Estimating a Dynamic Adverse-Selection Model: Labor-Force Experience and the Changing Gender Earnings Gap 1968–97.*

George-Levi Gayle and Limor Golan

Tepper School of Business, Carnegie Mellon University.

June 2010.

Abstract

This paper addresses two questions: What accounts for the gender gap in labor-market outcomes? What are the driving forces behind the changes in the gender-labor-market outcomes over the period 1968–97? It formulates a dynamic general equilibrium model of labor supply, occupational sorting and human capital accumulation in which gender discrimination and an earnings gap arise endogenously. It uses this model to quantify the driving forces behind the decline in the gender earnings gap and the increase in women's labor-force participation, professional-occupation representation and hours worked. It finds that labor-market experience is the most important factor explaining the gender earnings gap. In addition, statistical discrimination accounts for a large fraction of the observed gender earnings gap and its decline. It also finds that a large increase in aggregate productivity in professional occupations plays a major role in the increase in women's labor-force participation, professional-occupation representation and hours worked. Although of less importance, demographic changes account for a substantial part of the increase in female labor-force participation and hours worked, whereas home-production technology shocks do not.

Keywords: Gender earning gap, Discrimination, Occupation sorting, Labor-market experience, Structural estimation of dynamic games, Dynamic general equilibrium, Adverse selection, Ratchet effect.

1 Introduction

One of the most striking changes in the U.S. labor market over the last four decades is the significant decline in the gender wage gap in the 1970s and 1980s. The median gender wage differential

*We thank Orazio Attanasio, Lanier Benkard, Dan Bernhardt, Moshe Buchinsky, Zvi Eckstein, Larry Jones, John Kennan, Kevin Lang, Robert A. Miller, Derek Neal, Andrew Postlewaite, Victor Rios-Rull, Robert Shimer, Steve Spear, Yoram Weiss, and the participants of SOLE 2005, Tel-Aviv Summer Workshop 2005, SITE Theory-Based Micro-Econometric Modelling Workshop 2006, Econometric Society 2007, NBER Summer Workshop 2007, SITE Family Behavior and the Aggregate Economy Workshop 2007, CRES Empirical Microeconomics Conference at Washington University, and seminars at Purdue University, University of Minnesota, Carnegie Mellon, University of Pennsylvania, New York University, Boston University, Northwestern University, the Federal Reserve Bank of Chicago, IZA (Bonn), the Institute of Fiscal Studies (London), and University of Rochester for comments and suggestions.

dropped from about 40% in 1968 to 28% by 1992 and has remained constant since. Accompanying this decline were significant increases in women's labor-force participation, hours worked and professional-occupations representation. An important feature of the gender wage gap is its evolution over the life cycle as a function of labor-market experience. The gender wage gap is relatively small when workers are young and it increases with workers' age. In contrast, the gender wage gap for workers who work continuously full time decreases with age. The bottom right panel of Figure 3 illustrates these facts. This paper addresses two questions: What accounts for the labor-market gender gaps and their evolution over the life cycle? What are the main forces behind their changes over time?

Theoretical models of discrimination have a long tradition in economics going back to the seminal work by Becker (1971), Arrow (1972) and Phelps (1972). This literature can be divided in two categories, taste-based and statistical discrimination theory. Taste-based discrimination theory postulates that employers have a preference for one group. In the absence of impediments to competition, this theory predicts that group-based earnings gaps cannot be sustained in equilibrium. Statistical discrimination models emphasize that group differences can arise endogenously without any ex-ante differences across groups or preference for one group (see Coate and Loury, 1993, for this insight). This literature focuses primarily on productivity differences across racial groups (see Moro, 2003; Antonovics, 2004; Altonji, 2005). Two exceptions are Baron, Black, and Lowenstein (1993) and Albanesi and Olivetti (2009). The latter develops a static model of statistical gender discrimination in which effort in the labor market and hours worked at home are determined endogenously. The former is the first to formalize a statistical discrimination model based on the difference in the expected turnover rate between women and men. It shows that if employers expect women to have a higher turnover rate than men, then they offer women jobs which require less training and also pay less. Our paper, therefore, contributes to this literature by incorporating statistical discrimination, as formulated in Baron, Black, and Lowenstein, into a general equilibrium model with life-cycle labor-supply choices; this allows us to capture the evolution of the labor-market-outcomes gap over the life cycle.

There is an extensive empirical literature on the gender wage gap (see Altonji and Blank, 1999, for a survey). Most of this work describes the pattern of the gender wage gap and decomposes it into: the gap explained by differences in observed characteristics and the residual, typically attributed to discrimination. Although informed by theory, there is no fully specified behavioral model explaining the existence of a gender wage gap and its evolution over the life cycle in this literature. The exceptions are Bowlus (1997); Erosa, Fuster, and Restuccia (2005); and Flabbi (2010, forthcoming). Bowlus (1997) presents a static search model with no discrimination or human capital accumulation. It thus attributes the entire gender wage gap to differences in unobserved productivity between men and women. Flabbi (2010, forthcoming) adds taste-based discrimination to the Bowlus (1997) model. Bowlus (1997) and Flabbi (2010, forthcoming) do not account for the observed patterns of the gender earnings gap over the life cycle. Erosa, Fuster, and Restuccia (2005) develops and calibrates a partial equilibrium life-cycle model of labor-supply and fertility choice, explaining the increase in the gender wage gap over the life cycle with no discrimination; it does not explain, however, the decline in the wage gap for workers who work continuously full time.

There is a large and growing literature explaining the decline of the labor-market-outcomes gender gap. This literature can be divided into two main groups: papers that examine the increase in labor-force participation and hours of married women (see Jones, Manuelli, and McGrattan, 2003;[1] Greenwood, Seshardri, and Yorukoglu, 2005; Attanasio, Low, and Sanchez-Marcos, 2008; Albanesi and Olivetti, 2008; Fernández and Fogli, 2009) and papers that examine the decline in the gender wage gap and the increase in women's labor-force participation and hours. The work closest to ours is Lee and Wolpin (2010); it develops and estimates a dynamic general equilibrium model with separate demand and supply factors affecting the changes in the gender wage gap. This paper, however, does not account for the role of discrimination in the decline of the labor-market outcomes gender gap.

Our model includes three channels through which a gender earnings gap can arise: group differences in preferences, group differences in productivity, and discrimination. We use the framework to assess the importance of these channels in the observed gender gap in labor-market outcomes and its decline over time. We demonstrate identification and develop a three-step estimator of the model. To the best of our knowledge, this is the first paper to estimate a dynamic signaling model.

The model consists of forward-looking workers and firms. Each period workers make labor-market–participation, occupation and hours-worked decisions. The supply side of the model extends Mincer and Polachek's (1974) and Polachek's (1981) labor-supply model by incorporating privately observed labor-market–participation costs. The demand side consists of competitive firms that incur costs of hiring new workers. As a result of the workers' private information and the employers' hiring costs, the model gives rise to an adverse selection problem. Low–participation-cost workers have longer employment spells, which makes them more profitable to the employer. Therefore, employers use the observed labor-supply decisions as a signal of the worker's private information. This in turn provides incentives for workers with high-participation-costs to mimic the labor-supply behavior of workers with lower-participation-costs.

The model gives rise to statistical discrimination. Suppose women, on average, have higher participation costs. Then employers will expect them to have higher turnover rate than men, and will pay them lower wages. We measure discrimination as the difference between the labor-market outcomes of men and women under symmetric and private information. Thus, gender differences in earnings that arise due to observed group affiliations are referred to as discrimination, as opposed to gaps that arise due to differences in preferences and productive skills.

Several conclusions arise from the empirical analysis. First, private information plays an important role in women's labor-market outcomes. As part of our empirical strategy, we test for the presence of private information and find evidence for it. Second, human capital accumulated with labor-market experience is the most important factor explaining the gender earnings gap.

[1] Jones, Manuelli and McGrattan (2003) allow for an exogenous gender wage wedge that changes over time.

We conduct counterfactual experiments to assess the effect of different factors on the gender earnings gap, labor-market experience, and occupational sorting. We find that hiring costs amplify the gender differences in preferences and skills and is the largest factor affecting the gender earnings gap. Further, discrimination plays a big role in the observed gender earnings gap and its decline over time. This finding is in contrast to the results in Flabbi (2010) which finds that taste-based discrimination cannot account for the decline in the gender wage gap in the 80s and 90s. In addition, we find that signaling is an important factor in labor-force participation: Women participate less when information is symmetric than they do when it is private. Thus, in contrast to empirical work on private information in the insurance market, we find that private information is both statistically and quantitatively significant in the labor market.

The literature focuses on several factors that caused the changes in the labor market outcomes of women relative to men between the 70s and the 90s. The first factor is skill-biased technological changes that increase productivity in skill-intensive occupations. The second factor is a decline in the cost of producing home goods. The third factor is changes in education level, marriage and fertility over time. Our results confirm the importance of technology-biased productivity and demographic changes (see Lee and Wolpin, 2010 for similar results). A comparatively larger aggregate productivity increase in professional occupations accounts for much of women's increased representation in these occupations, labor-force participation and hours worked. Demographic changes reduced women's costs of participating in the labor market, increasing their participation and hours worked. In contrast to Greenwood, Seshardri and Yorukoglu (2005), our results do not support the hypothesis that changes in home production costs explain the increase in women's labor market experience and hours worked over the period we examine, but the time period in their paper does not completely overlaps with ours. This result, however, is similar to the findings in Jones, Manuelli and McGrattan (2003). Because the papers in this literature do not incorporate discrimination, our main contribution to this literature is the finding that changes in discrimination accounted for about 50% and 40% of the change in the increase in women's labor market experience in professional and nonprofessional occupations, respectively.

This paper is organized as follows. Section 2 describes the model. Section 3 presents the equilibrium analysis and the equilibrium labor-market gender gaps. The empirical implementation is presented in section 4. Section 5 shows that the model is identified and develops the estimator. Section 6 contains empirical results and counterfactual simulations. Section 7 concludes. The appendices present proofs, implementation details, asymptotic properties of our estimator, and a detailed data description.

2 Theoretical Model

The model we present extends Altug and Miller's (1998) general equilibrium model of dynamic labor supply and consumption. We incorporate occupation sorting, hiring costs and private information into their framework. The hiring costs and private information give rise to endogenous gender discrimination as formalized by Baron, Black, and Lowenstein (1993) in a partial equilibrium model.

The economy consists of infinitely lived firms and finitely lived workers. There exists a continuum of workers on the unit interval $[0,1]$ in each age–education cohort. These workers are divided into two observed gender groups, $i \in \{\mathrm{w}, \mathrm{m}\}$, women and men, respectively. For notational ease, we denote age and calendar year for each cohort by t ($t = 0, \ldots, T$). The theoretical model is written and solved for a given cohort, but the economy consists of a number of overlapping cohorts. There is free entry into the competitive labor markets. There are two occupations, $\tau \in \{\mathrm{P}, \mathrm{NP}\}$, professional and nonprofessional, each consists of a continuum of identical firms. Each firm offers one job in each period. The job offer maximizes the employer's lifetime expected discounted profits.

2.1 Workers' Problem

Choice Set At each discrete time t, an individual of gender i makes labor-market participation and occupation choices, deciding how many hours to work and how much to consume. Labor-market participation and occupation choices are discrete, while hours and consumption choices are continuous. We denote the participation decision by d_t, where $d_t = 1$ if the individual participates in the labor force in period t, and 0 otherwise. The occupation and participation indicators are defined as $I_{\tau t}(\tau \in \{\mathrm{P}, \mathrm{NP}\})$, which takes the value 1 if the worker is employed in occupation τ and 0 otherwise. Let $0 \leq h_t \leq 1$ denote the fraction of hours (normalized to be between 0 and 1) that the individual chooses to work. Denote the labor-supply decision vector by $a_t = (d_{t-1}, \{I_{\tau t-1}\}_{\tau \in \{\mathrm{P},\mathrm{NP}\}}, h_t)$. Finally, c_t denotes the individual's consumption.

Preferences Individuals have preferences over non-market hours, l_t, and consumption, c_t. non-market hours are the difference between the total time endowment and the labor-market hours, $l_t = 1 - h_t$. Preferences are additively separable in consumption and leisure contemporaneously. Consumption is additively separable over time, whereas leisure is not.

The other two time-varying vectors of individual characteristics that determine the utility associated with alternative-consumption and non-market-hours allocations are x_t, which is a $K \times 1$ vector, and $(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})'$ which is a 3×1 vector. Define $z_t = (a_0, ..., a_{t-1}, x_t)$ to be a vector with the first $t-1$ elements capturing the dependence of the utility on the past labor-supply choices; the last element, x_t, is independently distributed over the population and evolves according to a known group-specific transition distribution function, $F_{i0}(x_{t+1} \mid z_t)$. The vector $(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})'$ is independent across the population and time and is drawn from a population with a common distribution function, $F_1(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})$. The distribution functions $F_{i0}(x_{t+1} \mid z_t)$ and $F_1(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})$ are absolutely continuous with continuously differentiable densities $f_{i0}(x_{t+1} \mid z_t)$ and $f_1(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})$, respectively.

The current-period utility function at date t for an individual i is U_{it},

$$U_{it} = d_t u_{i0}(z_t, \zeta_t) + u_{i1}(l_t; z_t) + u_{i2}(c_t; x_t, \varepsilon_{2t}) + (1 - d_t)\varepsilon_{0t} + d_t \varepsilon_{1t}, \tag{1}$$

where $u_{i0}()$ represents the utility cost of participating in the labor force, and ζ_t is an aggregate shock to this cost which is drawn from a distribution $F_\zeta(\zeta_{t+1} \mid \zeta_t)$. The disutility of working is indexed by gender, allowing for possible gender preference differences. We assume the utility function is concave, continuous, and twice differentiable everywhere in l_t and c_t. $\beta \in (0,1)$ denotes the common subjective discount factor. Individuals maximize lifetime expected utility: $E_t\left[\sum_{s=t}^{T} \beta^{t-s} U_{is} \mid z_t\right]$, subject to the budget constraint described below.

Budget Sets We assume that the assets markets are competitive and complete, and that there are no frictions in the markets for loans.[2] These assumptions allow us to compactly write the lifetime individual budget constraint:

$$E_0 \left\{ \sum_{t=0}^{T} \beta^t \lambda_t \left[c_{wt} + c_{mt} - \overline{S}_t \right] \right\} \leq W, \tag{2}$$

where $\overline{S}_t$ is the total household labor-market income (or the individual's income if s/he is single), λ_t is the expected price of the contingent claim, and W is bequest net of inheritance.[3] We assume that λ_t is distributed according to the known distribution $F_\lambda(\lambda_{t+1} \mid \lambda_t)$.

2.1.1 Frisch Demand for Consumption and the Indirect Utility

The additive and intertemporal separability of consumption along with the complete assets markets assumption allow us to characterize the Frisch Demand for consumption separately from the labor-supply decisions. Let η_i denote the Lagrange multiplier associated with the budget constraint in equation (2) for each household member.[4] The first-order conditions with respect to individual consumption c_{it} are

$$\frac{\partial u_{i2}(c_{it}; x_t, \varepsilon_{2t})}{\partial c_{it}} = \eta_i \lambda_t, \tag{3}$$

for all $t \in \{0, 1, 2 \ldots\}$ and $i \in \{\mathrm{w}, \mathrm{m}\}$. Let $S_{i\tau t}(h_t, .)$ be the salary paid to an individual that works h_t hours in occupation τ in period t. We can now write the current-period utility as

$$U_{it} = d_t u_{i0}(z_t, \zeta_t) + u_{i1}(l_t, z_t) + \eta \lambda_t \sum_{\tau \in \{\mathrm{P}, \mathrm{NP}\}} I_{\tau t} S_{i\tau t}(h_t, .) + (1 - d_t)\varepsilon_{0t} + d_t \varepsilon_{1t}, \tag{4}$$

reducing the individual problem to making the labor-supply decisions, a_t.

[2] Other papers that use similar assumptions include Altug and Miller (1990, 1998), Card (1990), Mace (1991), Townsend (1994), and Altonji, Hayashi, and Kotlikoff (1996).

[3] See Altug and Labadie (1994, p. 305) for a formal derivation of the budget constraint under these assumptions.

[4] Note that η_i is also the inverse of the Pareto weight in the planner's problem.

2.2 Firms

Technology Each firm in each occupation produces output ($y_{\tau t}$) using labor (h_t) and a homogeneous capital ($K_{\tau t}$) augmented by human capital and a skill index, which is a function of the worker's production-relevant variables, $z_t^{\mathcal{P}}$. These variables are a subset of z_t and include the labor-market history and other variables that affect the productivity. Each occupation is subject to aggregate productivity shocks to the effective unit of capital; therefore, the law of motion of the effective unit of capital is

$$K_{\tau t} = K_{\tau t-1} + \triangle_{\tau t}, \tag{5}$$

where $\triangle_{\tau t}$ is identically and independently distributed over time. Specifically, production at time t in occupation τ is given by

$$y_{\tau t} = Y_{\tau}(h_t, z_t^{\mathcal{P}}, K_{\tau t}).$$

Output in each occupation is the sum of output produced by the firms. The output in each occupation, $y_{\tau t}$, is measured in terms of the relative prices of the occupation products.

There are employer-specific hiring costs. These costs capture all possible training, administrative and other net costs associated with hiring a new worker. We assume that within occupations the costs are the same for all employers, and denote them by γ_{τ}.

We define a job as the number of hours worked for a given worker's characteristics, $z_t^{\mathcal{P}}$, and assume that each job is offered in one occupation only. Specifically, a firm in occupation τ offers a contract for hours h to a worker with characteristics $z_t^{\mathcal{P}}$ only if the expected lifetime net productivity is greater than it is in the other occupation,

$$\begin{aligned} (6)\quad & Y_{\tau'}(h_t, z_t^{\mathcal{P}}, K_{\tau' t}) - \gamma_{\tau'}(1 - I_{\tau' t-1}) + E_t \sum_{s=t+1}^{T} \beta^{s-t} I_{\tau' s}\left[Y_{\tau'}(h_s, z_s^{\mathcal{P}}, K_{\tau' s}) - \gamma_{\tau'}(1 - I_{\tau' s-1})\right] > \\ & \qquad Y_{\tau}(h_t, z_t^{\mathcal{P}}, K_{\tau t}) - \gamma_{\tau}(1 - I_{\tau t-1}) + E_t \sum_{s=t+1}^{T} \beta^{s-t} I_{\tau s}\left[Y_{\tau}(h_s, z_s^{\mathcal{P}}, K_{\tau' s}) - \gamma_{\tau}(1 - I_{\tau s-1})\right] \end{aligned}$$

We also assume that $Y_{\mathrm{P}t}(h_t, z_t^{\mathcal{P}}, K_{\mathrm{P}t})$ and $Y_{\mathrm{NP}t}(h_t, z_t^{\mathcal{P}}, K_{\mathrm{NP}t})$ cross only once and that $Y_{\tau}(h_t, z_t^{\mathcal{P}}, K_{\tau t})$ is twice continuously differentiable. Hence, for every occupation τ and worker's characteristics, $z_t^{\mathcal{P}}$, there is a range of hours offered in this occupation, $h \subset (\underline{h}_{\tau t}(z_t^{\mathcal{P}}, K_{\tau t}), \overline{h}_{\tau t}(z_t^{\mathcal{P}}, K_{\tau t}))$. Therefore, a choice of hours implies an occupation choice. Because not all hours are offered in each occupation, some changes in hours worked may involve changing occupation. This assumption captures the limited flexibility of hours offered within some occupations.

State Variables Following Rust (1987), we separate the systematic observed state variables from the idiosyncratic unobserved variables, and assume that they are conditionally independent of each other. Let the observed state variables, ω_t, be defined as $\omega_t = (z_t, \zeta_t, \eta\lambda_t, K_{\mathrm{P}t}, K_{\mathrm{NP}t})$. Note that the third element, $\eta\lambda_t$, is the marginal utility of wealth.

Timing At the beginning of each period, the state variables ω_t and $(\varepsilon_{0t}, \varepsilon_{1t})$ are realized. The workers then make participation, occupation, and hours decisions. Observing workers' choices, firms simultaneously offer salaries to each worker. Workers observe the offers and choose a firm. Finally, production occurs and the agents consume. This structure repeats itself every period.

3 Equilibrium Analysis

3.1 Symmetric Information

Assume that the information is symmetric and that the workers and firms observe ω_t and $(\varepsilon_{0t}, \varepsilon_{1t})$. Workers and firms can only commit to one-period non-contingent contracts. A contract, $S_{i\tau t}(h_t; \omega_t)$, consists of a salary, schedule of hours and worker's characteristics. Let $\pi_{\tau t}$ denote the continuation value of hiring a new worker in occupation τ in period t:

$$\pi_{\tau t}(h_t, \omega_t) = \sum_{s=t}^{T} \beta^{s-t} E_t \left[Y_\tau(h_s, z_s^{\mathcal{P}}, K_{\tau s}) - S_{i\tau s}(h_s; \omega_s) \mid h_t, \omega_t \right] - \gamma_\tau. \tag{7}$$

Similarly, let $\pi^e_{\tau t}$ denote the continuation value to the current employer of retaining a worker:

$$\pi^e_{\tau t}(h_t, \omega_t) = \sum_{s=t}^{T} \beta^{s-t} E_t \left[\left(Y_\tau(h_s, z_s^{\mathcal{P}}, K_{\tau s}) - S_{i\tau s}(h_s; \omega_s) \right) \mid h_t, \omega_t \right]. \tag{8}$$

The only difference between the expected profits from a worker who is already working in the firm and a new worker is the hiring costs. Since workers are price takers, we can derive the optimal salary schedule by backward induction. The net revenue from hiring a worker in the final period, T, is $Y_\tau(h_T, z_T^{\mathcal{P}}, K_{\tau T}) - \gamma_\tau$. Because of the free-entry assumption, the salary is derived by equating the potential employer's profit with zero,

$$\pi_{\tau T} = Y_\tau(h_T, z_T^{\mathcal{P}}, K_{\tau T}) - S_{i\tau T}(h_T; \omega_T) - \gamma_\tau = 0. \tag{9}$$

Therefore, the current employer in period T earns a rent of $\pi^e_{\tau T} = \gamma_\tau$. Next, consider an offer to a worker with characteristics ω_{T-1} in period $T-1$, and define $p_{i\tau t+1}(h_t, \omega_t)$ to be the probability that the worker will remain in the firm in period $t+1$ given the information set available in period t. The salary in period $T-1$ is again derived by equating the outside employer's expected profit over the two periods to zero,

$$\pi_{\tau T-1} = Y_\tau(h_{T-1}, z_{T-1}^{\mathcal{P}}, K_{\tau T-1}) - \gamma_\tau - S_{i\tau T-1}(h_{T-1}, \omega_{T-1}) + \beta p_{i\tau t+1}(h_t, \omega_t) \pi^e_{\tau T} = 0. \tag{10}$$

We obtain the competitive salary schedule by substituting $\pi^e_{\tau T} = \gamma_\tau$,

$$S_{i\tau T-1}(h_{T-1}; \omega_{T-1}) = Y_\tau(h_{T-1}, z_{T-1}^{\mathcal{P}}, K_{\tau T-1}) - \gamma_\tau + \beta \gamma_\tau p_{i\tau T}(h_{T-1}, \omega_{T-1}). \tag{11}$$

Therefore, by induction, we obtain the competitive salary schedule,

$$S_{i\tau t}(h_t,\omega_t) = Y_\tau(h_t, z_t^{\mathcal{P}}, K_{\tau t}) - \gamma_\tau + \beta\gamma_\tau p_{i\tau t+1}(h_t,\omega_t), \tag{12}$$

for all $h_t \in (\underline{h}_{\tau t}, \overline{h}_{\tau t})$. The per-period net surplus that a worker generates after the first year of employment is $Y_\tau(h_t, z_t^{\mathcal{P}}, K_{\tau t}) - S_{i\tau t}(h_t,\omega_t) = \gamma_\tau$. The current employer's share each period is $\gamma_\tau - \beta\gamma_\tau p_{i\tau t+1}(h_t,\omega_t)$, and the worker's share is $\beta\gamma_\tau p_{i\tau t+1}(h_t,\omega_t)$. Because labor markets are competitive, the worker's discounted expected lifetime earnings are equal to the worker's expected net output in an outside firm.

Equilibrium Labor Supply Let V_{1i} and V_{0i} denote the ex-ante conditional valuation functions associated with the decisions to work and not to work, respectively. The ex-ante conditional valuation function is defined as

$$\begin{aligned} V_{ki}(\omega_t) + \varepsilon_{kt} \equiv \max_{\{h_s;\{I_{\tau s}\}_{\tau\in\{\mathrm{P,NP}\}}\}_{s=t}^{T}} E_t \Bigg\{ \sum_{s=t}^{T} \beta^{s-t} \Big[& d_s u_{i0}(z_s,\zeta_s) + u_{i1}(l_s,z_s) \\ & + \eta\lambda_s \sum_{\tau\in\{\mathrm{P,NP}\}} I_{\tau s} S_{i\tau s}(h_s,\omega_s) + d_s\varepsilon_{1s} + (1-d_s)\varepsilon_{0s} \Big] \Bigg| \omega_t, d_t = k \Bigg\} \end{aligned} \tag{13}$$

for $k \in \{0,1\}$, and the necessary condition for the optimal participation decision is

$$d_i^o(\omega_t, \varepsilon_{0t}, \varepsilon_{1t}) = \begin{cases} 1 & \text{if } V_{1i}(\omega_t) + \varepsilon_{1t} \geq V_{0i}(\omega_t) + \varepsilon_{0t} \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

The probability of participation conditional on ω_t is

$$p_i^o(\omega_t) = E[d_i^o \mid \omega_t] = \int_{-\infty}^{V_{1i}-V_{0i}} (\varepsilon_{0t} - \varepsilon_{1t})\, \mathrm{d}F_1(\varepsilon_{0t}, \varepsilon_{1t}) \equiv Q(V_{1i}(\omega_t) - V_{0i}(\omega_t)). \tag{15}$$

Using the Bellman principle, the ex-ante conditional valuation of participation is

$$\begin{aligned} V_{1i}(\omega_t) + \varepsilon_{1t} = \max_{h_t;\{I_t\}_{\tau\in\{\mathrm{P,NP}\}}} & u_{i0}(z_t,\zeta_t) + u_{i1}(l_t,z_t) + \eta\lambda_t \sum_{\tau\in\{\mathrm{P,NP}\}} I_{\tau t} S_{i\tau t}(h_t,\omega_t) \\ & + \beta E_t \left[p_i(\omega_{t+1}) V_{1i}(\omega_{t+1}) + (1 - p_i(\omega_{t+1})) V_{0i}(\omega_{t+1}) \right] \mid \omega_t, d_t = 1 \} \end{aligned} \tag{16}$$

and for non-participation it is

$$V_{0i}(\omega_t) + \varepsilon_{1t} = u_{i1}(1, z_t) + \beta E_t \left[p_i(\omega_{t+1}) V_{1i}(\omega_{t+1}) + (1 - p_i(\omega_{t+1})) V_{0i}(\omega_{t+1}) \right] \mid \omega_t, d_t = 0 \}. \tag{17}$$

Next, we characterize the optimal hours and occupation decisions beginning with the optimal hours choice h_{it}^* in *each* occupation. The first-order condition with respect to hours is

$$
\begin{aligned}
(18) \quad & -\frac{\partial u_{i1}(l_t, z_t)}{\partial h_t} + \eta\lambda_t \left[\frac{\partial Y_\tau(h_t, z_t^{\mathcal{P}}, K_{\tau t})}{\partial h_t} + \beta\gamma_\tau \frac{\partial p_{i\tau t+1}(h_t, \omega_t)}{\partial h_t} \right] = -\beta E_t \left\{ \frac{\partial V_{0i}(\omega_{t+1})}{\partial h_t} \right. \\
& \left. + p_i(\omega_{t+1}) \frac{\partial [V_{1i}(\omega_{t+1}) - V_{0i}(\omega_{t+1})]}{\partial h_t} + \frac{\partial p_i(\omega_{t+1})}{\partial h_t} [V_{1i}(\omega_{t+1}) - V_{0i}(\omega_{t+1})] \mid \omega_t, h_t = h_{it}^*, I_{\tau t} = 1 \right\}
\end{aligned}
$$

and the occupation choice is given by

$$
(19) \qquad I_{i\tau}^0(\omega_t) \equiv I\{\underline{h}_{\tau t} < h_{it}^*(\omega_t) < \overline{h}_{\tau t}\}.
$$

Notice that equation (18) is obtained by maximizing separately over each open interval $(\underline{h}_{\tau t}, \overline{h}_{\tau t})$ and then choosing the occupation that yields the highest expected life-time utility. Under the assumption in (6), any choice of hours, h_t, conditional on z_t *maps* into a unique occupation choice in period t implying equation (19). The occupation choice affects the probability of working in the firm next period and therefore future salaries; this probability is given by

$$
(20) \qquad p_{i\tau t+1}(h_t, \omega_t) = \int Q(\omega_{t+1}) I_{i\tau}^0(\omega_{t+1}) f_{i0}(\omega_{t+1} \mid \omega_t, a_t) \, \mathrm{d}\omega_{t+1}
$$

The optimality conditions in Equations (16)–(19) describe the current and intertemporal trade-offs between consumption and non-market hours. If the hiring costs were zero, the third term on the left hand side of equation (18) would be zero, reducing it to the intertemporal labor supply decisions considered by Polacheck (1981) and Altug and Miller (1998). The left hand side of Equation (18) describes the standard effect of an extra hour's work on the trade-off between current utility from non-market hours and consumption. The marginal change in salary includes, in addition to the increase in output, the effect on future participation in the occupation, illustrating the new source of dynamic self-selection. The right hand side of equation (18) demonstrates the dynamic effect of a marginal change in hours on the probability of participation and the continuation values of working and not working, through the effect on labor market experience and non-market hours. Labor market experience affects the future productivity of workers whereas the stock of non-market hours affects the future disutility from working.

3.2 Asymmetric Information

Now assume that the worker observes both ω_t and $(\varepsilon_{0t}, \varepsilon_{1t})$, but the firms only observed ω_t^* which is a subset of the worker's state variables ω_t, and do not observe the iid shocks to the utility, $(\varepsilon_{0t}, \varepsilon_{1t})$. Similarly let x_t^* denote a subset of x_t that the firms observe; therefore the vector of information the firms observe is $\omega_t^* = (a_0, \ldots, a_{t-1}, K_{\mathrm{P}t}, K_{\mathrm{NP}t}, x_t^*)$. We assume that all production-relevant variables, $z_t^{\mathcal{P}}$, are observed. To make the complete-assets-markets assumption consistent with private information, we assume anonymous trades in the assets markets.

This information structure transforms the model into a signaling game. Hence, we use the perfect Bayesian equilibrium in this section of the model. At the beginning of each period, firms form a (common) set of prior beliefs on each individual worker's type, $\mu_{it}(\omega_t \mid \omega_t^*)$. Upon observing the worker's labor-market actions, a_t, firms update their beliefs about each worker's type. We denote the *posterior* beliefs by $\widetilde{\mu}_{it}(\omega_t \mid \omega_t^*, a_t)$. Notice that $\widetilde{\mu}_{it}(\omega_t \mid \omega_t^*, a_t)$ is used to form the prior beliefs in period $t+1$ because the types are persistent over time and evolve according to $F_{0i}(\omega_{t+1} \mid \omega_t, a_t)$.

Definition 3.1 (Equilibrium) *A perfect Bayesian equilibrium consists of labor-market strategies, σ_{it}, Frisch demand for consumption, c_{it}^o, the workers' contract-offer choices, $\{S_{i\tau t}\}_{\tau\in\{\mathrm{P,NP}\}}$ and a common belief system such that:*

1. *Each player's strategy is optimal given that player's beliefs and other players' strategies.*
2. *The posterior beliefs, $\widetilde{\mu}$, satisfy Bayes' rule when possible,*

$$\widetilde{\mu}_{it}(\omega_t \mid \omega_t^*, a_t) = \frac{\mu_{it}(\omega_t \mid \omega_t^*)\sigma_{it}(a_t \mid \omega_t)}{\int \mu_{it}(\omega_t \mid \omega_t^*)\sigma_{it}(a_t \mid \omega_t)\,\mathrm{d}\omega_t}, \tag{21}$$

   and, for all histories, types and actions,

$$\widetilde{\mu}_{it}(\omega_t \mid \omega_t^*, a_t) = \widetilde{\mu}_{it}(\omega_t \mid \omega_t^*, \widehat{a}_t) \text{ if } a_t = \widehat{a}_t.$$

3. *At the beginning of period $t+1$, firms form priors about the worker's type in that period based on past history:*

$$\mu_{it+1}(\omega_{t+1}|\omega_t^*, a_t) = f_{i0}(\omega_{t+1} \mid \omega_t, a_t)\widetilde{\mu}_{it}(\omega_t \mid \omega_t^*, a_t). \tag{22}$$

Existence is nontrivial in the model because within each period, we have the classic adverse selection problem (as formulated by Rothschild and Stiglitz, 1976). As pointed out by Hellwig (1987), this nonexistence result is sensitive to the timing of the players' moves. In our model, workers move first deciding how much to work, and then the firms make offers for these hours; the workers then choose which offer to accept. This difference in timing transforms the within-period game from a screening model into a signaling game, and the nonexistence result of the original Rothschild–Stiglitz model is broken. The proposition below establishes existence.

Proposition 3.1

1. *For all $h_t \in (\underline{h}_{\tau t}, \overline{h}_{\tau t})$, the optimal salary is given by*

$$S_{i\tau t}^0(h_t; \omega_t^*) = Y_\tau(h_t, z_t^{\mathcal{P}}, K_{\tau t}) - \gamma_\tau + \beta\gamma_\tau\widetilde{p}_{i\tau t+1}(h_t, \omega_t^*). \tag{23}$$

   Under asymmetric information the workers' optimal strategies are as described in the symmetric information case except for the following changes in the optimal labor-supply and *consumption decisions: The salaries in the state variables entering the valuation functions in equations (13) - (19) are now replaced by the salary in equation (23).*

2. *There exists a perfect Bayesian equilibrium characterized by the workers' strategies* $\left\{a_t^0, c_{it}^o\right\}$ *and the optimal salary in equation (23) in which the employers' beliefs satisfy the consistency requirement (Definition 3.1 (2–3)). An implication of the above is that*

$$\widetilde{p}_{i\tau t+1}(h_t, \omega_t^*) = \int Q(\omega_{t+1}) I_{i\tau}^0(\omega_{t+1}) f_{i0}(\omega_{t+1} \mid \omega_t, a_t) \widetilde{\mu}_{it}(\omega_t \mid \omega_t^*, a_t) \, \mathrm{d}\omega_{t+1}. \tag{24}$$

To highlight the difference asymmetric information makes, consider the first-order condition for hours under asymmetric information and compare it to the first-order condition under symmetric information (equation (18))

$$\begin{aligned} &-\frac{\partial u_{i1}(l_t, z_t)}{\partial h_t} + \eta\lambda_t \left[\frac{\partial Y_\tau(h_t, z_t^{\mathcal{P}}, K_{\tau t})}{\partial h_t} + \beta\gamma_\tau \frac{\partial \widetilde{p}_{i\tau t+1}(h_t, \omega_t^*)}{\partial h_t}\right] = -\beta E_t \left\{\frac{\partial V_{0it+1}(\omega_{t+1})}{\partial h_t}\right. \\ &\left. + p_{it+1}\frac{\partial [V_{1it+1}(\omega_{t+1}) - V_{0it+1}(\omega_{t+1})]}{\partial h_t} + \frac{\partial p_{it+1}}{\partial h_t}[V_{1it+1}(\omega_{t+1}) - V_{0it+1}(\omega_{t+1})] \mid h_t = h_{it}^*, I_{\tau t} = 1\right\}. \end{aligned} \tag{25}$$

If the information was symmetric, then $\frac{\partial \widetilde{p}_{i\tau t+1}(h_t, \omega_t^*)}{\partial h_t} = \frac{\partial p_{i\tau t+1}(h_t, \omega_t)}{\partial h_t}$. When the information is asymmetric, a marginal change in the amount of hours worked changes the employers' beliefs on the worker's type. These beliefs are a function of the current and past history of labor-supply decisions. The current choice of hours, therefore, has a dynamic effect on the future valuation functions on the right-hand side of equation (25) through the change in the publicly observed state variables, ω_{t+1}^*. As is standard in dynamic adverse selection models without commitment, the effect of current actions on future salary offers gives rise to the ratchet effect. The ratchet effect occurs because, for a low attachment type, revealing information on ω_t reduces the future expected rents (contract offers), thus, providing incentives to "pool" and hide the private information (a dynamic adverse selection model with similar features is analyzed in Laffont and Tirole (1988)). Notice that participation and occupation choices have similar signaling elements as they affect ω_{t+1}^* in a similar way.

The optimal contract provides insurance to the risk-adverse workers; hence it dominates contracts which impose all the risk on the workers. Therefore, selling the job to the worker, with partial commitment or without commitment, is not optimal when the workers are risk averse and the employers are risk neutral. While some partial-commitment contracts, such as those requiring that the worker makes a transfer upon leaving the job, may ameliorate the adverse-selection problem, they do not fully solve it (see for example Dionne and Doherty, 1994). We choose not to allow for partial-commitment contracts for two reasons. First, these termination transfers are not observed in the data. Second, if it is possible to write the long-term contract as a sequence of spot contracts in which a payment is made every period, our contract would be observationally equivalent.

Labor-Market-Outcome Gender Gaps In the model, gender differences in labor-market outcomes can arise due to preference and productivity differences. In our model, the distribution of x_t affects the utility function and is gender specific. Even if a man and a woman have the same characteristics and the information is symmetric, a difference in the distribution of x_t affects the probability of participation, and therefore the earnings. This gender difference in salaries can, in turn, give rise to differences in labor-supply decisions, and therefore, human capital of men and women.

When the information is asymmetric, statistical discrimination provides an additional source of gender earnings gaps. Women who have high probability of remaining in the firm may face lower earnings than men with a lower probability because, on average, women with similar observable characteristics are more likely to leave than men. The model may also give rise to discriminatory equilibria because of the possibility of multiple equilibria, even if ex-ante there are no differences between men and women of a given cohort.[5]

Change in the Gender Gap The following changes in exogenous factors could account for the narrowing in the observed gender earnings gap over time: occupation-specific aggregate productivity, demographic characteristics, costs of participation, and education. Suppose that women's participation costs are larger than men's, and suppose that there is an increase in the overall productivity within an occupation. Such an increase affects the earnings of all workers because $Y_\tau(h_t, z_t^{\mathcal{P}}, K_{\tau t})$ increases. If males' participation rate is high relative to women's, it is possible that women's human capital will increase more than men's, leading to a reduction in the gender earnings gap. The relative increase in women's participation can cause a relative increase in employers' beliefs, further reducing the gender earnings gap. Demographic changes, such as a decline in fertility, and a decline in the cost of producing home goods can have a similar effect on the change in the patterns of female labor market participation and the gender earnings gap.

An additional factor that the asymmetric information introduces and can account for the narrowing gender earnings gap is changes in beliefs across cohorts. While the equilibrium is characterized for each cohort separately, we observe in the data several overlapping cohorts allowing us to quantify the difference in employers' beliefs across cohort.

4 Empirical Implementation

The theoretical model is written and solved for a single cohort; the empirical implementation, however, uses data from a number of overlapping cohorts. A cohort is characterized by the year of birth and years of completed education. The large number of cohorts in the data makes it undesirable to enumerate all the cohorts; instead, we note that there is a one-to-one mapping between a cohort and age, current year and years of completed education. The functions that depend on cohort are the transition function of the time varying characteristics, $F_{i0}(x_{t+1} \mid z_t)$, and the employers beliefs,

[5] See Tirole (1996) for a dynamic-adverse-selection and statistical-discrimination model. The difference between this model and Tirole's arises because the matching in Tirole's model between firms and workers is random.

$\widetilde{p}_{i\tau t+1}(h_t, \omega_t^*)$. These two functions depended on age and years of education; to make them depend on cohorts we add a year dummy. These functions are estimated nonparametrically using kernel density estimation.

The private information we use in the estimation consists of data that the econometrician observes as well as persistent heterogeneity which is unobserved to the econometrician. To select which variables are private information, we first choose variables which are not observed by potential employers and the worker is not obligated to report truthfully by law.[6] We then develop a test for private information which allows us to test the assumption that these variables are private information. The persistent unobserved heterogeneity is the inverse of the Pareto weights (η_i) that enter the Frisch demand for consumption. It depends on the family budget constraint and is therefore private information by nature. We estimate this persistent unobserved heterogeneity in a first stage as fixed effect, and use it in the final stage as *if* the econometrician observes it. This estimation procedure is possible because of the additive separability of consumption and non-market hours in the per-period utility specification (see section 5 for more details). Below, we describe the data and parametric assumptions we make about the model's underlining primitives.

4.1 Data

This paper uses the Panel Study of Income Dynamics (PSID) because it includes a long panel of matched data on individuals' working, marriage and child birth histories for overlapping cohorts. The two main disadvantages of using the PSID is the lack of nonfood consumption and good job-to-job–transition data. As a result, job switches in the data are occupation changes and transition in and out of the labor force.[7]

The data is taken from the Family File, the Childbirth and Adoption History File, the Retrospective Occupation File, and the Marriage History File of the PSID. The sample contains individuals who were either the *Head* or *Wife* of a household in the year of the interview. Individuals are classified into two occupation categories, professional and nonprofessional. We only keep White individuals between the ages of 25 and 65 in our sample. After eliminating those with missing values, we are left with 15,702 individuals between the years 1968 and 2007 of which 46% are women. However, we only have annual labor-market data for the years 1968 to 1997. The construction of our sample and the definitions of the variables are described in greater details in Appendix C

The average annual earnings for men increased by roughly 58% over the period, from \$40,000 per year in year-2000-constant dollars in 1968, to \$63,000 in 1997. Meanwhile, the average annual earnings for women increased by around 113% over the same period, from \$16,200 in 1968 to \$34,000 in 1997. As Figure 2 shows, the earnings gap declined by around 19% (10%) in professional (nonprofessional) occupations over the period. Note that the earnings gap is normally 50% larger than the wage gap because women not only earn less, but work fewer hours per year than men. We

[6] Notice that while the PSID survey data are readily available, the specific individual information recorded in them is anonymous and cannot be obtained by potential employers.

[7] This is consistent with the equilibrium job turnover in our model in which workers do not change employers within an occupation.

therefore, focus on the earnings gap in order to capture both dimensions.

Table 1 contains summary statistics of our main labor-market and human-capital variables. The participation rate for men is relatively constant over the sample period with a slight decline toward the end. In contrast, the participation rate for women increased significantly, from 54% in 1968 to 76% in 1997. The average annual hours worked by men is also relatively constant, but the average annual hours worked by women increased by roughly a third, from 1,400 hours per year in 1968 to 1,868 hours in 1997. Although the hours-worked gap between women and men has narrowed significantly, it remains large. The gender gap in the average years of completed education has almost completely closed by 1997.

Women's representation in the professional occupations increased by roughly 64% over the sample period, going from 28% of the occupation in 1968 to around 46% of the occupation by 1997. At the same time, the fraction of women in the nonprofessional occupations remained constant over the period.

Table 2 contains summary statistics of our main demographic and wealth variables. The sample includes the household size and the age distribution and number of children. Both the household size and number of children declined, but the decline in the number of young children is the most pronounced. Roughly 80% of our final sample is married in any given year of the sample.

Our measure of consumption is food consumption. Food consumption expenditures for a given year is obtained by summing the values of annual food expenditures for meals at home, annual food expenditures for eating out, and the value of food stamps received for the year. Household food consumption has declined over the period while the per capita food consumption has increased.

4.2 Consumption

Following the literature on the estimation of the consumption Euler equation (see Browning and Lusardi, 1996, for a survey), we use the Constant Relative Risk Aversion (CRRA) utility function:

$$u_{i2}(c_{nt}, x_{nt}, \varepsilon_{2nt}) = \exp(x'_{nt}B_4 + \varepsilon_{2nt})c_{nt}^{\alpha}/\alpha, \tag{26}$$

where $1-\alpha$, the coefficient of relative risk aversion, must be positive in order for the utility function to be concave. The function $\exp(x'_{nt}B_4 + \varepsilon_{2nt})$ is effectively an adult equivalence scale; it is used in the literature to get around the problem that consumption is typically measured at the household level, but the object of interest is the marginal utility of consumption at the individual level. Substituting the utility function into equation (3) gives

$$\exp(x'_{nt}B_4 + \varepsilon_{2nt})c_{nt}^{\alpha-1} = \eta_n \lambda_t. \tag{27}$$

We take logs and get a log-linearized Euler equation:

$$\ln c_{nt} = x'_{nt}(1-\alpha)^{-1}B_4 + (\alpha-1)^{-1}\ln \eta_n + (\alpha-1)^{-1}\ln \lambda_t + (1-\alpha)^{-1}\varepsilon_{2nt} \tag{28}$$

This equation captures, in a simple and parsimonious way, several important features of our model worth noting. x'_{nt} captures changes in the individual circumstances over the life cycle. Therefore, if an individual changes households (e.g., marriage, divorce), the individual Pareto weights η_n^{-1} would be identified assuming that the utility functional form is invariant across these different household configurations controlling for x'_{nt}. The above condition also captures implicitly the intertemporal substitution (the elasticity of intertemporal substitution is $(\alpha-1)^{-1}$), precautionary-saving and bequest motives.

4.3 Production Function

The output for occupation τ in year t measured in relative prices is specified as

$$(29)\quad \begin{aligned} Y_\tau(h_{nt}, z_{nt}^{\mathcal{P}}, K_{\tau t}) = K_{\tau t} + b_{\tau 1} h_{nt} + b_{\tau 2} h_{nt}^2 + \textstyle\sum_{r=1}^{\rho} b_{\tau 3r} h_{nt-r} \sum_{r=1}^{\rho} b_{\tau 4r} d_{nt-r} + b_{\tau 5} age_{nt} \\ + b_{\tau 6} age_{nt}^2 + b_{\tau 7} age_{nt} \times education_n + \nu_n, \end{aligned}$$

where ν_n is an individual effect fixed across occupations and time. The aggregate shock to productivity, $K_{\tau t}$, enters the production function linearly, capturing the relative increases in productivity or changes in the relative prices of goods produced in an occupation. We allow for the possibility of nonlinearity in the productivity of current and past hours. This specification was chosen for its strong support in the empirical literature (see, among others, Eckstein and Wolpin, 1989; Altug and Miller, 1998). It implies that human capital is general in nature but has occupation-specific returns.

4.4 Utility from Non-market Hours

Following the literature, we assume that the idiosyncratic preference shocks, $(\varepsilon_{0nt}, \varepsilon_{1nt})$, are distributed as a Type I extreme value with variance parameter σ^2 and mean zero. This distributional assumption for the preference shocks implies that the inverse of the $E[d_i^0 \mid \omega_t]$ is given by

$$(30)\qquad Q^{-1}(p_i^o(\omega_t)) = \sigma \ln[p_i^o(\omega_t)/(1-p_i^o(\omega_t))],$$

and the expected values of the shocks, given the state and decisions, have the following forms

$$(31)\qquad E[\varepsilon_{0t} \mid \omega_t, d_t^o = 0] = \varphi_0(p_i^o(\omega_t)) = \frac{\xi}{\sigma} - \sigma \ln[(1-p_i^o(\omega_t))]$$

and

$$(32)\qquad E[\varepsilon_{1t} \mid \omega_t, d_t^o = 1] = \varphi_1(p_i^o(\omega_t)) = \frac{\xi}{\sigma} - \sigma \ln[p_i^o(\omega_t)],$$

where ξ is the Euler constant.

We parametrize the cost of working to include a linear additive aggregate shock that does not depend on gender, non-separability in past participation, and gender differences in the effect of the observed time-varying characteristics. It takes the form of

$$u_{i0}(z_t, \zeta_t) = \zeta_t + \sum_{s=1}^{2} \kappa_{is} d_{t-s} + x_t' B_{i1}.$$

The specification of the utility from non-market hours follows the literature on non-separable utility (see, among others, Hotz, Kydland, and Sedlacek, 1988; Altug and Miller, 1990; Becker, Grossman, and Murphy, 1994) but allows the effect to differ by gender. Specifically, we assume the following functional form.

$$u_{i1}(l_t; z_t) = x_t' l_t B_{i2} + \theta_{i0} l_t^2 + \sum_{s=1}^{2} \theta_{is} l_t l_{t-s} \tag{33}$$

This specification allows for concavity in current hours (i.e., $\theta_{i0} < 0$) and for past and current hours to be either compliments or substitutes (i.e., $\theta_{is} > 0$ or $\theta_{is} < 0$).

4.5 State Variables

A large state space is a major difficulty in estimating dynamic programming models. The literature typically addresses this problem by either restricting the agent's observed state variables or by assuming functional forms of the production and utility functions that limit their dependence on past decisions. We can only partially adopt these strategies, however, because our model is a signaling game and past decisions are central to modeling reputation. Specifically, in a perfect Bayesian equilibrium, the beliefs depend on all past decisions, not just on those relevant to the payoffs (i.e., utility and output). In order to make the estimation feasible and at the same time capture the signaling content of past decisions, we restrict the beliefs to depend on the complete labor-market history in the past three years and the total number of years worked in the occupation. This restriction, along with the Markovian assumption on the transition of the exogenous variables, allows the model to satisfy the stationarity property called *finite state dependence*.

A decision process is said to exhibit finite state dependence if there exists a finite sequence of future decisions that lead to the same state variables irrespective of what decision is taken today. This property can be illustrated with the following example. Consider two individuals making the discrete decisions to work or not to work in two consecutive years. Suppose that the state variables are restricted to include only the number of years of experience, and that both individuals have five years of experience in the current period. Suppose individual A works in the current period and does not work in the next period while individual B does not work in the current period but works in the next period. Entering the next period, individual A has six years of experience while individual B has five years of experience. At the beginning of the third period, however, both individuals have six years of experience. This decision process is said to exhibit finite state dependence since the two different decision sequences (work, not work) and (not work, work) will result in both individuals having the same work experience entering the third period, irrespective of the decision today. This definition can be generalized to account for stochastic transition of the state variables. In that case, a decision process is said to exhibit *stochastic finite state dependence* if there exists a finite sequence of future decisions that leads to the same distribution of the state variables irrespective of what decision is taken today.

5 Estimation and Identification

One of the main problems in estimating game-theoretic models is the possibility of multiple equilibria. Multiple equilibria induces indeterminacy in standard estimation criterion functions that maps the structural parameters of the model to the observed distribution in the data. One solution the literature proposes is based on the following intuition. Conditional on other players' equilibrium strategies, each player's decision becomes a single-agent maximization problem (i.e., the best response function). This maximization problem is a necessary condition which holds in all equilibria; hence, an estimator of the structural parameters, based on this necessary condition, will be well defined if the model is identified.

A model is identified if we can show that the structural parameter vector, θ, with the true population value θ^* can be written as a functional of the conditional distribution of the observable variable. In addition, that functional has to return the value θ^* in all the admissible structures of the model (see Chesher, 2007, for more details). We show below that our model is identified and develop an estimator that gets around the problem of multiple equilibria. In the online appendix we show that our identification results are more general, and that the model is semiparametrically identified.

5.1 Identification

The assumption that the utility function is additively separable in non-market hours and consumption is necessary for identification. Using the variation in consumption over time, across individuals and across different household configurations, allows the identification of the Pareto weights η_n^{-1}(unobserved heterogeneity) and the aggregate price of consumption λ_t from equation (28).

The production functions and hiring costs are identified from the variation in salaries across occupations and gender for individuals with different patterns of future labor supply (i.e., propensity to change occupation, labor-market participation and amount of hours worked). We assume that the unobserved individual-specific effect, ν_n, is correlated with all the time-invariant publicly observed variables such as gender and completed education, but not with the time-varying variables (see Mundlak, 1961, 1978; Chamberlain, 1984). Taking the expectation over ν_n of the zero-profit condition transforms the earnings equation into a partial linear-panel-data model, where the nonlinear part is the expected value of ν_n conditional on the time-invariant publicly observed variables. We can now use standard panel-data transformations to show that the time-varying production function parameters and the hiring cost are identified. We then use the individual-level zero-profit condition to identify the individual-specific fixed effects. Appendix B contains a detailed description of this result.

Each equilibrium generates specific conditional choice probabilities. In principal, any set of observed choices in the data could be generated by a mixture of different equilibria. Following the literature on estimation of games, we assume that conditional on observed characteristics,[8] the data is generated by only one equilibrium (see Bajari, Benkard, and Levin, 2007; Aguirregabiria and Mira, 2007; Pesendorfer and Schmidt-Dengler, 2008). Under this assumption, the conditional choice probabilities are identified. Notice, however, that this assumption allows for different equilibria to be played across different cohorts and across different types of workers.

The utility from non-market hours is identified from the variation in the hours worked, labor-force participation and occupation choices made over time given different salaries. This identification result relies on the above assumptions on the equilibrium selection, and the stochastic finite state dependence of the state variables. To set some notation, let $\omega_{kt}^{(s)}$ denote the state in period $t+s$ if, at time t, the k^{th} option is taken—that is, $d_t = k$—and the decisions along the finite sequence that leads to the same state variable are taken in the following s periods. Denote by $p_{kit}^{(s)}$, the probability that $d_{t+s} = 1$ conditional on $\omega_{kt}^{(s)}$—i.e., $p_{kit}^{(s)} = E\left[d_{t+s} \mid \omega_{kt}^{(s)}\right]$. Then the optimal participation decisions described in equation (15) and the hours Euler equations, (18) and (25) (depending on whether information is symmetric or asymmetric) become

$$
\begin{aligned}
(34) \quad m_{i2t} \equiv \eta\lambda_t \sum_{\tau\in\{\text{P},\text{NP}\}} I_{\tau t}^{o} S_{i\tau t}^{o} + \sigma \sum_{s=1}^{3} \beta^s \ln\left(\frac{1-p_{1it}^{(s)}}{1-p_{0it}^{(s)}}\right) - \sigma \ln[p_{it}/(1-p_{it})] \\
+ \zeta_t + \sum_{s=1}^{2} \kappa_{is} d_{t-s} + x_t' B_{i1} - x_t' h_t B_{i2} - \theta_{0i}\left(1-l_t^2\right) - \sum_{s=1}^{2} \theta_{si} h_t (l_{t-s} + \beta^s)
\end{aligned}
$$

and

$$
(35) \quad m_{i3t} \equiv \eta\lambda_t \sum_{\tau\in\{\text{P},\text{NP}\}} I_{\tau t}^{o} \frac{\partial S_{i\tau t}^{o}}{\partial h_t} + \sigma \sum_{s=1}^{3} \beta^s \left(1-p_{1it}^{(s)}\right)^{-1} \frac{\partial p_{1it}^{(s)}}{\partial h_t} - x_t' B_{i2} - 2\theta_{0i} l_t + \sum_{s=1}^{2} \theta_{si}(l_{t-s} + \beta^s),
$$

where $E[m_{ijt}] = 0$ for $j \in \{2, 3\}$. These equations are obtained by applying the Hotz and Miller (1993) inversion and the alternative valuation-function representation to the participation equation (15), and hours Euler equations, (18) or (25), in Altug and Miller (1998). Appendix B derives the above equations. By inspection, the model is identified once the consumption and salary equations, the beliefs and the conditional choice probabilities are identified.

5.2 Estimation

Consumption We estimate equation (28) using standard panel-data estimation techniques. We estimate the marginal utility of wealth, $(1-\alpha)^{-1}\eta_n$, for each individual in our sample. Because

[8] Observable characteristics are either characteristics directly observed in the data or characteristics that can be identified from the data separately, like the individual fixed effects in production and consumption.

the standard fixed-effect estimates (used in Heckman and MaCurdy, 1980) are biased for small T, we use consumption data for the period 1969–2007; these consumption data are available annually for the years 1998–2008. We are unable to use data for the years after 1997 to estimate the rest of the model because other variables needed for the estimation are not available annually.

Equilibrium Beliefs, Their Derivatives and Conditional Choice Probabilities The equilibrium beliefs for each occupation, $\widetilde{p}_{in\tau t+1}$, are computed as a nonlinear regression of the product of the next-period participation and the occupation choice index, $d_{nt+1} \times I_{n\tau t+1}$, on the current-period's publicly observed state variables and hours worked, h_{nt}, conditional on working today in occupation τ. The nonparametric estimates of $\widetilde{p}_{in\tau t+1}$, denoted by $\widetilde{p}^{\mathcal{N}}_{in\tau t+1}$, are computed using the kernel estimator. We then estimate their derivatives using the standard nonparametric-derivative kernel estimator (see Pagan and Ullah, 1999).

In contrast to the beliefs, the conditional choice probabilities are defined from the workers' perspective and not the employers'. The elements included in x_{nt} are the number of individuals in the family unit, the number of children younger than three, the number of children between three and fourteen, age, years of completed education, marital status, spouse's years of education, and gender. The conditional choice probabilities, p_{int}, are computed using nonlinear regressions of the participation index, d_{nt}, on the current state-variables vector, $\omega^{\mathcal{N}}_{nt} \equiv (z^{\mathcal{N}}_{t,}{}' K^{\mathcal{N}}_{\mathrm{P}t}, K^{\mathcal{N}}_{\mathrm{NP}t}, \eta^{N}_{n} \lambda^{\mathcal{N}}_{t})'$, where the superscript $\mathcal{N}$ denotes an estimated quantity.

Estimation of the Finite-State Path Probabilities and their Derivatives We first characterize all the different possible sequences of choices that can lead to the same labor-market history at a certain point in time due to the assumption of finite state dependence. The hypothetical labor-market history is defined as

$$
\begin{aligned}
(36) \quad z^{(s)}_{1nt} = (&NY_{n1t-1+s}, NY_{n2t-1+s}, d_{nt-3+s}I_{n1t-3+s}, d_{nt-3+s}I_{n2t-3+s}, \ldots, d_{nt-1}I_{n1t-1}, \\
&d_{nt-1}I_{n2t-1}, I_{n1t}, I_{n2t,}, 0, \ldots, 0, h_{nt-3+s}, \ldots, h_{nt-1}, h^{*}_{nt}, 0, \ldots, 0, x'_{nt+s})
\end{aligned}
$$

and

$$
\begin{aligned}
(37) \quad z^{(s)}_{0nt} = (&NY_{n1t-1+s}, NY_{n2t-1+s}, d_{nt-3+s}I_{n1t-3+s}, d_{nt-3+s}I_{n2t-3+s}, \ldots, d_{nt-1}I_{n1t-1}, \\
&d_{nt-1}I_{n2t-1}, 0, 0, I_{n1t}, I_{n2t,}, 0, \ldots, 0, h_{nt-3+s}, \ldots, h_{nt-1}, 0, h^{*}_{nt}, 0, \ldots, 0, x'_{nt+s})
\end{aligned}
$$

where

$$NY_{n\tau t-1+s} = NY_{n\tau t-1} + I_{n\tau t}$$

for $s = 1, 2, 3$. The vector $z^{(s)}_{1nt}$ is the state vector of an individual in period $t+s$, with a state vector z_{nt-1} at period t who chooses to work in period t making optimal hours and occupation choices and then chooses not to work in the $s-1$ periods following period t. Similarly, the vector $z^{(s)}_{0nt}$ is the state vector of an individual with characteristics vector z_{nt-1} entering period t after choosing *not* to work in period t, but choosing to work in period $t+1$; the hours and occupation choices in period $t+1$ are the same as the optimal hours and occupation choices *the individual would have made in period t if s/he worked.* For the following $s-2$ periods after period $t+1$, this individual chooses not to work. Notice that these two sequences of decisions will lead to the same labor-market history in period $t+4$. Assuming a Markovian transition of the variables x_{nt}, the decision sequences satisfy the stochastic finite state dependence assumption. Let us define the following participation indices that correspond to the decision sequences that lead to a state vector $z_{knt}^{(s)}$.

$$d_{1nt}^{(s)} = (1 - d_{nt-1}) \times \ldots \times (1 - d_{nt-s-1}) \times d_{nt-s}$$

and

$$d_{0nt}^{(s)} = (1 - d_{nt-1}) \times \ldots \times d_{nt-s-1} \times (1 - d_{nt-s}).$$

Note that $d_{1nt}^{(s)}$ and $d_{0nt}^{(s)}$ take the value 1 if the individual entering period t has followed a decision path identical to the paths specified in $z_{1nt}^{(s)}$ and $z_{0nt}^{(s)}$. Let $\omega_{knt}^{(s)\mathcal{N}} \equiv (z'_{nt+s}, K_{\text{Pt}}^{\mathcal{N}}, K_{\text{NPt}}^{\mathcal{N}} \eta_n^{\mathcal{N}} \lambda_{t+s}^{\mathcal{N}})'$ for $k = \{0, 1\}$ be the empirical counterpart of the hypothetical state. Recall that $p_{kit}^{(s)} = E[d_{t+s} \mid \omega_{kt}^{(s)}]$; hence, it can be estimated as nonlinear regressions of the participation index, d_{nt}, on the hypothetical state, $\omega_{knt}^{(s)\mathcal{N}}$, conditional on $d_{knt}^{(s)} = 1$. We then evaluate the term $\partial p_{i1nt}^{(s)} / \partial h_{nt}$, using the standard nonparametric derivative kernel estimator (Pagan and Ullah, 1999).

Earnings Equations We estimate the parameters of the earnings equations using the specification of the production function, the assumption that employers observe the occupation and hours choices in the past three years as well as the total number of years worked in each occupation. We derive a moment condition using the assumption of the free-entry condition. From the free-entry condition, the expected lifetime profits of an employer from each contract is zero, conditional on the employer's information. Defining $d_{n\tau t} = I_{n\tau t} \times d_{nt}$, we derive the moment condition based on this condition:

(38)

$$m_{n\tau t}(\theta_{w\tau}) \equiv S_{in\tau t} - K_{\tau t} - b_{1\tau} h_{nt} - b_{2\tau} h_{nt}^2 - \sum_{r=1}^{\rho} b_{3r\tau} h_{nt-r} - \textstyle\sum_{r=1}^{\rho} b_{\tau 4r} d_{nt-r} - Z_{nt}{}' B_{5\tau} - v_n + \gamma_\tau - d_{nt+1} I_{n\tau t+1} \beta \gamma_\tau$$

where $\theta_{w\tau}$ denotes the $(7+T)$-dimensional vector of parameters we estimate. Using equation (38), we obtain a set of orthogonality conditions which can be exploited to estimate $\theta_{w\tau}$ using standard panel data techniques of optimal instrumental variables or two-stage least squares.

Utility of Non-market Time The remaining parameters are estimated by GMM using the empirical counterpart of equations (34) and (35). They are constructed by substituting the estimated quantities: β, $\eta_n \lambda_t$, $S_{i\tau t}^o()$, p_{int}, $p_{i0nt}^{(s)}$, and $\frac{\partial p_{i1nt}^{(s)}}{\partial h_{nt}}$. The remaining details of the implementation, the estimation of p_{int} and $p_{i0nt}^{(s)}$, and the $\frac{\partial p_{i1nt}^{(s)}}{\partial h_{nt}}$, and the asymptotic properties of the estimator are in the sections A3 and A4 in the online appendix.

6 Empirical Results

The estimation results are shown in Tables 3 through 6 and Figures 1 through 3. The results from the earnings equation are reported in Table 4, and Table 6 reports the non-market hours utility parameters, the risk-aversion parameter, and the variance of the idiosyncratic preference shocks; these results are consistent with the empirical regularities that women sort more into nonprofessional occupations, spend more hours on non-market activities and accumulate less human capital than men. The estimates of the aggregate productivity shocks in professional and nonprofessional occupations are consistent with the increase in women's representation in professional occupations over time.

We show that the symmetric and asymmetric information models are nested and reject the hypothesis that the data were generated by the symmetric-information model. We then show that the asymmetric-information model accurately predicts the gender earnings gap, age–earnings profiles, and the changes in the gender earnings gap over time and over the life cycle. Figures 2 and 3 show the fit of the model in several dimensions.

6.1 Earnings Equations

Asymmetric Information, Specification and Goodness-of-Fit Tests We begin by developing a three-step test of the asymmetric information. Under the hypothesis that the information is symmetric, the zero-profit condition implies that $E_t[m_{n\tau t}(\theta_{w\tau}) \mid \omega_t, i, d_{nt}I_{n\tau t} = 1] = 0$; under the asymmetric-information hypothesis: $E_t[m_{n\tau t}(\theta_{w\tau}) \mid \omega_t^*, i, d_{nt}I_{n\tau t} = 1] = 0$. Under asymmetric information, the employers have a subset of the information available to them under symmetric information. Therefore, if the model is identified under asymmetric information, the variables which are assumed to be private information serve as overidentifying restrictions under the null hypothesis that the information is symmetric. This leads to a natural nested test of both models using the standard J-test in an optimal GMM setting.

In the first step, the overidentifying restrictions test rejects the earnings equation specification under symmetric information (the J-statistic is 76.3). That is, it rejects the hypothesis that the production function takes the functional form we specified and that the variables *Number of individuals in the Household*, *Number of Kids, Martial Status*, *Marginal Utility of Wealth*, *Spouse Labor-Market Income and Spouse Education* are included in the information set of the employer.

In the second step, we perform an overidentifying restrictions test for the earnings equation under the null that the information is asymmetric. There are different overidentifying restrictions built into the asymmetric information model; as noted in the identification section, we only need one of the exclusion restrictions, gender or cohort, to achieve identification. Using as instruments five years of labor-market histories, gender and cohort, we were not able to reject the overidentifying restrictions.

Since the production function is similar under symmetric and asymmetric information, the only difference between the two specifications of the earnings equation is the information structure.

Rejecting the information structure in the symmetric information case is not sufficient to show that this information is private. These variables should only be included in the information set of the employer if they affect profits. In the third step, we test and confirm that these variables predict future occupation participation. We can, therefore, conclude that the earnings equation specified under the symmetric information assumption is rejected because the variables above affect the employers' profits yet they are not priced. For the rest of the paper, we use the asymmetric model; we estimate the earnings equation using both a two-stage least squares (2SLS) and an optimal GMM. The results are identical so we only present the 2SLS results below.

We assess the model's fit by looking at how well it predicts three important features of the data: the (unconditional) evolution of the gender earnings gap over the life cycle and its evolution over the life-cycle conditional on continuously working full time, the age–earnings profiles, and the change in the gender earnings gap over time. Figure 2 shows that the model does a good job in predicting the trend in the gender earnings gap over time. Figure 3 shows that the model predicts well the increase in the gender earnings gap over the life cycle, and its decline for men and women who work continuously full-time. The main force driving the first result is the increase gender gap in human capital, while the latter is driven by the faster increase in earnings of women who consistently work full time relative to men's.

Parameter Estimates The estimation results in Table 4 reveal comparatively larger returns to experience in professional occupations, faster depreciation of human capital in nonprofessional occupations, and a larger hiring cost in professional occupations. For example, the returns to working full time (40 hours per week) versus part-time (20 hours per week) in the previous year are \$2,751 in professional occupations versus \$2,284 in nonprofessional occupations; this difference is larger for experience accumulated with a two-year lag. The cost of hiring a new worker (the discount factor is assumed to be 0.95) is \$4,502.65 in professional occupations versus \$3,174.74 in nonprofessional occupations. The three factors we describe above affect occupational sorting. To see that, consider a male and a female choosing an occupation, and suppose that the female's probability of an employment-spell interruption and her time spent on non-market activities are larger than the male's. The salary cost in occupation τ caused by the hiring costs and probability of interruption of employment spells, $\gamma_\tau(1 - \beta\widetilde{p}_{i\tau t+1}(h_t, \omega_t^*))$, is comparatively lower in nonprofessional occupations. In addition, the depreciation rate of human capital is higher in the nonprofessional occupations. The returns to labor-market experience, however, are larger in professional occupations. Thus, women, in equilibrium, are more likely to sort into nonprofessional occupations, which penalize interruption of employment spells less, despite the lower returns to experience. Men, on the other hand, accumulate more human capital, and have less frequent interruptions of employment spells than women; this reduces the salary cost associated with the hiring costs, and therefore, they are more likely to sort into professional occupations. Consistent with other studies (see Card and DiNardo, 2002; Lee and Wolpin, 2010), Figure 1 reveals a significant increase in aggregate productivity in professional occupations since the mid 80s, and a smaller increase in the 90s in nonprofessional occupations; these findings are consistent with women's increase in representation in professional occupations.

Other Source of Discrimination Our model captures statistical discrimination linked to the probability of future interruption of the employment spell. There are two other notable types of discrimination we do not consider: Becker's taste-based discrimination, and statistical discrimination linked to unobserved productivity (see Coate and Loury, 1993). In its simplest form, Becker's model would imply a fixed differential wage for men and women; if taste does not change over time, it should appear as a gender gap in our individual-specific productivity estimates. The second form of discrimination can affect our estimates in the following way. Suppose that there are unobserved productivity gender differences, and suppose that education is used as a costly signal of productivity; then employers would pay workers their expected productivity given their education. In our sample, however, education level does not change over the life cycle; therefore, this form of discrimination should appear in the individual-specific productivity estimates.

Table 5 shows the regression results of the unobserved individual-specific effects on education and gender. The results reveal that the coefficient on the male dummy is negative, hence ruling out a taste-based discrimination in its simplest form. The coefficient on the interaction of the male dummy and education is positive, however, indicating that there is a gender differential in the compensation for a given education level. This finding is consistent with the hypothesis that the education signal of productivity is different for males and females. This effect, however, is not large.

6.2 Utility of Non-market Time

Before we proceed, it is worth noting that we cannot reject the overidentifying-restrictions test for the participation and hours Euler equations. As in other studies that used the PSID and assume additive separability of consumption and non-market hours (see Zeldes, 1989; Lawrence, 1991; Shea, 1995; and among others), our estimate of the elasticity of intertemporal substitution is -1.01 (i.e., $-1/(1-\alpha)$), but is not statistically significantly different from zero. That is, we can not reject that the utility takes the log form; this result is presented in column (3) in Table 6.

Reviewing Column (1) in Table 6 shows that there is no clear trend in the cost of participation over time; this is in contrast to the results in Greenwood, Seshardri, and Yorukoglu (2005); our time period, however, does not completely overlap with theirs. Reviewing Columns (2) and (3) of Table 6 reveals that there is a significant non-separability in both the disutility of working and the marginal utility of non-market hours, and these patterns differ by gender. For example, there are significant complementarities between current and previous year's participation and non-market hours for women; we do not find these complementarities in participation for men. These findings are consistent with Becker's (1965) theory of home-production division of labor and habit formation. Kids and marriage have the expected effect on the cost of participation and the utility from non-market hours for women, with young kids increasing the cost of participation and older kids increasing the utility from non-market hours for women; we find that the opposite holds for men.

6.3 Dynamic Decomposing Gender Gap

Why Is There a Gender Earnings Gap? To quantify the effect of labor-market frictions and asymmetric information on the observed gender earnings gap, labor supply, and occupation composition, we simulate the model with no hiring cost and with symmetric information. Under these two scenarios, the model has a unique equilibrium which can be found by backward induction. We compare the outcomes in each scenario to the observed outcomes in the data, which are assumed to be generated by the model with asymmetric information. In the model with no hiring costs, wages are the marginal productivity. Therefore, the gender labor-market differences in the model with no hiring costs are driven by differences in preference and skills only; the difference in the labor-market outcomes gender gap between the no-hiring-costs scenario and the observed data is therefore due to labor-market frictions. We then compare the outcomes of the model with symmetric information to the observed outcomes in the data and quantify the extent to which the gender gap is due to statistical discrimination. These results are presented in Figure 4 and Tables 7 and 8.

Figure 4 reveals that, on average, 70% of the earnings gap is due to hiring costs and about 48% is due to statistical discrimination in professional occupations. This is because, in the regime with no hiring cost and under the symmetric information regime, the earnings gap would have been 12% and 19%, respectively, instead of the 40% our model predicts.[9] Repeating the exercise for nonprofessional occupations reveals that 44% of the earnings gap is due to hiring costs and about 13% is due to statistical discrimination. Tables 7 and 8 show that, under both counterfactual regimes, women would have participated less in the labor force, but those who participate would have worked more hours. These results suggest that the asymmetric information affects the selection of women into the labor market; on the margin, women with higher costs of participation participate more. This is because participation provides a signal that the woman is more "attached." That is, the returns to labor-market participation are larger due to the reputation effect. Thus, the decline in the earnings gap relative to the symmetric information and frictionless-markets regimes is also due to the selection of women with either higher productivity or lower costs of participation into the labor markets.

What Caused the Changes in the Gender Earnings Gap? In order to examine the sources of the decline in the gender earnings gap over time, we decompose it into the following factors: human capital, beliefs and other (i.e. unobserved productivity, and age–education cohort composition). This is done using the estimated earnings equation. To quantify the effect of each of these factors, we hold it at the 1972:1978 median levels, allowing all other factors to change. Figure 5 reveals that the decline in the gender gap in human capital almost entirely explains the decline

[9] Figure 4 shows that under no-hiring cost regime the average median female to male earnings ratio is 88%. The gap is therefore 12%. The average predicted earnings gap over these three time periods is 40%. Therefore, the gap associated with labor-market frictions is 28/40. The other calculations are conducted in the same way.

in the gender earnings gap; the gender earnings gap would not have declined much if the human capital gap remained at its 1970s level. Changes in beliefs had a smaller effect, and the exogenous factors (other) had the smallest effect on the observed decline in the gender earnings gap. Since human capital is endogenous in the model, it is affected by discrimination. We therefore use our model to quantify the effect of the different factors on the change in human capital.

We face the standard problems of solving games with multiple equilibria under counterfactual regimes: First, it is not clear that one can find all the possible equilibria, and second, it may not be possible to select which equilibrium will be played. Fortunately, we can get around this problem because when there are no hiring costs the model has a unique equilibrium. When there are no hiring costs, wages are the marginal productivity, and therefore, all the changes in labor-supply decisions are driven by changes in demographics, participation cost shocks, and productivity shocks. We can therefore quantify the effect of changes in these factors on changes in human capital by holding each factor at the 1970s levels and allowing all other factors to change. Human capital is measured using its value in production according to the index: $b_{\tau 1} h_{nt} + b_{\tau 2} h_{nt}^2 + \sum_{r=1}^{\rho} b_{\tau 3r} h_{nt-r} + \sum_{r=1}^{\rho} b_{\tau 4r} d_{nt-r}$. We then calculate the gender gap in this index, and decompose its changes over the three time periods.

Similarly, to examine the effect of changes in beliefs over time, we can solve the model under the symmetric information regime. Performing this counterfactual analysis allows us to quantify the effect of changes in beliefs on human capital. Changes in beliefs over time and cohorts capture the effect of social and norm changes on the gender earnings gap and labor-market participation; our analysis, therefore, complements the work by Fernández (2007), Fogli and Veldkamp (2007) and Fernández and Fogli (2009), which models explicitly the causes of these changes. The results of this decomposition are presented in Table 9. They reveal that labor-market frictions (hiring cost) and asymmetric information amplify the effect of gender differences in preferences and skills, and are jointly responsible for roughly 50% of the change in human capital in both occupations and time periods; hiring cost accounts for the majority of this change. Social changes, through their effects on beliefs, account for roughly 13% of the change. The other major source of change is demographics, including changes in education, fertility and marriage, and accounts for about 30% of the decline in the human capital gender gap. Finally, the effect of productivity shocks is more important in professional than in nonprofessional occupations. Overall, we find that the demand-side factors played a greater role than the supply-side factors in the decline of the labor-market outcomes gender gaps and greater than found in the literature (see Lee and Wolpin, 2010, for example); the main reason for this difference is that we account for the endogenous effect of discrimination.

7 Conclusion

This paper finds that the most important factor affecting the gender earnings gap is the gender differences in human capital accumulated in both the labor market and at home. The paper quantifies the contribution of different factors that affected the gender earnings gap. We find that while differences in preferences and skills are important, asymmetric information has a large effect. In addition, we find that changes in discrimination patterns between the 70s and the 90s had a large effect on the decline in the gender gap in human capital, and hence on the gender earnings gap.

Because our model nests the symmetric information case, we propose a test for the existence of asymmetric information. There is a large literature on testing for asymmetric information and adverse selection, and there is no consensus regarding the importance of it in the labor markets or in other markets. While some papers find the existence of private information to be statistically significant, it is not always quantitatively important. Our test is statistically significant, and we are able to show that asymmetric information is quantitatively important to the existence of the gender earnings gap.

One source of the changes in the earnings gap has been the decline in the education gender gap. While education is exogenous in our model, we find evidence suggesting it is an important issue to be further explored. Importantly, the paper finds evidence that educated males earn more than females after accounting for statistical discrimination and human capital differences. This suggests that education choice and education signaling may have a role in the observed gender earnings gap and the changing patterns in the education gap. This is beyond the scope of this paper and is left for future research.

A Theoretical Results Proofs

Proof of Proposition 3.1. We establish optimally of the players' strategies below. We begin with a description of the off-equilibrium-path beliefs, and assign belief of a type (which is equivalent to the private information vector of characteristics) to choices which are not optimal and are not observed on the equilibrium path. ■

Assumption A.1 *Let $\overline{\omega}_{\tau t}(\omega_t^*)$ be the type with the lowest costs of working and highest returns (both current and future expected) given the observable characteristics ω_t^*; this type will work the most hours in occupation τ. Similarly, define $\underline{\omega}_{\tau t}(\omega_t^*)$ to be the type with the highest disutility from hours worked and lowest returns given the observable characteristics ω_t^* (this type also works the smallest fraction of time in the occupation). Let $\overline{\mu}_\tau = \mu_{it}(\overline{\omega}_{\tau t} \mid \omega_t^*)$ be the beliefs (probability assigned to this type by employer) about the type $\overline{\omega}_{\tau t}(\omega_t^*)$, and $\underline{\mu}_\tau = \mu_{it}(\underline{\omega}_{\tau t} \mid \omega_t^*)$ be the beliefs about the type $\underline{\omega}_{\tau t}(\omega_t^*)$. Suppose that $\forall \omega_t$, $\sigma_{it}(a_t \mid \omega_t) = 0$, then $\mu_{it}(\omega_t \mid \omega_t^*) = \underline{\mu}_\tau$ if $h_t < \underline{h}_{\tau t}$ and $\mu_{it}(\omega_t \mid \omega_t^*) = \overline{\mu}_\tau$ if $h_t > \overline{h}_{t\tau}$.*[10]

Lemma A.1 *Given any choice of hours, a worker accepts the contract with the highest salary. An employed worker remains with the current employer if there is a tie. If not employed, the worker randomizes between identical offers.*

Proof of Lemma A.1. An increase in salary enters the value function through the Frisch demand

[10] If the optimal hours have full support for any given observable characteristics, then there is no off-equilibrium-path analysis for hours.

for consumption:

$$(39)\quad V_{ki}(\omega_t)+\varepsilon_{kt}\equiv \max_{\{h_s;\{I_{\tau s}\}_{\tau\in\{\mathrm{P,NP}\}}\}_{s=t}^{T}} E_t\left\{\sum_{s=t}^{T}\beta^{s-t}\left[d_s u_{i0}(z_s,\zeta_s)+u_{i1}(l_s,z_s)\right.\right.$$
$$\left.\left.+\,\eta\lambda_s\sum_{\tau\in\{\mathrm{P,NP}\}} I_{\tau s}S_{i\tau s}(h_s,\omega_s)+d_s\varepsilon_{1s}+\ (1-d_s)\varepsilon_{0s}\right]\middle|\,\omega_t,d_t=k\right\}$$

Since $\eta\lambda_t>0$, the current utility is increasing in $S_{i\tau t}(h_t;\omega_t^*)$. We need to show that given any beliefs and any h_t, the continuation value of the worker is non-decreasing in salary. To see that, recall that given the hours choice, one occupation is chosen (by assumption). Changing employers within an occupation with the same hours worked does not change the beliefs. Furthermore, we assume that salary is not observed by outside employers; hence, it is not part of employment history and does not affect beliefs. Therefore, the continuation value is non-decreasing in salary, and accepting the highest salary given hours is optimal. ■

Proof of Proposition 3.1 (1). We begin with the workers' equilibrium strategies. Notice that the decision-theoretic solution to optimal consumption, c_t^0, is an optimal response given the contracts offered on the equilibrium path by construction. It is also the optimal consumption behavior off the equilibrium path. An optimal consumption strategy response to a one-period unanticipated salary shock is also an optimal response to a single deviation by employer. Thus, the optimal consumption plan in equation (3) is optimal.

Next we show that the labor-supply decisions, a_t^0, are optimal. Using the Bellman principle, the ex-ante value function for an individual who chooses to participate in the labor force in period t and behaves optimally thereafter is given by equation (13). Equations (13)–(19) describe the necessary conditions for optimality on-the-equilibrium-path once the salary schedule in the symmetric-information case in equation (12) is replaced with the optimal salary schedule in the asymmetric-information case described in equation (23).

Since this is a perfect Bayesian equilibrium, we need to show optimality of the strategies off the equilibrium path. By Assumption A.1, workers who work fewer (more) hours than the minimal (maximal) hours are compensated according to the beliefs component attached to the marginal type who works the least (most) hours, $\underline{\omega}_{\tau t}(\omega_t^*)$ $(\overline{\omega}_{\tau t}(\omega_t^*))$. By construction, the costs and benefits from working more than $\overline{h}_{\tau t}(\omega_t^*)$ or less than $\underline{h}_{\tau t}(\omega_t^*)$ are suboptimal given the on-equilibrium-path strategies and beliefs; because the beliefs component will not increase payoffs beyond working $\overline{h}_\tau$ or $\underline{h}_\tau$, such deviations cannot strictly increase payoffs.

Next, we show that the contract described in equation (23) satisfies the zero-profit condition and is optimal. Equation (23) is derived from the zero-profit condition using backward induction. We derive it the same way we derive equations (7)–(12) for the symmetric-information case; the only difference is that the employer's information set is now ω_t^* instead of ω_t.

In order to establish optimality, we need to show that given that other firms offering the competitive rate, the worker's strategy, and the firm's beliefs, there is no single unilateral profitable deviation from the competitive rate that strictly increases the expected profits. First, we show that by offering a lower salary, the firm cannot increase its profit. From Lemma A.1, workers accept the highest offer; thus, a deviation to a lower salary implies the worker rejects the offer and the payoff is zero. Notice that by assumption firms cannot credibly commit to pay above the market rate in the future.

Consider a firm offering a salary $\widetilde{S}_{i\tau t}$ for h_t such that $\widetilde{S}_{i\tau t}>S_{i\tau t}^0(h_t;\omega_t^*)$. The worker's state variables, ω_t, are not a function of past salaries. Therefore, at $t+1$, the worker's state variable remains ω_{t+1} and competing firms offer $s_{i\tau t+1}^0(h_{t+1};\omega_{t+1}^*)$; therefore, $Q(V_{1it+1}(\omega_{t+1})-V_{0it+1}(\omega_{t+1}))$,

$I^0_{i\tau t+1}(\omega_{t+1})$ and $h_{it+1}(\omega_{t+1})$ also remain the same. To see that the current salary does not change future labor-supply decisions note that the function $Q()$ is only a function of $(\varepsilon_{0t}, \varepsilon_{1t})$ by Lemma 1 of Hotz and Miller (1993). Hence, it is not affected by current salary. Also, because of the additive separability of leisure and consumption, the complete-assets–markets assumption and the assumption that the workers move first (announcing hours), the functions $V_{1it+1}()$ and $V_{0it+1}()$ remain the same.

Given that past salaries are not observed by outside employers, the beliefs, $\mu_{it+1}(\omega_{t+1} \mid \omega^*_{t+1})$, are unchanged. Therefore the probability of participation next period remains unchanged:

$$\widetilde{p}_{i\tau t+1}(h_t, \omega^*_t \mid \widetilde{s}_{\tau t}) = \int Q(\omega_{t+1}) I^0_{i\tau}(\omega_{t+1}) f_{i0}(\omega_{t+1} \mid \omega_t, a_t) \widetilde{\mu}_{it}(\omega_t \mid \omega^*_t, a_t) \, \mathrm{d}\omega_{t+1} = \widetilde{p}_{i\tau t+1}(h_t, \omega^*_t)$$

Since we only need to check for a single deviation (after which the employer follows the optimal salary schedule), as established in equation (23), the continuation expected profit can be written as

$$\begin{aligned}
\pi(\widetilde{S}_{\tau t}; \omega^*_t) = {} & Y_\tau(h_s, z^{\mathcal{P}}_s, K_{\tau s}) - \widetilde{S}_{\tau t}(h_s; \omega_s) - \gamma_\tau \\
& + \sum_{s=t+1}^{T} \beta^{s-t} E_t \left[Y_\tau(h_s, z^{\mathcal{P}}_s, K_{\tau s}) - S^0_{i\tau s}(h_s; \omega_s) \mid h_t, \omega_t \right] \\
& < Y_\tau(h_s, z^{\mathcal{P}}_s, K_{\tau s}) - S^0_{i\tau s}(h_s; \omega_s) - \gamma_\tau \\
& + \sum_{s=t+1}^{T} \beta^{s-t} E_t \left[Y_\tau(h_s, z^{\mathcal{P}}_s, K_{\tau s}) - S^0_{i\tau s}(h_s; \omega_s) \mid h_t, \omega_t \right] = 0.
\end{aligned}$$

Hence, there is no unilateral profitable deviation from the competitive salary schedule. ■

Proof of Proposition 3.1 (2). The optimality of the strategies described in the first part of the proposition is only a subset of the necessary conditions for existence of equilibrium. Next, we establish a set of (further) necessary and sufficient conditions for existence of equilibrium in our model. Note that the optimal hours worked and participation is a function of the firms' beliefs about the next period's participation. To see this, consider the beliefs about period T's participation:

$$\widetilde{p}_{i\tau T} = \int Q(\omega_T) I^0_{i\tau}(\omega_T) f_{i0}(\omega_T \mid \omega_{T-1}, a_{T-1}) \widetilde{\mu}_{iT-1}(\omega_{T-1} \mid \omega^*_{T-1}, a_{T-1}) \, d\omega_T. \tag{40}$$

Recall that ω_{T-1} and hence $\widetilde{p}_{i\tau T}$ are a function of the labor-market history $a_{0,\dots} a_{T-1}$. However a^0_{T-1}, the equilibrium labor-market decisions, are a function of the sequence $\widetilde{p}_{i\tau T}, \dots, \widetilde{p}_{i\tau 2}$. Therefore, $\widetilde{p}_{i\tau T}$ is defined as an implicit function of itself. In fact, there is a triangular system of implicit equilibrium beliefs of the following form.

$$\begin{aligned}
\widetilde{p}_{i\tau T} &= \Gamma_{iT}(\widetilde{p}_{i\tau T}, \dots, \widetilde{p}_{i\tau 2}) \\
\widetilde{p}_{i\tau T-1} &= \Gamma_{iT-1}(\widetilde{p}_{i\tau T-1}, \dots, \widetilde{p}_{i\tau 2}) \\
&\vdots \\
\widetilde{p}_{i\tau 2} &= \Gamma_{i2}(\widetilde{p}_{i\tau 2}),
\end{aligned} \tag{41}$$

where Γ_{it} is the RHS of equation (24). ■

Corollary A.1 *A necessary and sufficient condition for existence of equilibrium is that there exist a fixed point in* $\{\widetilde{p}_{i\tau 2}, \dots, \widetilde{p}_{i\tau T}\}^{\tau \in \Upsilon}_{i \in \{w,m\}}$ *of the system of equations in equation (41) for all* τ *and* i.

Proof of Corollary A.1. In order to prove this, we first show necessity. Suppose there exists an equilibrium in which equation (41) does not have a fixed point. Then take any t, τ, and i. The probability of a worker remaining in the firm at $t+1$ is either higher or lower than $\widetilde{p}_{i\tau t+1}$. By equation (23) and because on the equilibrium path the beliefs are consistent, the zero-expected-profit condition holds. Since $\widetilde{p}_{i\tau t+1}$ is not equal to the probability of next-period participation, the zero-profit condition is violated. Hence, this state cannot constitute an equilibrium.

Next we show sufficiency. Suppose equation (41) has a fixed point. Then, for any t, τ, and i by part 1 of the proposition, the competitive salary schedule exists. Given the competitive salary schedule, the worker's strategies for hours, participation, occupation, and consumption exist and are unique. Hence conditions 1, 2, and 3 of Definition 3.1 hold (mutual best responses by construction, the beliefs satisfy Bayes' law). ■

Existence of equilibrium in our model is established by showing that there exists a fixed point to the system of equations in (41).

Proof of Existence. Given the triangular nature of the system of equations in (41), it is sufficient to show existence for each equation in its own variable.

Existence of a solution to the worker's consumption and hours problem follows immediately from continuity and strict concavity of the utility function, and the fact that there is a solution to the worker's problem for any set of contracts offered. Next, note that for any period t, occupation τ, and gender i, $\widetilde{p}_{i\tau t+1}$ is the solution to

$$\begin{aligned}\widetilde{p}_{i\tau t+1} \quad = \quad & \int_{\omega_t}\left\{\int_{\omega_{t+1}} f_{i0}(\omega_{t+1}\mid\omega_{t+1})Q(V_{1it+1}(\omega_{t+1})-V_{0it+1}(\omega_{t+1}))\right.\\ & \left. I_{i\tau t+1}(\underline{h}_{\tau t+1}<h^*_{it+1}(\omega_{t+1})<\overline{h}_{\tau t+1})\,\mathrm{d}\omega_{t+1}\right\}\widetilde{\mu}_{it}(\omega_t\mid\omega_t^*,a_t)\,\mathrm{d}\omega_t. \qquad (42)\end{aligned}$$

Here we only make explicit the arguments of interest. Note that $\widetilde{p}_{i\tau t+1}:[0,1]$, and that the left-hand side is also defined on the compact interval $[0,1]$. Hence by Brouwer's fixed-point theorem, continuity of the RHS in $\widetilde{p}_{i\tau t+1}$ suffices to guarantee a solution to each one of the equations separately.

To show continuity, recall that $\underline{\omega}_{\tau t+1}$ is the *marginal type* for which $h^*_{\tau t+1}(\omega_{t+1})\equiv\underline{h}_{\tau t+1}$, and $\overline{\omega}_{\tau t+1}$ is the type for which $h^*_{\tau t+1}(\omega_{t+1})\equiv\overline{h}_{\tau t+1}$. Note that $h^*_{it+1}(\omega_{t+1})$ is continuous and invertible in ω_{t+1} as the utility function is continuous and differentiable. Thus, we can write, $\underline{h}^{-1}_{\tau t+1}(\omega_{t+1})$ $=\underline{\omega}_{\tau t+1}$ and $\overline{\omega}_{\tau t+1}\equiv\overline{h}^{-1}_{\tau t+1}(\omega_{t+1})$. Since $I_{i\tau t+1}(.)$ is an indictor function, we can rewrite the inner integral as

$$\int_{\underline{\omega}_{\tau t+1}}^{\overline{\omega}_{\tau t}} f_{i0}(\omega_{t+1}\mid\omega_{t+1})Q_{t+1}()\,\mathrm{d}\omega_{t+1}.$$

Since $h_t(\widetilde{p}_{i\tau t+1})$ is continuous in $\widetilde{p}_{i\tau t+1}$ and $Q(h_t(\widetilde{p}_{i\tau t+1}),.)$ is continuous in h_t, we only need to show that the functions $\underline{h}^{-1}_{\tau t+1}$ and $\overline{h}^{-1}_{\tau t+1}$ are continuous in $\widetilde{p}_{i\tau t+1}$. From the continuity of the production function in each occupation in all factors of production, $\underline{h}_{\tau t+1}$ and $\overline{h}_{\tau t+1}$ are continuous in h_t and $h_t(\widetilde{p}_{i\tau t+1},.)$ is continuous in $\widetilde{p}_{i\tau t+1}$. Hence, their inverses are continuous in $\widetilde{p}_{i\tau t+1}$. Therefore, there exists a solution to every period's beliefs separately.

The fact that there exists a one-to-one mapping between the posterior beliefs and the implied participation probability in equation (42) comes directly from the equilibrium requirement that the beliefs satisfy Bayes' rule; this condition holds by construction. Therefore, the expected profit condition on salary is also correct on the equilibrium path. ■

B Identification

B.1 Derivation of the Non-market-Hours Moment Condition

Combining equations (15), (30), (31) and (32) with the ex-ante valuation function (13) allows us to write the ex-ante equilibrium value function for any initial state ω:

$$\begin{aligned} V_{1i}(\omega_t) = \zeta_t &+ \sum_{r=1}^{2} \kappa_{ir} d_{t-r} + x_t' B_{i1} + x_{t'} l_t B_{i2} + \theta_{i0} l_t^2 + \sum_{r=1}^{2} \theta_{ir} l_t l_{t-r} \\ &+ \eta\lambda_t \sum_{\tau \in \{\text{P,NP}\}} I_{\tau t} S_{i\tau}(h_t, \omega_t^*) + E_t \left\{ \sum_{s=1}^{3} \beta^s \left[x_{t'} B_{i2} + \theta_{i0} + \sum_{r=1}^{2} \theta_{ir} l_{t+s-r}^{(1)} + \frac{\xi}{\sigma} - \sigma \ln\left(1 - p_{1it}^{(s)}\right) \right] \right\} \\ &+ \beta^4 E_t \left\{ V_{0i}(\omega_{t+4}) + \frac{\xi}{\sigma} - \sigma \ln\left(1 - p_{1it}^{(4)}\right) \right\} \end{aligned}$$

and

$$\begin{aligned} V_{0i}(\omega_t) = x_{t'} B_{i2} + \theta_{i0} + \sum_{s=1}^{2} \theta_{is} l_{t-s} + E_t \left\{ \sum_{s=1}^{3} \beta^s \left[x_{t'} B_{i2} + \theta_{i0} + \sum_{r=1}^{2} \theta_{is} l_{t+s-r}^{(0)} + \frac{\xi}{\sigma} - \sigma \ln\left(1 - p_{0it}^{(s)}\right) \right] \right\} \\ + \beta^4 E_t \left\{ V_{0i}(\omega_{t+4}) + \frac{\xi}{\sigma} - \sigma \ln\left(1 - p_{0it}^{(4)}\right) \right\}. \end{aligned}$$

A proof of this representation can be found in Altug and Miller (1998).

Next, we characterize, using the above, the necessary conditions for equilibrium (participation and hours). First, we characterize the equilibrium relationship from (15). Substituting the above ex-ante function representation into (15) and using (30) gives

$$\begin{aligned} (43) \quad \sigma \ln\left(\frac{p_{it}}{1 - p_{it}}\right) \equiv \eta^o \lambda_t^o \sum_{\tau \in \{\text{P,NP}\}} I_{\tau t} S_{i\tau t}^o + \sigma E_t \left[\sum_{s=1}^{3} \beta^s \ln\left(\frac{1 - p_{i1t}^{(s)}}{1 - p_{i0t}^{(s)}} \right) \right] + \zeta_t \\ + \sum_{s=1}^{2} \kappa_{is} d_{t-s} + x_t' B_{i1} - x_t' h_t B_{i2} - \theta_{0i}\left(1 - l_t^2\right) - E_t \left[\sum_{s=1}^{2} \theta_{si} h_t (l_{t-s} + \beta^s) \right]. \end{aligned}$$

Note that all the elements from period 4 onward are the same irrespective of whether action 1 or 0 is taken today by stochastic finite state dependence. Hence, they fall out of the above equation and we get equation (34). Similarly, using above ex-ante function representation, the necessary condition for equilibrium hours can be rewritten as

$$(44) \quad \eta^o \lambda_t^o \sum_{\tau \in \{\text{P,NP}\}} I_{\tau t} \frac{\partial S_{i\tau t}^o}{\partial h_t} + \sigma E_t \left[\sum_{s=1}^{3} \beta^s \left(1 - p_{i1t}^{(s)}\right)^{-1} \frac{\partial p_{i1t}^{(s)}}{\partial h_t} \right] - z_t' B_{i2} - 2\theta_{i0} l_t + E_t \left[\sum_{s=1}^{2} \theta_{si} (l_{t-s} + \beta^s) \right] = 0.$$

Note that, again by stochastic finite state dependence, all the elements from period 4 onward fall out of the above equations and this gives us equation (35).

B.2 Salary Equation Identification

Consider the optimal salary under asymmetric information.[11] Combining equations (23) and (29),

$$(45) \quad S_\tau(h_{nt}, \omega^*_{nt}) = K_{\tau t} + b_{\tau 1} h_{nt} + b_{\tau 2} h^2_{nt} + \textstyle\sum_{r=1}^{\rho} b_{\tau 3r} h_{nt-r} \sum_{r=1}^{\rho} b_{\tau 4r} d_{nt-r} + b_{\tau 5} age_{nt} \\ + b_{\tau 6} age^2_{nt} + b_{\tau 7} age_{nt} \times education_n + \nu_n - \gamma_\tau + \beta\gamma_\tau \widetilde{p}_{i\tau t+1}(h_{nt}, \omega^*_{nt}).$$

Integrating out ν_n gives us

$$(46) \quad E[S_{i\tau}(h_{nt}, \omega^*_{nt}) \mid h_{nt}, \{\omega^*_{nt} \backslash \nu_n\}] = K_{\tau t} + b_{\tau 1} h_{nt} + b_{\tau 2} h^2_{nt} + \sum_{r=1}^{\rho} b_{\tau 3r} h_{nt-r} + \sum_{r=1}^{\rho} b_{\tau 4r} d_{nt-r} \\ + b_{\tau 5} age_{nt} + b_{\tau 6} age^2_{nt} + b_{\tau 7} age_{nt} \times education_n + E[\nu_n \mid i, education_n] \\ - \gamma_\tau + \beta\gamma_\tau E[\widetilde{p}_{i\tau t+1}(h_{nt}, \omega^*_{nt}) \mid h_{nt}, \{\omega^*_{nt} \backslash \nu_n\}].$$

where $\{\omega^*_{nt} \backslash \nu_n\}$ is the set of all public variables except ν_n. Let $\overline{\nu}_n = E[\nu_n \mid gender, education_n]$, $S_{i\tau}(h_{nt}, \{\omega^*_{nt} \backslash \nu_n\}) = E[S_{i\tau}(h_{nt}, \omega^*_{nt}) \mid h_{nt}, \{\omega^*_{nt} \backslash \nu_n\}]$ and $\widetilde{p}_{i\tau t+1}(h_{nt}, \{\omega^*_{nt} \backslash \nu_n\}) = E[\widetilde{p}_{i\tau t+1}(h_{nt}, \omega^*_{nt}) \mid h_{nt}, \{\omega^*_{nt} \backslash \nu_n\}]$. We now get

$$(47) \quad S_{i\tau}(h_{nt}, \{\omega^*_{nt} \backslash \nu_n\}) = K_{\tau t} + b_{\tau 1} h_{nt} + b_{\tau 2} h^2_{nt} + \textstyle\sum_{r=1}^{\rho} b_{\tau 3r} h_{nt-r} \sum_{r=1}^{\rho} b_{\tau 4r} d_{nt-r} + b_{\tau 5} age_{nt} \\ + b_{\tau 6} age^2_{nt} + b_{\tau 7} age_{nt} \times education_n + \overline{\nu}_n - \gamma_\tau + \beta\gamma_\tau \widetilde{p}_{i\tau t+1}(h_{nt}, \{\omega^*_{nt} \backslash \nu_n\}).$$

Since both $S_\tau(h_{nt}, \{\omega^*_{nt} \backslash \nu_n\})$ and $\widetilde{p}_{i\tau t+1}(h_{nt}, \{\omega^*_{nt} \backslash \nu_n\})$ are identified from data, equation (47) is a standard panel-data model with unobserved effect $\overline{\nu}_n$. Thus, under standard assumptions, $(K_{\tau t}, b_{\tau 1}, ...b_{\tau 7}, \gamma_\tau, \beta\gamma_\tau)$ are identified. Integrating equation (45) over all the information except ν_n gives

$$(48) \quad E[S_{in\tau} \mid \nu_n] = E[K_{\tau t}] + b_{\tau 1} E[h_{nt} \mid \nu_n] + b_{\tau 2} E[h^2_{nt} \mid \nu_n] + \textstyle\sum_{r=1}^{\rho} b_{\tau 3r} E[h_{nt-r} \mid \nu_n] \\ + \textstyle\sum_{r=1}^{\rho} b_{\tau 4r} E[d_{nt-r} \mid \nu_n] + b_{\tau 5} E[age_{nt} \mid \nu_n] + b_{\tau 6} E[age^2_{nt} \mid \nu_n] \\ + b_{\tau 7} E[age_{nt} \times education_n \mid \nu_n] + \nu_n - \gamma_\tau + \beta\gamma_\tau E[\widetilde{p}_{i\tau t+1} \mid \nu_n]$$

Note that each of the components that are conditional on ν_n is identified by looking at the same individual over time. Therefore, ν_n is identified.

C Data Description

We used data from the Family File, the Childbirth and Adoption History File, the Retrospective Occupation File, and the Marriage History File of the PSID. The Family File contains a separate record for each member of each household included in the survey in a given year, but includes labor income, hours worked, and years of completed education only for Heads and Wives. The Childbirth and Adoption History File contains information collected in the 1985–2007 waves of the PSID regarding histories of childbirth and adoption. The file contains details about childbirth and adoption events of eligible people living in a PSID family at the time of the interview in any wave from 1985 through 2007. Each set of records for a specified individual contains all known cumulative data about the timing and circumstances of his/her childbirth or adoption experience up to and including 2007, or those waves during that period when the individual was in a responding

[11] All the arguments remain the same in the symmetric-information case.

family unit. If an individual has never had any children, one record indicates that report. Note that *eligible* refers individuals of childbearing age in responding families. Similarly, the 1985–2007 Marriage History file contains retrospective histories of marriages for individuals of marriage-eligible age living in PSID families between 1985 and 2007. Each set of records for a specified individual contains all known cumulative data about the timing and circumstances of his/her marriages up to and including 2007, or those waves during that period when the individual was in a responding family unit.

Our sample selection started from the Childbirth and Adoption History File. We then drop any individual who was in the survey for four years or less. We then drop all individuals who were older than 65 in 1967, and then drop all individuals that were less than 25 years old in 2007. We then drop all individuals who were neither Head nor Wife in our sample for at least four years.

There were coding errors for the different measures of consumption in the PSID from which we construct our consumption measure. In particular, our measure of food consumption expenditures for a given year is obtained by summing the values of annual food expenditures for meals at home, annual food expenditures for eating out, and the value of food stamps received for the year. We measured consumption expenditures for year t by taking 0.25 times the value of this variable for the year $t-1$ and 0.75 times its value for the year t; this step accounts for the fact that the survey questions used to elicit information about household food consumption were asked sometime in the first half of the year, while the response is dated in the previous year.

The variables used in the construction of the measure for total expenditures are also subject to the problem of truncation described above because the way they are coded in the 1983 PSID data tapes. The truncation value for the value of food stamps received for that year is \$999.00; the relevant value for this variable in the subsequent years and for the value of food consumed at home and eating out is \$9,999.00, however. We also use variables describing various demographic characteristics of the individuals in our sample. The dates of birth of the individuals were obtained from the Child Birth and Adoption file.

The race of the individual and the region of residence at the time of the interview were obtained from the Family portion of the data record. We defined the region variable as the geographical region in which the household resided at the time of the annual interview. This variable is not coded consistently across the years. For 1968 and 1969, the values 1 to 4 denote the regions Northeast, North central, South, and West. For 1970 and 1971, the values 5 and 6 denote the regions Alaska and Hawaii and a foreign country, respectively. After 1971 a value of 9 indicates missing data, but no person-years data were lost due to missing data for these variables. We also drop all observations of individuals coded as living in regions 5 and 6.

We used the family variable *Race of the Household Head* to code the race variable in our study. For the interviewing years 1968–1970, the values 1 to 3 denote White, Black, and Puerto Rican or Mexican, respectively; the value 7 denotes other (including Asian and Philippino), and 9 denotes missing data. For 1971 and 1972, the third category is redefined as Spanish-American or Cuban and between 1973 and 1984, just as Spanish American. After 1984, the variable was coded in such a way that 1–6 correspond to the categories White, Black, Hispanic, American Indian, Aleutian or Eskimo, and Asian or Pacific Islander, respectively. A value of 7 denotes the other category, a value of 9 denotes missing data. We used all available information for all the years to assign the race of the individual for years in the sample when that information was available. We then drop all individuals that were not coded as White.

The marital status of a woman in our subsample was determined from the Marriage History File. The number of individuals in the household and the total number of children within that household were also determined from the family-level variables of the same name. In 1968, a code for missing data (equal to 99) was allowed for the first variable, but in the other years, missing data were assigned. The second variable was truncated above the value of 9 for the interviewing years 1968 and 1971. After 1975, this variable denotes the actual number of children in the family unit. Household income was measured from the PSID variable, *Total Family Money Income*, which included taxable income of Head and Wife, total transfers of Head and Wife, taxable income of others in the family units and their total transfer payments.

We used the PSID Retrospective Occupation File to obtain a consistent three-digit occupational code for our sample. First we eliminated all self-employed, dual-employed, government workers, Farmers and Farm Managers, Farm Laborers and Farm Foremen, Armed Forces, and Private Household workers. The professional occupation is made up of the following classifications: Professional, Technical, and Kindred Worker; Managers and Administrators, Except Farm Managers; and some categories of Sales Workers. The Sales Workers included in professional occupations are Advertising and Salesmen, Insurance Agents Brokers and Underwriters, and Stock and Bond Salesmen. The nonprofessional occupation consists of the following classifications: Sales Workers (not included in Professional); Clerical and Kindred Workers; Craftsmen and Kindred workers; Operatives, Except Transport; Transport Equipment Operatives; Laborers, Except Farm; and Service Workers, Except Private Household.[12]

We used two different deflators to convert the nominal quantities such as average hourly earnings, household income, and so on to real values. First, we defined the (spot) price of food consumption to be the numeraire good at t in the theoretical section. We accordingly measured real food consumption expenditures and real wages as the ratio of the nominal consumption expenditures and wages, and the annual chain-type price deflator for food consumption expenditures published in Table t.12 of the National Income and Product Accounts. On the other hand, we deflated variables such as the nominal value of home ownership or nominal family income with the chain-type price deflator for total personal consumption expenditures.

References

[1] Aguirregabiria, V., and P. Mira (2007): “Sequential Estimation of Dynamic Discrete Games,” *Econometrica*, 75, 1–53.

[2] Albanesi, S., and C. Olivetti (2008): "Gender Roles and Medical Progress", Working Paper, Columbia University.

[3] Albanesi, S., and C. Olivetti (2009): "Home Production, Market Production and the Gender Wage Gap: Incentives and Expectations", *Review of Economic Dynamics*, 12, 80–107.

[4] Altonji, J. (2005): “Employer Learning, Statistical Discrimination and Occupational Attainment,” *American Economic Review*, 95(2), 112–117.

[5] Altonji, J. and M. Blank (1999): Race and gender in the labor market, in: O. Ashenfelter and D. Card ed., *Handbook of Labor Economics*, vol. 3, Elsevier, Amsterdam, 3143–3259.

[6] Altonji, J., F. Hayashi, and L. Kotlikoff (1996): “Risk Sharing Between and Within Families,” *Econometrica*, 64, 261–294.

[7] Altug, S., and P. Labadie (1994): *Dynamic Choice and Asset Markets*, Academic Press, San Diego, CA.

[12] See the PSID wave XIV—1981 documentation, Appendix 2: Industry and Occupation Codes for a detailed description of the classifications used in the paper.

[8] Altug, S., and R. A. Miller (1990): "Household Choices in Equilibrium," *Econometrica*, 58, 543–570.

[9] Altug, S., and R. A. Miller (1998): "The Effect of Work Experience on Female Wages and Labour Supply," *Review of Economic Studies*, 65, 45–85.

[10] Antonovics, K. (2004): *Statistical Discrimination and Intergenerational Income Mobility*, mimeo, University of California, San Diego.

[11] Arrow, K. (1972): "The Theory of Discrimination," in: O. A. Ashenfelter and A. Rees, eds., *Discrimination in Labor Markets*, Princeton University Press, Princeton, NJ, pp. 3–33.

[12] Attanasio, O., H. Low, and V. Sanchez-Marcos (2008): "Explaining Changes in Female Labor Supply in a Life-Cycle Model," *American Economic Review*, 98(4), 1517–1552.

[13] Bajari, P., L. Benkard, and J. Levin (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75(5), 1331–1370.

[14] Baron, J., D. Black, and M. Lowenstein (1993): "Gender Differences in Training, Capital and Wages," *Journal of Human Resources*, 28, 343–364.

[15] Becker, G. (1965): "A Theory of Allocation of Time," *Economic Journal*, 75, 496–517.

[16] Becker, G. (1971): *The Economics of Discrimination*, 2nd edition, The University of Chicago Press, Chicago, IL.

[17] Becker, G., M. Grossman and K. Murphy (1994): "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, 84(3), 396–418.

[18] Bowlus, A. J. (1997): "A Search Interpretation of Male–Female Wage Differentials," *Journal of Labor Economics*, 15(4), 625–657.

[19] Browning, M., and A. Lusardi (1996): "Household Saving: Micro Theories and Micro Facts," *Journal of Economic Literature*, 34(4), 1797–1855.

[20] Card, D. (1990): "Labour Supply With a Minimum Threshold," *Carnegie–Rochester Series on Public Policy*, 33, 137–168.

[21] Card, D., and J. DiNardo (2002). "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics*, 20, 733–783.

[22] Chamberlain, G. (1984). "Panel Data," in: Z. Griliches and M. D. Intriligator eds., *Handbook of Econometrics*, vol. 2. Amsterdam: North-Holland, pp. 1247–1318.

[23] Chesher, A. (2007): "Instrumental Values," *Journal of Econometrics*, 139, 15–34.

[24] Coate, S., and G. Loury (1993): "Will Affirmative Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, 83, 1220–1240.

[25] Dionne, G., and N. A. Doherty (1994): "Adverse Selection, Commitment, and Renegotiation: Extension to and Evidence from Insurance Markets," *Journal of Political Economy*, 102(2), 209–235.

[26] Eckstein, Z., and K. Wolpin (1989): "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience," *Review of Economic Studies*, 56, 375–390.

[27] Erosa, A., L. Fuster, and D. Restuccia (2005): *A Quantitative Theory of the Gender Gap in Wages*, Working Paper 05-09, Federal Reserve Bank of Richmond, September.

[28] Fernández R., (2007): "Culture as Learning: The Evolution of Female Labor-Force Participation over a Century," Working Paper, New York University.

[29] Fernández, R., and A. Fogli (2009): "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics*, 1, 146–177.

[30] Flabbi, L. (2009): "Prejudice and gender differentials in the US labor market in the last twenty years", Journal of Econometrics, 156(1), 190-200.

[31] Flabbi, L. (forthcoming): "Gender Discrimination Estimation in a Search Model with Matching and Bargaining", International Economic Review.

[32] Fogli, A. and L. Veldkamp, (2007): "Nature or Nurture? Learning and the Geography of Female Labor-Force Participation," NBER Working Papers 14097.

[33] Greenwood, J., A. Seshardri, and M. Yorukoglu (2005): "Engines of Liberation," *Review of Economic Studies*, 72(1), 109–133.

[34] Heckman, J. J., and T. E. MaCurdy (1980): "A Life-Cycle Model of Female Labour Supply," *Review of Economic Studies*, 47, 47–74.

[35] Hellwig, M. (1987): "Some Recent Developments in the Theory of Competition in Markets With Adverse Selection," *European Economic Review*, 31, 319–325.

[36] Hotz, J., F. Kydland and G. Sedlacek (1988): "Intertemporal Preferences and Labor Supply" *Econometrica*, 56(2), 335–360.

[37] Hotz, V. J., and R. A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 60, 497–529.

[38] Jones, L., R. Manuelli, and E. McGrattan (2003): *Why are Married Women Working so Much?* Report #317, Federal Reserve Bank, Minneapolis.

[39] Laffont, J. J., and J. Tirole, (1988): "The Dynamics of Incentive Contract," *Econometrica*, 56(5), 1153-1175.

[40] Lawrence, E. (1991): "Poverty and the Rate of Time Preference," *Journal of Political Economy*, 99(1), 54–77.

[41] Lee, D., and K. Wolpin (2010): "Accounting for Wage and Employment Changes in the U.S. from 1968–2000: A Dynamic Model of Labor-Market Equilibrium," *Journal of Econometrics*, 156, 68–85.

[42] Mace, B. (1991): "Full Insurance in the Presence of Aggregate Uncertainty," *Journal of Political Economy*, 99, 928–956.

[43] MaCurdy, T. E. (1981): "An Empirical Model of Labour Supply in a Life-Cycle Setting," *Journal of Political Economy*, 89, 1059–1085.

[44] Mincer, J., and S. W. Polachek (1974): "Family Investments In Human Capital: Earnings of Women," *Journal of Political Economy*, 82(2), 76–108.

[45] Moro, A. (2003): "The Effect of Statistical Discrimination on Black–White Wage Inequality: Estimating a Model With Multiple Equilibria," *International Economic Review*, 44, 467–500.

[46] Mundlak, Y. (1961): "Empirical production function free of management bias," *Journal of Farm Economics*, 43, 44–56.

[47] Mundlak, Y. (1978): "On the pooling of time series and cross-section data," *Econometrica*, 46, 69–85.

[48] Pagan, A., and P. Ullah (1999): *Nonparametric Econometrics*, Cambridge University Press, London.

[49] Pesendorfer, M., and P. Schmidt-Dengler (2008): "Asymptotic Least Squares Estimators for Dynamic Games," *Review of Economic Studies*, 75, 901–928.

[50] Phelps, E. (1972): "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62, 659–661.

[51] Polachek, S. W. (1981): "Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure," *Review of Economics and Statistics*, 63(1), 60–69.

[52] Rothschild, M., and J. Stiglitz (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, 90, 629–649.

[53] Rust, John (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033.

[54] Shea, J. (1995): "Union Contracts and the Life-Cycle Permanent Income Hypothesis," *American Economic Review*, 95(1), 186–200.

[55] Tirole, J. (1996): "A Theory of Collective Reputation," *Review of Economic Studies*, 63, 1–22.

[56] Townsend, R. (1994): "Risk and Insurance in Village India," *Econometrica*, 62, 539–591.

[57] Zeldes, S. P. (1989): "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy*, 97(2), 3005–46.

Table 1: Summary of Labor-Market and Human-Capital Variables

	Participation		Hours		Earnings		Fraction of Women		Education	
	Male	Female	Male	Female	Male	Female	Professional	Nonprofessional	Male	Female
1968	0.93	0.54	2,244	1,401	39.8	16.2	0.28	0.45	12.2	11.7
	(0.25)	(0.50)	(631)	(731)	(24.9)	(10.8)			(3.4)	(2.6)
1969	0.96	0.60	2,240	1,371	41.2	16.2	0.29	0.48	12.1	11.7
	(0.18)	(0.49)	(610)	(739)	(26.6)	(11.2)			(3.4)	(2.6)
1970	0.97	0.64	2,216	1,332	41.6	16.2	0.30	0.49	12.1	11.8
	(0.18)	(0.48)	(593)	(758)	(26.5)	(11.0)			(3.4)	(2.6)
1971	0.96	0.63	2,175	1,382	41.7	17.0	0.32	0.48	12.2	11.8
	(0.20)	(0.48)	(636)	(750)	(24.9)	(11.6)			(3.3)	(2.6)
1972	0.95	0.62	2,155	1,389	41.7	17.4	0.32	0.49	12.2	11.8
	(0.21)	(0.49)	(636)	(728)	(26.2)	(11.8)			(3.3)	(2.6)
1973	0.96	0.60	2,188	1,411	43.2	17.8	0.30	0.47	12.2	11.9
	(0.19)	(0.49)	(633)	(720)	(26.5)	(11.2)			(3.3)	(2.6)
1974	0.95	0.62	2,130	1,424	43.2	18.1	0.32	0.47	12.2	11.9
	(0.23)	(0.49)	(641)	(713)	(28.4)	(12.5)			(3.3)	(2.6)
1975	0.92	0.62	2,230	1,415	42.0	17.7	0.32	0.48	12.3	11.9
	(0.27)	(0.49)	(641)	(726)	(30.6)	(11.9)			(3.2)	(2.6)
1976	0.92	0.62	2,092	1,395	40.9	17.8	0.35	0.49	12.3	12.0
	(0.24)	(0.49)	(677)	(726)	(31.1)	(12.4)			(3.2)	(2.5)
1977	0.91	0.61	2,119	1,418	42.6	18.1	0.35	0.50	12.4	11.8
	(0.27)	(0.49)	(668)	(706)	(31.0)	(12.4)			(3.1)	(2.5)
1978	0.87	0.62	2,115	1,454	44.2	18.6	0.33	0.46	12.4	12.0
	(0.33)	(0.49)	(650)	(737)	(32.8)	(12.5)			(3.1)	(2.5)
1979	0.91	0.63	2,141	1,472	43.5	18.7	0.36	0.49	12.4	12.1
	(0.29)	(0.48)	(675)	(711)	(30.3)	(12.7)			(3.1)	(2.5)
1980	0.91	0.65	2,112	1,450	42.4	18.5	0.38	0.50	12.4	12.1
	(0.29)	(0.48)	(651)	(726)	(28.2)	(12.4)			(3.0)	(2.5)
1981	0.91	0.64	2,199	1,642	42.3	19.9	0.37	0.45	12.6	12.2
	(0.28)	(0.48)	(578)	(607)	(28.3)	(13.7)			(2.8)	(2.4)
1982	0.91	0.64	2,166	1,630	41.3	19.7	0.36	0.46	12.6	12.3
	(0.29)	(0.48)	(576)	(617)	(28.3)	(12.1)			(2.8)	(2.4)
1983	0.90	0.65	2,136	1,632	40.3	20.2	0.37	0.47	12.6	12.3
	(0.30)	(0.48)	(600)	(628)	(31.2)	(13.7)			(2.8)	(2.3)
1984	0.90	0.67	2,142	1,635	40.7	20.4	0.38	0.47	12.6	12.3
	(0.30)	(0.47)	(586)	(628)	(32.7)	(13.8)			(2.7)	(2.3)
1985	0.90	0.70	2,188	1,646	42.9	20.6	0.40	0.47	12.6	12.3
	(0.30)	(0.45)	(615)	(680)	(39.9)	(13.1)			(2.7)	(2.3)
1986	0.90	0.70	2,192	1,665	44.0	21.6	0.39	0.48	12.7	12.3
	(0.30)	(0.46)	(576)	(678)	(39.5)	(15.1)			(2.7)	(2.3)
1987	0.90	0.70	2,215	1,690	45.2	22.5	0.39	0.48	12.7	12.3
	(0.30)	(0.46)	(612)	(662)	(41.5)	(15.1)			(2.6)	(2.3)
1988	0.90	0.71	2,230	1,691	46.7	23.2	0.41	0.48	12.7	12.4
	(0.30)	(0.45)	(594)	(671)	(51.4)	(15.3)			(2.6)	(2.3)
1989	0.89	0.72	2,221	1,703	47.7	23.7	0.41	0.47	12.7	12.4
	(0.31)	(0.45)	(610)	(676)	(54.0)	(16.6)			(2.6)	(2.3)
1990	0.88	0.72	2,251	1,683	48.0	23.8	0.41	0.48	12.7	12.4
	(0.32)	(0.45)	(579)	(631)	(50.7)	(17.4)			(2.6)	(2.2)
1991	0.87	0.72	2,259	1,807	47.2	23.7	0.42	0.43	12.7	12.5
	(0.33)	(0.49)	(576)	(641)	(41.5)	(18.7)			(2.6)	(2.3)
1992	0.87	0.74	2,221	1,815	47.2	24.1	0.43	0.50	12.8	12.6
	(0.33)	(0.44)	(606)	(682)	(44.8)	(18.2)			(2.6)	(2.3)
1993	0.91	0.76	2,205	1,801	51.2	29.1	0.45	0.49	12.7	12.5
	(0.30)	(0.46)	(632)	(609)	(58.6)	(23.2)			(2.5)	2.3
1994	0.91	0.77	2,250	1,820	52.6	29.3	0.46	0.48	12.7	12.6
	(0.28)	(0.42)	(591)	(580)	(55.2)	(21.2)			(2.4)	(2.2)
1995	0.92	0.78	2,260	1,833	54.0	30.3	0.47	0.48	12.7	12.6
	(0.27)	(0.41)	(604)	(600)	(56.6)	(23.4)			(2.5)	(2.2)
1996	0.93	0.79	2,276	1,829	58.0	33.1	0.48	0.47	12.9	2.8
	(0.25)	(0.40)	(587)	(606)	(59.3)	(26.8)			(2.5)	(2.2)
1997	0.90	0.76	2,280	1,868	62.7	34.0	0.46	0.46	12.9	12.8
	(0.32)	(0.47)	(550)	(563)	(91.3)	(28.2)			(2.4)	(2.2)

Note: Standard deviation in parentheses. Earnings in thousands of year-2000 US$.

Table 2: Summary of Demographic and Wealth Variables

	Food Consumption	Family Size	Age	Number of Kids ≤ 5 years old	Number of Kids > 5 and < 17 years old	Marital Status
1968		4.0	37.8	0.56	0.94	0.85
		(1.9)	(10.7)	(0.82)	(1.3)	(0.35)
1969	7.7	4.0	38.5	0.53	0.93	0.86
	(3.7)	(1.9)	(10.9)	(0.83)	(1.3)	(0.35)
1970	7.7	3.8	38.6	0.49	0.87	0.85
	(3.6)	(1.8)	(11.3)	(0.79)	(1.3)	(0.36)
1971	7.5	3.7	39.0	0.44	0.85	0.83
	(3.5)	(1.8)	(11.6)	(0.76)	(1.2)	(0.37)
1972	7.4	3.7	39.3	0.42	0.80	0.82
	(3.5)	(1.8)	(11.8)	(0.72)	(1.2)	(0.38)
1973	7.4	3.6	39.5	0.39	0.77	0.82
	(3.4)	(1.8)	(12.2)	(0.69)	(1.1)	(0.38)
1974	7.3	3.4	39.8	0.37	0.71	0.82
	(3.4)	(1.8)	(12.5)	(0.68)	(1.1)	(0.39)
1975	6.9	3.3	39.8	0.35	0.66	0.81
	(3.2)	(1.7)	(12.6)	(0.67)	(1.0)	(0.39)
1976	6.8	3.3	39.7	0.35	0.62	0.80
	(3.2)	(1.7)	(12.6)	(0.68)	(1.0)	(0.39)
1977	6.7	3.2	39.7	0.34	0.60	0.79
	(3.3)	(1.6)	(12.6)	(0.68)	(0.96)	(0.40)
1978	6.5	3.2	38.9	0.41	0.51	0.77
	(3.5)	(1.6)	(12.7)	(0.72)	(0.87)	(0.42)
1979	6.7	3.1	39.8	0.34	0.53	0.77
	(3.3)	(1.5)	(12.5)	(0.66)	(0.88)	(0.42)
1980	6.6	3.1	39.9	0.35	0.50	0.78
	(3.3)	(1.5)	(12.5)	(0.69)	(0.84)	(0.42)
1981	6.4	3.1	38.8	0.39	0.50	0.80
	(3.1)	(1.4)	(11.9)	(0.69)	(0.82)	(0.39)
1982	6.3	3.1	38.9	0.38	0.50	0.80
	(3.1)	(1.4)	(11.8)	(0.69)	(0.82)	(0.40)
1983	6.3	3.1	39.0	0.38	0.51	0.80
	(3.1)	(1.4)	(11.8)	(0.67)	(0.84)	(0.40)
1984	6.3	3.1	39.1	0.38	0.51	0.80
	(3.1)	(1.4)	(11.7)	(0.70)	(0.84)	(0.40)
1985	6.5	3.1	39.6	0.37	0.53	0.80
	(3.8)	(1.4)	(11.5)	(0.68)	(0.85)	(0.40)
1986	6.4	3.1	40.2	0.37	0.56	0.81
	(3.2)	(1.4)	(11.2)	(0.69)	(0.87)	(0.40)
1987	6.5	3.1	40.6	0.36	0.57	0.81
	(3.1)	(1.3)	(10.8)	(0.67)	(0.87)	(0.40)
1988	6.6	3.1	41.4	0.34	0.58	0.81
	(3.0)	(1.3)	(10.6)	(0.66)	(0.87)	(0.40)
1989	6.5	3.1	42.1	0.33	0.59	0.81
	(2.9)	(1.3)	(10.3)	(0.63)	(0.87)	(0.39)
1990	6.6	3.1	42.7	0.30	0.61	0.81
	(3.2)	(1.3)	(10.1)	(0.62)	(0.89)	(0.39)
1991	6.4	3.0	43.7	0.29	0.58	0.77
	(3.1)	(1.4)	(10.2)	(0.60)	(0.87)	(0.43)
1992	6.7	3.1	44.0	0.29	0.62	0.82
	(3.8)	(1.3)	(9.6)	(0.61)	(0.89)	(0.39)
1993	6.5	3.1	43.0	0.30	0.54	0.76
	(3.2)	(1.5)	(11.0)	(0.62)	(0.85)	(0.43)
1994	7.1	3.1	42.0	0.29	0.54	0.74
	(6.0)	(1.4)	(9.9)	(0.61)	(0.84))	(0.44)
1995	6.3	3.1	43.0	0.27	0.54	0.74
	(8.8)	(1.4)	(11.0)	(0.61)	(0.84)	(0.44)
1996	6.2	3.0	43	0.25	0.54	0.74
	(9.2)	(1.4)	(11.0)	(0.57)	(0.84)	(0.44)
1997	7.0	3.0	44.0	0.22	0.53	0.75
	(11.0)	(1.4)	(10.0)	(0.53)	(0.84)	(0.43)

Note: Standard deviation in parentheses. Household Income and Food Consumption in thousands of year-2000 US$

Table 3: Panel Estimates. Dependent Variable: Log Household Annual Food Consumption, 1969–2007

Age	0.00447
	(0.0164)
Age Squared	0.000280***
	(2.15e−05)
Number of Individuals Living in the Household	0.226***
	(0.00367)
Number of Kids Between Age 6 and 14 in the Household	−0.136***
	(0.00565)
Number of Kids Between Age 0 and 5 in the Household	−0.181***
	(0.00573)
Northeast Regional Dummy	0.0842**
	(0.0415)
South Regional Dummy	−0.0846***
	(0.0306)
West Regional Dummy	−0.0779**
	(0.0322)
Constant	6.869***
	(0.648)
Individual-Specific Effects	Yes
Observations	424656
R-squared	0.524
Number of id	15702

Note: Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Data from PSID, each regression also includes yearly dummies.

Table 4: 2SLS Estimates of Earnings Equation. Dependent Variable: Average Weekly Earnings, 1970–1996

Variables	Prof	Nonprof
Hours Worked	20.23***	12.63***
	(1.200)	(0.469)
Hours Worked Squared	−0.156***	−0.0736***
	(0.0141)	(0.00640)
Hours Worked Lagged	2.645***	2.196***
	(0.292)	(0.107)
Hours Worked Twice Lagged	2.374***	1.440***
	(0.313)	(0.109)
Age	36.27**	10.72*
	(14.08)	(5.904)
Age Squared	−0.559***	−0.252***
	(0.0564)	(0.0170)
Age X Years of Completed Education	0.997**	1.328***
	(0.402)	(0.205)
Dummy = 1 if Worked Last Year	−16.59	−28.01***
	(14.06)	(6.995)
Dummy = 1 if Worked Two Years Ago	−50.66***	−15.84***
	(13.22)	(5.639)
Prob. of Working Next in Occupation	82.26***	58.00*
	(23.21)	(37.27)
Individual Specific Effects	Yes	Yes
Constant	−939.5*	−512.5***
	(480.8)	(197.2)
Observations	37618	81272
R-squared	0.258	0.284
Number of id	5600	10271

Note: Standard errors are in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Data from PSID, each regression also includes yearly dummies.

Table 5: OLS Estimates of Individual-Specific Effects from Earnings Equation. Dependent Variable: Fixed Effect

Variables	Prof	Nonprof
Male Dummy	−51.42***	−5.843
	(18.06)	(15.64)
Years of Completed Education	−150.3***	−165.7***
	(6.322)	(5.686)
Years of Completed Education Squared	5.426***	6.194***
	(0.233)	(0.220)
Years of Completed Education X Male Dummy	25.26***	17.09***
	(1.316)	(1.251)
Constant	816.2***	805.1***
	(42.24)	(36.29)
Observations	5600	10271
R-squared	0.138	0.106

Note: Standard errors are in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: GMM Estimates Utility Function. Moment Conditions: Hours Euler Equation and Labor-Market Participation Equations

$$u_{it0}(z_{nt},\zeta_t) = \zeta_t + \sum_{r=1}^{2} \kappa_{ir} d_{t-r} + x'_{nt} B_{i1};\; u_{i1}(z_{nt}, l_{nt}) = x'_{nt} l_{nt} B_{i2} + \theta_{i0} l_{nt}^2 + \sum_{s=1}^{2} \theta_{is} l_{nt} l_{nt-s}$$

Time Effect	(1)	Participation Cost	(2)	Marginal Utility of Leisure	(3)
1974	0.429***	Constant	−11.30***	Constant	0.671***
	(0.122)		(0.345)		(0.0128)
1975	−0.110	Individuals in Household	−0.0894***	Individuals in Household	$6.93e{-}05$
	(0.116)		(0.0211)		(0.000362)
1976	−0.00819	Kids Age 0 to 5	−1.308***	Kids Age 0 to 5	−0.000429
	(0.109)		(0.142)		(0.00220)
1977	0.147	X Male Dummy	0.128	X Male Dummy	−0.0126***
	(0.105)		(0.149)		(0.00228)
1978	−0.454***	Kids Age 6 to 14	0.110	Kids Age 6 to 14	0.00778***
	(0.104)		(0.0882)		(0.00116)
1979	−0.163*	X Male Dummy	−0.682***	X Male Dummy	−0.0164***
	(0.0989)		(0.0941)		(0.00125)
1980	0.0407	Age	0.734***	Age	0.000385
	(0.0986)		(0.0164)		(0.000262)
1981	−0.0723	Age Squared	−0.00893***	Age Squared	$-5.50e{-}06$*
	(0.0968)		(0.000196)		$(3.15e{-}06)$
1982	0.0101	Years of Education	0.163***	Years of Education	−0.00792***
	(0.0971)		(0.0198)		(0.000386)
1983	0.0203	X Male Dummy	0.224***	X Male Dummy	−0.00120***
	(0.0978)		(0.0252)		(0.000430)
1984	0.0228	Marital Status Dummy	−1.093***	Marital Status Dummy	−0.0692***
	(0.0967)		(0.326)		(0.00536)
1985	0.232**	X Male Dummy	−15.77***	X Male Dummy	−0.0412***
	(0.0973)		(0.345)		(0.00202)
1986	0.0490	Years of Education of Spouse	−0.263***	Years of Education of Spouse	0.00542***
	(0.0961)		(0.0126)		(0.000101)
1987	−0.0571	X Male Dummy	0.774***	X Male Dummy	−0.00476***
	(0.0943)		(0.0295)		(0.000276)
1988	−0.114	Part. Dummy last year	4.712***	Squared	−0.00429***
	(0.0922)		(0.104)		$(5.62e{-}05)$
1989	−0.338***	X Male Dummy	−4.846***	X Leisure Last Year	0.000935***
	(0.0906)		(0.167)		$(5.67e{-}05)$
1990	−0.199**	Part. Dummy Two Years Lagged	0.0548	X Male Dummy	0.000971***
	(0.0994)		(0.0894)		$(5.80e{-}05)$
1991	−0.0233	X Male Dummy	0.761***	X Leisure Two Years Lagged	0.000811***
	(0.103)		(0.172)		$(4.49e{-}05)$
1992	0.0384			X Male Dummy	−0.000505***
	(0.107)	Risk Aversion	0.01		$(5.84e{-}05)$
1993	0.281**		(0.012)	Inverse of Variance of Shock	0.000665***
	(0.121)	Observations	80986		$(7.69e{-}05)$

Note: Standard errors are in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Participation and Occupation Composition

Women Participation			
Source	1974:1978	1984:1988	1992:1996
Raw	62	70	72
Hiring Cost	56	62	61
Private Information	51	57	56
Fraction of Women			
Professional			
Raw	34	40	45
Hiring Cost	30	35	37
Private Information	28	38	41
Nonprofessional			
Raw	48	48	47
Hiring Cost	42	40	41
Private Information	45	46	45

Table 8: Average Annual Hours Worked

	Professional					
	1974:1978		1984:1988		1992:1996	
Source	Women	Men	Women	Men	Women	Men
Raw	$1,640$	$2,201$	$1,904$	$2,226$	$1,902$	$2,297$
Hiring Cost	$1,975$	$2,017$	$2,057$	$2,098$	$2,049$	$2,087$
Private Information	$1,813$	$2,088$	$1,980$	$2,090$	$1,988$	$2,120$

	Nonprofessional					
	1974:1978		1984:1988		1984:1988	
	Women	Men	Women	Men	Women	Men
Raw	$1,424$	$1,998$	$1,635$	$2,117$	$1,773$	$2,184$
Hiring Cost	$1,580$	$2,000$	$1,790$	$2,060$	$1,820$	$2,076$
Private Information	$1,510$	$1,970$	$1,640$	$1,930$	$1,768$	$1,897$

Table 9: Decomposition of Change in Gender Human Capital Gap:
$b_{\tau 1}h_{nt} + b_{\tau 2}h_{nt}^2 + \sum_{r=1}^{\rho} b_{\tau 3r}h_{nt-r} + \sum_{r=1}^{\rho} b_{\tau 4r}d_{nt-r}$
(Median Women Value over Median Men Value (%))
1974–1978: 1984–1988: 1992–1996

	Professional		Nonprofessional	
Source	1984–1988	1992–1996	1984–1988	1992–1996
Hiring Cost	38	37	30	29
Private Information	12	13	13	14
Demographic	28	25	38	39
Home Production Shock	2	1	1	1
Production Shock	18	22	11	10

Figure 1: Aggregate Occupation-Specific Productivity Shocks

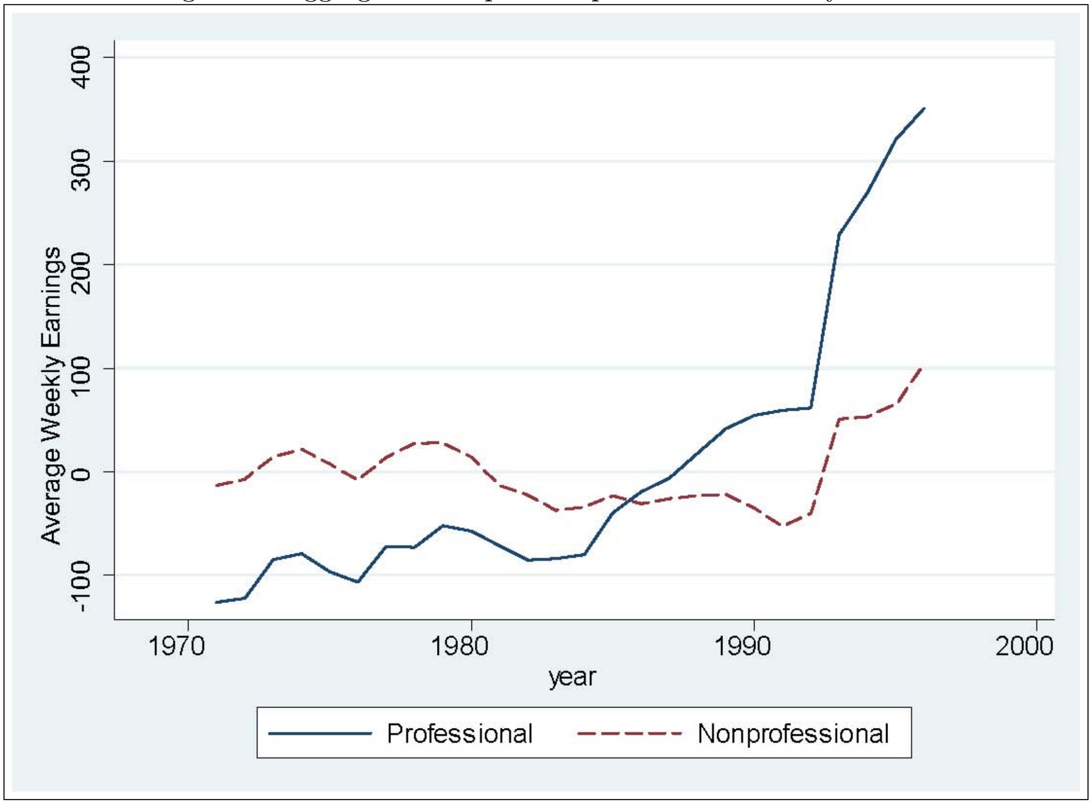

Figure 2: Gender Earnings Gap Data and Model Predicted

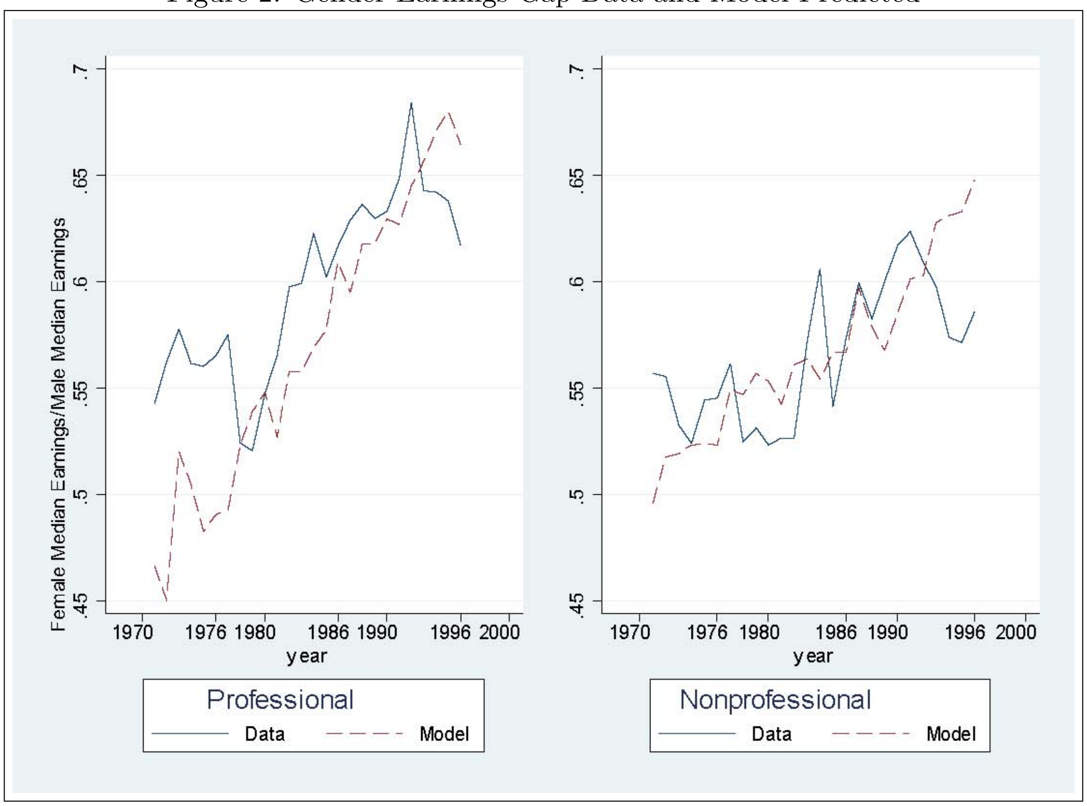

Figure 3: Wage Gap and Wage-Age Profiles

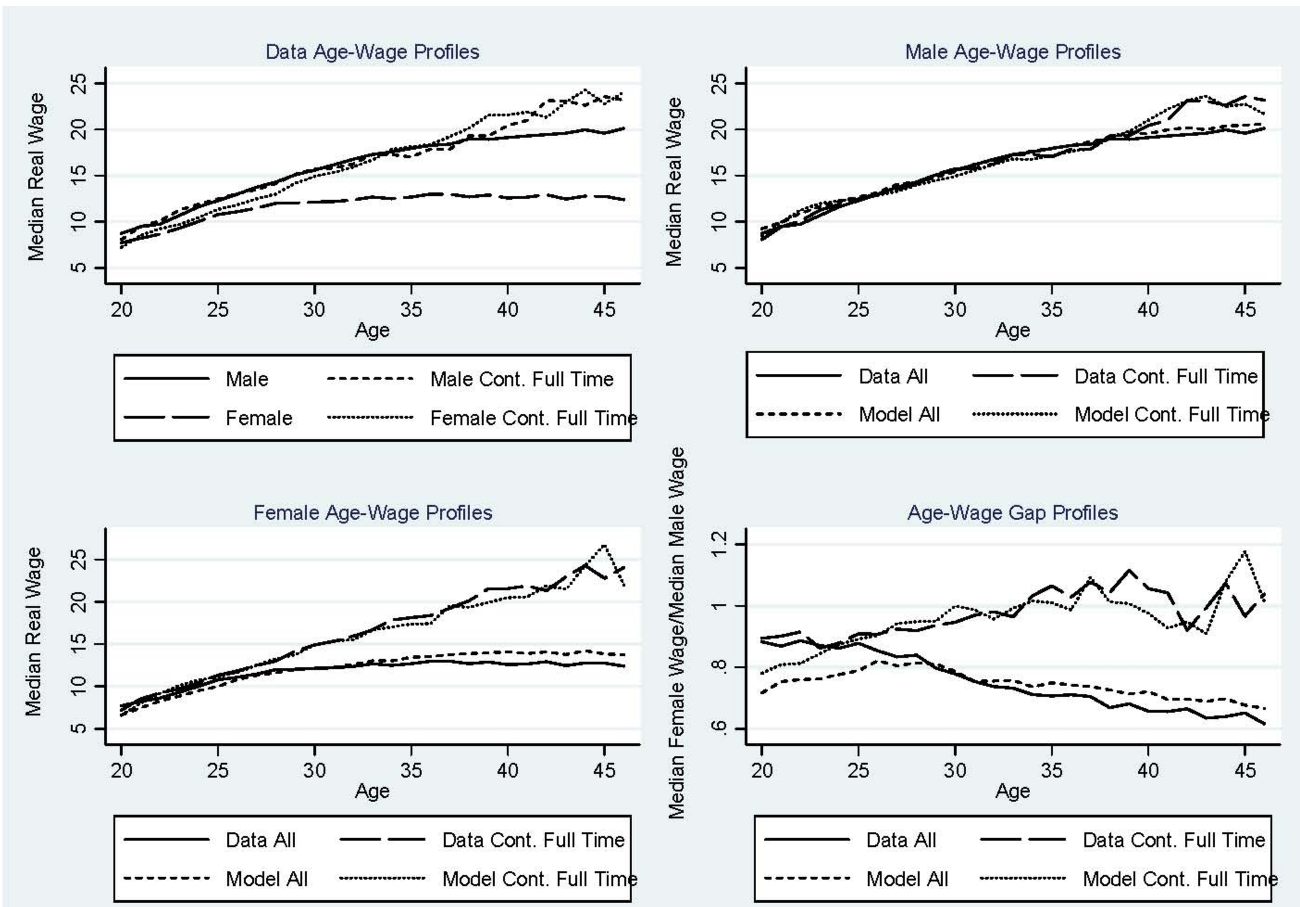

Figure 4: Counterfactual Decomposition of Gender Earnings Gap

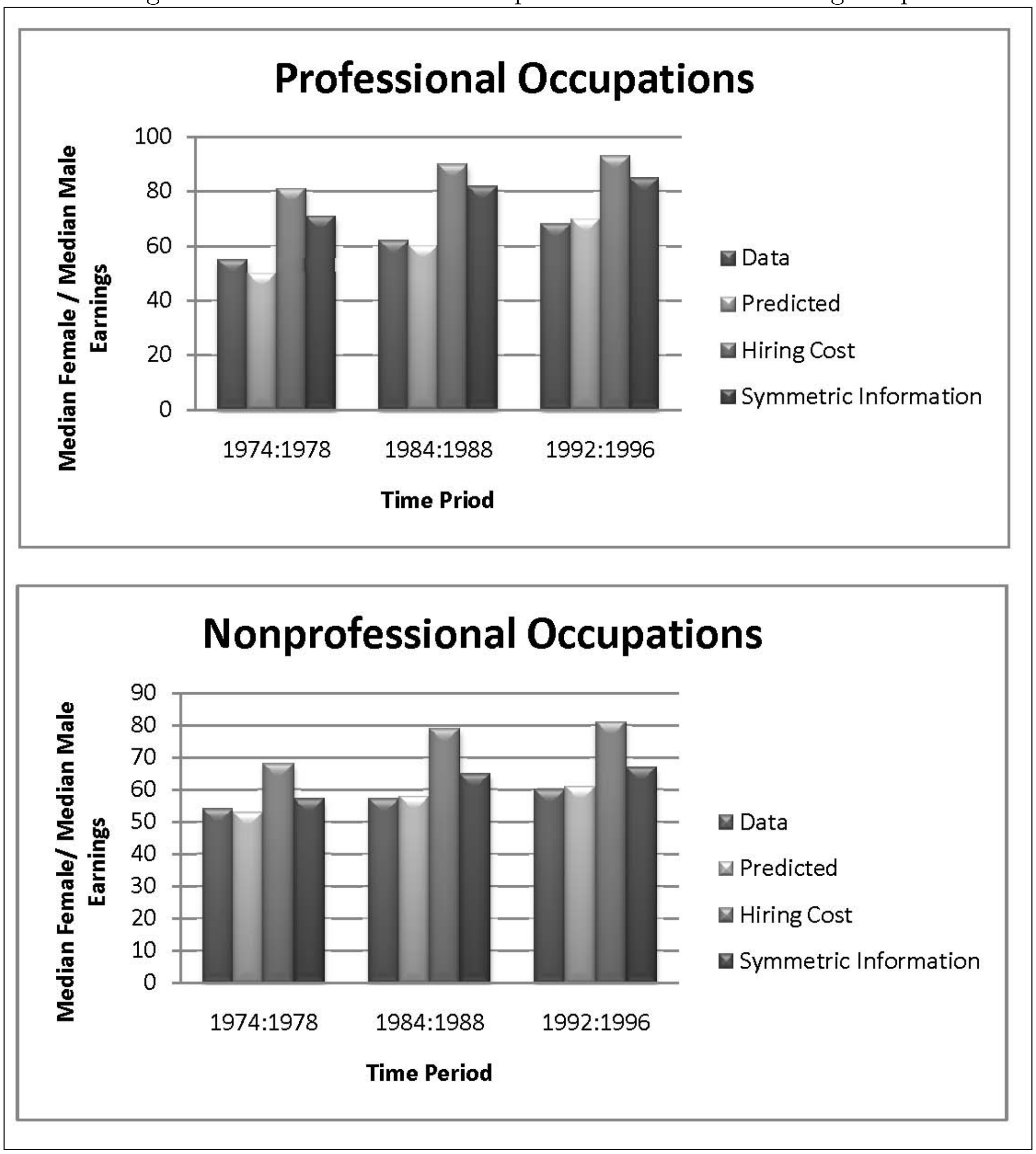

Figure 5: Decomposition of the Change in Gender Earnings Gap