. The second and third column report, respectively, the frequency with which the gender wage ratio is lowest (second column) and highest (third column) across the four categories. In each column we report the results obtained both for the sample of married individuals and for the sample of individuals who were never married.

We find that for married workers there is a systematic pattern in the ranking of gender earnings gaps over the four occupational categories across all the industries. In particular, we find that the female/male median earnings ratio is lowest for sales occupations in 13 industries - 68% of the cases. In the 6 remaining industries (32% of the cases) the occupation that display the lowest ratio of female to male median earnings is high management. On the other end, when we look at the frequency with which the gender wage ratio is highest across all industries we find the opposite pattern. Workers in production occupations display the largest ratio in 13 industries - 68% of the cases. In the 6 remaining industries the ratio is highest in low level management occupations. As shown in column 1, for the sample of married workers the average female/male ratio of median earnings in all industries varies substantially across the four occupation categories, from a minimum of 0.57 in sales to a maximum of 0.79 in production.

Table 5: Gender differences in earnings across industries for four broad occupational categories. (Full time, year round workers.)

	% female/male earnings ratio		% times lowest ratio		% times highest ratio	
	<i>married</i>	<i>single</i>	<i>married</i>	<i>single</i>	<i>married</i>	<i>single</i>
higher level mgmt	65	94	32	5	0	53
lower level mgmt	74	86	0	16	32	16
sales	57	86	68	21	0	16
production	79	83	0	47	68	5

For the sample of never married workers the gender gap is much lower and varies less across industries and occupations. Single women earn at least 83% of single men’s median earnings across all industries. In contrast to our findings for married workers, we find that the gender earnings ratio for single workers is largest in high management occupations in 10 of the 17 industries - 53% of the cases. For the 7 remaining industries we find that the earning ratio is highest in sales (3 times), low management (3 times) and production (1 time). The ratio is lowest in production occupations for 9 industries - 47% of the cases. For the 8 remaining industries we find that the earning ratio is lowest in sales (4 times), low management (3 times) and high management (1 time). Note that for this sample we can only consider 17 industries since in two industries, Agriculture and Forestry and Mining, there are two few observations in sales occupations.

Our findings suggest that there is a considerable variation in the female/male ratio of median earnings across industries, and across the four occupational categories within an industry, even after controlling for human capital variables. We report these ratios by industry, occupation and marital status in Table 6. For each industry, we include the value for the occupation with the lowest and highest ratio, as well as for all occupations in the Census. The variation of the median earnings ratio across industries and occupational categories is striking. For married workers it ranges from 0.20 for sales occupation in the Agriculture and Forestry industry to 1.012 for low management occupation in Mining. For never married workers the range is smaller, from a minimum of .55 for sales occupation in the Utilities industry to a maximum of 1.23 for sales occupation in the Public Administration. This variation could be exploited to calibrate industry specific values of the production technology parameters in our model to the gender earnings ratio in different industries and analyze the resulting structure of earnings by occupation. We plan to pursue this in further research, as well as to extend the analysis to a more disaggregate classification by industry.

Table 6: Variation in female/male median earnings ratios by industries, occupation, and marital status

	Married			Never Married		
	Lowest	Highest	All	Lowest	Highest	All
Accomodation and Food	0.38 (sales)	0.78 (prod.)	0.61	0.82 (slaes)	1 (high mgt)	0.86
Administrative and Support	0.64 (sales)	0.79 (prod.)	0.70	0.85 (prod.)	0.91 (sales)	0.90
Agriculture, forestry	0.20 (sales)	0.98 (low mgt.)	0.75			
Arts, Ent. & Recreation	0.63 (sales)	0.82 (prod.)	0.73	0.81 (high mgt)	1.12 (low mgt)	0.92
Construction	0.51 (high mgt.)	0.82 (prod.)	0.59	0.67 (low mgt)	0.94 (high mgt)	0.89
Educational Services	0.60 (sales)	0.83 (low mgmt.)	0.72	0.88 (prod.)	0.95 (high mgt)	0.91
Finance and Insurance	0.55 (sales)	0.82 (prod.)	0.54	0.83 (sales)	0.92 (high mgt)	0.85
Health Care & Soc. Assistance	0.54 (sales)	0.74 (prod.)	0.65	0.62 (sales)	0.97 (high mgt)	0.91
Information	0.68 (high mgt.)	0.82 (prod.)	0.64	0.87 (prod.)	0.91 (high mgt)	0.90
Manufacturing	0.60 (sales)	0.74 (low mgt.)	0.59	0.61 (prod.)	0.95 (high mgt)	0.83
Mining	0.56 (high mgmt)	1.01 (low mgt.)	0.59			
Other Services (no Public. Adm.)	0.44 (sales)	0.73 (low mgt.)	0.57	0.72 (prod.)	1.02 (low mgt)	0.85
Profess,Scientific&Tech. Services	0.61 (sales)	0.81 (prod.)	0.61	0.86 (prod.)	0.90 (high mgt)	0.86
Public Administration	0.64 (sales)	0.82 (prod.)	0.69	0.88 (prod.)	1.23 (sales)	0.87
Real Estate and Rental/Leasing	0.58 (sales)	0.83 (prod.)	0.70	0.88 (prod.)	1.02 (high mgt)	0.95
Retail Trade	0.44 (sales)	0.68 (prod.)	0.51	0.71 (prod.)	0.91 (high mgt)	0.82
Transportation and Warehousing	0.55 (sales)	0.82 (prod.)	0.72	0.85 (low mgt)	1 (high mgt)	0.91
Utilities	0.64 (high mgt.)	0.83 (prod.)	0.64	0.55 (sales)	0.88 (prod.)	0.80
Wholesale Trade	0.61 (high mgt.)	0.85 (low mgt.)	0.62	0.70 (low mgt)	0.97 (high mgt)	0.88

3.2 Evidence from the PSID

There are two additional predictions that we need to document. First, according to our model, we should observe a negative correlation between the female/male ratio of home hours and the female/male earnings ratio across industries/occupations. Second, our model also predicts a negative correlation between relative home hours of *husbands* and *wives* and their relative earnings, as would be the case for other efficient models of intra-household allocation. Since the Census does not include information on home hours we use PSID data for the late 1990s to document this fact. As we did with the Census data set, we select our sample to include all individuals between 15 and 64 years of age who are not in school and who are not in the armed forces. In our analysis we study gender earnings ratios at the occupation/industry level and by marital status. This level of disaggregation requires a larger sample size than the one available in each wave of the PSID. Hence, we do not exploit the panel dimension of the data set but simply pool together all the individuals in the 1994 to 2001 waves of the PSID.²¹ The resulting statistics can be interpreted as medium run averages of the relevant variables.

The two main variables in our analysis are weekly home hours and annual labor earnings. The earnings variable is top coded in the PSID. We follow the procedure in Katz and Murphy (1992) and impute annual earnings at 1.45 times the annual topcode amount for individuals with top coded labor earnings. We also exclude workers with real weekly earnings below \$67 in 1982 dollars from the sample. Our sample includes 49301 observations of which 33,379 are married and only 7,728 are single. We use PSID weights to compute the relevant statistics.

3.2.1 Earnings gaps and home hours across occupations

In this section we present evidence on the correlation between the female/male ratio of median earnings and the female/male ratio of average home hours across occupations/industries. The PSID coding of occupations differs from the one available from the Census 2000. In our analysis, we construct occupation/industry categories that are as close as possible to the ones used in our Census analysis. In order to obtain a comparable classification we combine information on occupations and industries from the PSID. In particular, we consider the following categories: management positions in administration, management positions in banking, finance and in the clerical sector, lower level management occupations, professional occupations (engineers, architects, lawyers, and medical doctors), technical occupations (in the health sector, engineering, and social sciences), occupations in community/social services, social scientists and university professors, teachers other than college professors, occupations in arts and entertainment, design, sports and the media, sales occupations, clerical occupations, craftsmen, operatives, physical laborers, laborers in services excluding private households, and laborers in private households.

The results of this analysis are presented in Figure 5. We find that lower gender earnings ratios are associated with higher female/male ratios of home hours for workers in management occupations in banking and finance and in the clerical sector, and for workers in sales occupations. This finding is consistent with the evidence from the Census. However, we find that for

²¹Since the PSID started collecting data bi-annually since 1997 our sample includes 7 waves.

workers in operative occupations and for laborers higher gender earnings ratios are associated with higher female/male ratios of home hours. We think that wealth effects might explain this observation. The correlation between the two ratios is twice as large for married workers than for never married workers. For married workers, the two ratios display a negative correlation of -0.31. The correlation coefficient is significant at the 1 per cent level. For single workers the correlations is equal to -0.15. We interpret these statistics as suggestive evidence in favor of the mechanism that we highlight in our model.

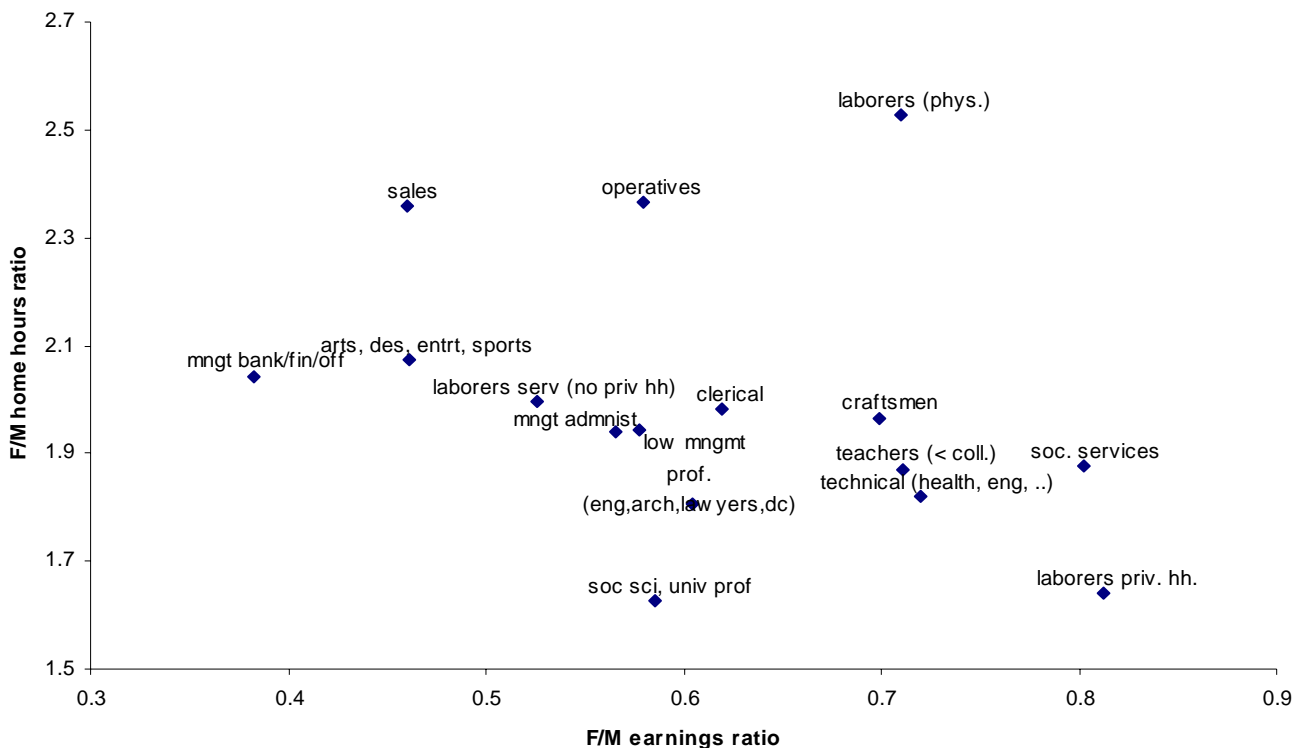


Figure 5: Correlation between F/M earnings ratio and F/M weekly home hours

The use of PSID data for this analysis has two drawbacks. First, even when we pool together all the the waves of the PSID we do not have enough observations to perform the analysis at the same level of disaggregation as in the Census analysis. Second, the information on home hours in the PSID is not obtained from time diaries, rather the survey respondent is asked to provide a measure of weekly hours worked by him- or herself and by his or her spouse (if married.) As a consequence, the home hours variable is potentially subject to serious measurement problems. In order to address both concerns, we plan to use the American Time-Use data set (ATUS) recently made available by the Bureau of Labor Statistics for the year 2003. This unique data set combines very detailed time-use information at the individual level with the information on labor market variables typically available in the Current Population Survey. This data set also

has the advantage to provide a much larger number of observations than the PSID. However, the ATUS data set does not include time-use information for both husbands and wives in the CPS sample (only one of the two partners is asked to provide time-use information). Hence we will not be able to use it to analyze patterns of relative home hours and earnings across married couples.

3.2.2 Relative wages and home hours for married couples

Our model predicts that husband-wife ratio of earnings should be negatively correlated with their home hours ratio, that is, the spouse with higher earnings will contribute fewer hours to the production of the home public good. This prediction is supported by the PSID evidence. We find that there is a negative and significant (at the 1 per cent level) correlation between the husband/wife ratio of home hours and the husband/wife ratio of total labor income. The correlation coefficient ranges from 0.12 for the whole sample to approximately 0.25 for the sub-sample of couples where the husband’s labor income is in the upper half of the male earnings distribution. This is also true when we concentrate on couples where both the husband and the wife work full time.

Table 7 reports on the fraction of married couples that display a negative relationship between earnings and home hours of the spouses. We find that 86.3% of married couples in the PSID sample display a correlation which is consistent with our model’s prediction.²² We also investigate how this fraction varies by the husband’s position in the male earnings distribution. The fraction of married couples for which we observe a negative correlation between these two ratios ranges from 73% in the bottom quartile to approximately 95% in the top quartile.²³

Table 7: Composition of the PSID sample

	All	Quartile of Male Earnings Distribution			
		<25	25-50	50-75	>75
couples with negative $\text{corr}\left(\frac{w_m}{w_f}, \frac{h_m}{h_f}\right)$	86.3%	73%	85.3%	92%	95.3%

4 Conclusion

We plan to extend our analysis along several dimensions. We have shown that in equilibria with gender discrimination, given that firms use gender as a screening device, discrimination actually reduces the incentive problem for firms, eliminating adverse selection. We conjecture that this property of the model implies that equilibria without discrimination are unstable. We

²²In 3.7% of the remaining households the husband is both the main earner and the main provider of home hours. This leaves a 10% of households where wives have both higher earnings and higher home hours than their husbands.

²³Findings are similar if we consider the sample of white married couples.

plan to formally analyze this conjecture in future drafts of the paper. Moreover, in our current framework, there are no costs associated with gender discrimination. We plan to extend the analysis to environments where discrimination is potentially associated with efficiency losses stemming from the misallocation of workers.

Our model also predicts that gender earnings gaps should be higher in industries/occupations in which the incentive problem is more severe. Using Census data, we have shown that gender earnings differentials are greatest in high management and sales occupations across all industries. We argue that these occupations are likely to be the ones in which incentive problems are most stringent. We plan to extend this analysis to a more disaggregate level.

References

- [1] Joseph G. Altonji, and Rebecca M. Blank, 1999, "Race and Gender in the Labor Market", Handbook of Labor Economics, Volume 3. Elsevier.
- [2] Jess Benhabib, Richard Rogerson, and Randall Wright, 1991. Homework in Macroeconomics: Household Production and Aggregate Fluctuations. *The Journal of Political Economy*, Vol. 99, No. 6, 1166-1187.
- [3] Linda Babcock and Sarah Laschever, 2003. Women Don't Ask: Negotiation and the Gender Divide. Princeton University Press.
- [4] Gary S. Becker, 1985, "Human Capital, Effort and the Sexual Division of Labor", *Journal of Labor Economics*, Vol. 3, No.1, 33-58.
- [5] Gary S. Becker, and Kevin Murphy, 2000, "Social Economics: Market Behavior in a Social Environment", Harvard University Press.
- [6] Theodore Bergstrom, 1997, "A Survey of Theory of the Family" Handbook of Family and Population Economics", Elsevier.
- [7] Catalyst, 1995, "Knowing the Territory: Women in Sales. A Catalyst Performance Report", Catalyst Publications, New York.
- [8] Pierre-Andre Chiappori, 1988. Introducing Household Production in Collective Models of Labor Supply. *Econometrica*, Vol. 56, No. 1, 63-90.
- [9] Pierre-Andre Chiappori, 1997. Introducing Household Production in Collective Models of Labor Supply. *The Journal of Political Economy*, Vol. 105, No. 1, 191-209.
- [10] Stephen Coate, and Glen Loury, 1993, "Will Affirmative Action Policies Eliminate Negative Stereotypes?" *The American Economic Review*, Vol. 83, No.5, 1220-1240.
- [11] Drogin, Richard, 2003, "Statistical Analysis of Gender Patterns in Wal-Mart Workforce". Plaintiff's Expert Statistical Report, *Dukes vs Wal-Mart*, <http://www.walmartclass.com>.

- [12] Patrick Francois, 1998, "Gender Discrimination without Gender Difference: Theory and Policy Responses", *Journal of Public Economics*, Vol. 68, No.1, 1-32.
- [13] Claudia Goldin, 1986, "Monitoring Costs and Occupational Segregation by Sex: A Historical Analysis," *Journal of Labor Economics*, Vol. 4, 1-7.
- [14] Claudia Goldin, 1990. *Understanding the Gender Wage Gap: An Economic History of American Women*. Oxford University Press.
- [15] Bengt Holmstrom and Paul Milgrom, 1991. Multitask Principal Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics and Organization*, Volume 7.
- [16] Jean-Jack Laffont, and David Martimort, 2002, "The Theory of Incentives. The Principal Agent Model", Princeton University Press.
- [17] Light, Audrey and Manuelita Ureta, 1992, "Panel Estimates of Male and Female Job Turnover Behavior: Can Female Non-Quitters Be Identified?," *The Journal of Labor Economics*, Vol. 10, No. 2, 156-181
- [18] Shelly Lundberg and Richard Starz, 1983, "Private Discrimination and Social Intervention in Competitive labor Market", *The American Economic Review*, Vol. 73, No. 3, 340-347.
- [19] W. Bentley MacLeod, and James M. Malcomson, 1988. Reputation and Hierarchy in Dynamic Models of Employment. *The Journal of Political Economy*, Vol. 96, No. 4, 832-854.
- [20] Marjorie McElroy, and Mary Jean Horney, 1981, "Nash Bargained Household Decisions: Toward a Generalization of the Theory of Demand", *International Economic Review*, 22, 333-349.
- [21] Marilyn Manser and Murrey Brown, 1980, "Marriage and Household Decision Making: A Bargaining Analysis", *International Economic Review*, 21, 31-44.
- [22] June O'Neill, 2003. The Gender Wage Gap in Wages, circa 2000. *American Economic Review*, Papers and Proceedings, June.

5 Appendix A

Proof of Proposition 1

Conditions (10) and (11) imply that Γ , the fraction of earnings given by incentive pay satisfies:

$$\Gamma := \frac{\tilde{w} f(n, 1)}{\bar{w} + \tilde{w} f(n, 1)} = \tilde{w}.$$

We are interested in:

$$\frac{d\tilde{w}}{dh} = \frac{\partial \tilde{w}}{\partial h} + \frac{\partial \tilde{w}}{\partial n} \frac{dn}{dh}.$$

Condition (10) implies:

$$\frac{\partial \tilde{w}}{\partial h} = \tilde{w} \frac{v_h(h, n, 1) - v_h(h, n, 0)}{v(h, n, 1) - v(h, n, 0)} < 0,$$

by (2). In addition:

$$\frac{\partial \tilde{w}}{\partial n} = \tilde{w} \left[\frac{v_n(h, n, 1) - v_n(h, n, 0)}{v(h, n, 1) - v(h, n, 0)} - \frac{f_n(n, 1) - f_n(n, 0)}{f(n, 1) - f(n, 0)} \right].$$

Hence, if $v_{ne} \leq 0$ and $f_{ne} \geq 0$, $\frac{\partial \tilde{w}}{\partial n} \leq 0$; if $v_{ne} > 0$ and $f_{ne} < 0$, $\frac{\partial \tilde{w}}{\partial n} > 0$. Then, if $dn/dh \leq 0$, $v_{ne} \leq 0$ and $f_{ne} \geq 0$ guarantees $\frac{d\tilde{w}}{dh} \geq 0$. Incentive pay depends on h as follows:

$$\frac{d[\tilde{w}f(n, 1)]}{dh} = \frac{d\tilde{w}}{dh} f(n, 1) + \tilde{w} f_n(n, 1) dn/dh.$$

The result follows. ■