Estimating a Dynamic Adverse-Selection Model: Labour-Force Experience and the Changing Gender Earnings Gap 1968-1997

Author(s): GEORGE-LEVI GAYLE and LIMOR GOLAN

Source: *The Review of Economic Studies*, Vol. 79, No. 1 (January 2012), pp. 227-267
Published by: Oxford University Press
Stable URL: http://www.jstor.org/stable/41407050
Accessed: 27-09-2017 21:59 UTC

Review of Economic Studies (2012) **79**, 227–267 doi: 10.1093/restud/rdr019

Advance access publication 8 September 2011

Estimating a Dynamic Adverse-Selection Model: Labour-Force Experience and the Changing Gender Earnings Gap 1968–1997

GEORGE-LEVI GAYLE and LIMOR GOLAN

Tepper School of Business, Carnegie Mellon University

First version received June 2008; *final version accepted April* 2011 (*Eds.*)

This paper addresses two questions: What accounts for the gender gap in labour-market outcomes? What are the driving forces behind the changes in the gender labour-market outcomes over the period 1968–1997? It formulates a dynamic general equilibrium model of labour 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 female labour-force participation, the proportion of women working in professional occupations, and hours worked. It finds that labour-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 female labour-force participation, number of hours worked, and the proportion of females working in professional occupations. Although of less importance, demographic changes account for a substantial part of the increase in female labour-force participation and hours worked, whereas home production technology shocks do not.

Key words: Gender earnings gap, Statistical discrimination, Occupational sorting, Human Capital, Structural estimation of dynamic games, Dynamic general equilibrium, Adverse selection, Signalling.

JEL Codes: J01, J2, J31, J71, C5, C72, C13, D82

1. INTRODUCTION

One of the most striking changes in the U.S. labour market over the last four decades is the significant decline in the gender wage gap in the 1970's and 1980's. The median gender wage differential dropped from about 40% in 1968 to 28% by 1992 and has remained constant since. Accompanying this decline were significant increases in women's labour-force participation and hours worked and the proportion of women working in professional occupations. Basic empirical analysis suggests that the patterns of the decline in the gender wage gap are intertwined with the patterns of increasing female labour-force participation and hours worked. This fact is strikingly illustrated by examining the life cycle evolution of the gender wage gap as a function of labour-market experience. As illustrated in the bottom right panel of Figure 3, the 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 and closes by the age of 36. To explain these patterns, the paper formulates a life cycle general equilibrium model of labour supply, occupational sorting, and human-capital accumulation in which gender discrimination and the earnings gap arise endogenously. The model includes three channels through which a gender earnings gap can arise: group differences in preferences, group differences in productivity, and discrimination. It then uses the framework to assess the importance of these channels in the observed gender gap in labour-market outcomes and its decline over time. In addition, it demonstrates identification and develops a three-step estimator of the model. To the best of our knowledge, this is the first paper to estimate a dynamic signalling model.

The model consists of forward-looking workers and firms. Each period, workers make labour-market participation, occupation, and labour-hours decisions. The supply side of the model extends the labour-supply model of Mincer and Polachek (1974) and Polachek (1981) by incorporating privately observed labour-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 labour-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 labour-supply behaviour of workers with lower participation costs. The model gives rise to a notion of statistical discrimination proposed by Barron *et al.* (1993). 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 labour-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.

The empirical analysis shows that human capital accumulated with labour-market experience is the most important factor explaining the gender earnings gap and that private information plays an important role in women's labour-market outcomes. The human capital accumulated with experience affects the earnings and the gender pay gap because it increases the worker's productivity and because of its signalling effect. Because of the signalling element, private information has a direct effect on the salaries and the gender pay gap and an indirect effect through the gender gap on the human capital accumulated with experience because workers internalize the effect their labour-market decisions have on employers' beliefs. Empirically, the total direct and indirect effects of the private information on the gender earnings gap are large.

Figure 3 demonstrates the data variations used to identify these effects. The gender wage gap increases over the life cycle largely because the gap in the human capital accumulated increases with age. At the same time, women who work continuously full time have faster wage growth than men with the same labour-market experience, driving the decline in the wage gap for these workers. The productivity and private-information effects of human capital on the gender pay gap are disentangled using the data variation in future labour-market participation patterns of otherwise identical men and women. The faster wage growth of women who work continuously full time relative to men's identifies the signalling effect.

One of the main difficulties in estimating game-theoretic models is the possibility of multiple equilibria. We use a multistage estimation procedure to overcome the indeterminacy caused by a multiplicity of equilibria. We then use the estimates to conduct counterfactual experiments to assess the effect of different factors on the gender earnings gap, labour-market experience, and occupational sorting. These counterfactuals were conducted under variations of our model where there is a unique equilibrium, namely when the information is symmetric and when there is no hiring cost. The outcomes under these scenarios were then compared to the observed outcomes in the data, allowing us to ascertain the effects of the private information from other factors affecting the gender earnings gap and its decline. 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 1980's and 1990's. In addition, we find that signalling is an important factor in labour-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 labour market.

The literature focuses on several factors that caused the changes in the labour-market outcomes of women relative to men between the 1970's and the 1990's. 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 the increase in the proportion of women working in these occupations, their labour-force participation, and their hours worked. Demographic changes reduced women's costs of participating in the labour market, increasing their participation and hours worked. In contrast to Greenwood *et al.* (2005), our results do not support the hypothesis that changes in home production costs explain the increase in women's labour-market experience and hours worked over the period we examine, but the time period in their paper does not completely overlap ours. This result, however, is similar to the findings in Jones *et al.* (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 increase in women's labour-market experience in professional and non-professional occupations, respectively.

Our model builds on the theoretical literature on statistical discrimination pioneered by Arrow (1972) and Phelps (1972).[1] The theoretical and empirical statistical discrimination literature that emphasizes that group differences can arise endogenously without any *ex ante* differences across groups focuses primarily on productivity differences across racial groups (see Coate and Loury, 1993, for this insight and Moro 2003; Moro and Norman, 2004; Antonovics, 2004; Altonji, 2005, among others). Two exceptions are Barron *et al.* (1993) and Albanesi and Olivetti (2009). Our paper contributes to this literature by incorporating statistical discrimination, as formulated in Barron *et al.* (1993), into a general equilibrium model with life cycle labour-supply choices; this allows us to capture the evolution of the labour-market-outcomes gap over the life cycle. The estimator we develop is akin to a number of estimators in the literature on the estimation of discrete games and single-agent models (Hotz and Miller, 1993; Altug and Miller, 1998; Pesendorfer and Schmidt, 2003; Bajari *et al.*, 2007; Pakes *et al.*, 2007). Our estimator is different, however, in that we are estimating a perfect Bayesian

1. A second category of discrimination literature is the taste-based discrimination theory pioneered by Becker (1971) which 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.

equilibrium of a dynamic adverse-selection model while these papers estimate Markov perfect equilibrium games.

While there is an extensive empirical literature on the gender wage gap (see Altonji and Blank, 1999, for a survey), few papers specify a fully behavioural model explaining the existence of a gender wage gap and its evolution over the life cycle in this literature. Bowlus (1997) presents a static search model with neither discrimination nor human capital accumulation, and Flabbi (2009, 2010) adds taste-based discrimination to the Bowlus (1997) model. None of these papers accounts for the observed patterns of the gender earnings gap over the life cycle. Erosa *et al.* (2005) calibrates a partial equilibrium life cycle model of labour supply and fertility choice with no discrimination; it does not explain, however, the decline in the wage gap for workers who work continuously full time. The literature explaining the decline of the labour-market-outcomes gender gap can be divided into two main groups: papers that examine the increase in labour-force participation and hours of married women (see Jones *et al.*, 2003;[2] Greenwood *et al.*, 2005; Attanasio *et al.*, 2008; Albanesi and Olivetti, 2008; Fernandez and Fogli, 2009) and papers that examine the decline in the gender wage gap and the increase in women's labour-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, but it does not account for the role of discrimination.

This paper is organized as follows. Section 2 describes the model. Section 3 presents the equilibrium analysis and the equilibrium labour-market gender gaps. The empirical implementation is presented in Section 4. Section 5 shows identification 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 (1998) general equilibrium model of dynamic labour supply and consumption. We incorporate occupational sorting, hiring costs, and private information into their framework. The hiring costs and private information give rise to endogenous gender discrimination as formalized by Barron *et al.* (1993) in a partial equilibrium model.

The economy consists of infinitely lived firms and finitely lived workers. 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 age–education cohort by t ($t = 0, \ldots, T$). The theoretical model is written and solved for a given age–education cohort, but the economy consists of a number of overlapping cohorts. There is free entry into the competitive labour markets. There are two occupations, $\tau \in \{\mathrm{P}, \mathrm{NP}\}$, professional and nonprofessional, each of which 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 labour-market participation and occupation choices, deciding how many hours to work and how much to consume. Labour-market participation and occupation choices are discrete, while hours and

2. Jones *et al.* (2003) allow for an exogenous gender wage wedge that changes over time.

consumption choices are continuous. We denote the participation decision by d_t, where $d_t = 1$ if the individual participates in the labour force in period t and 0 otherwise. The occupation and participation indicators are defined as $I_{\tau t}(\tau \in \{\text{P}, \text{NP}\})$, 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 labour-supply decision vector by $a_t = (d_{t-1}, \{I_{\tau t-1}\}_{\tau \in \{\text{P,NP}\}}, h_t)$. 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 labour-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 $z_t = (a_0, \dots, a_{t-1}, x_t)$ to be a vector with the first $t - 1$ elements capturing the dependence of the utility on the past labour-supply choices; the last element, x_t, is independently distributed over the population and evolves according to a known gender-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 labour force and ζ_t is an aggregate shock to this cost drawn from the 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 that 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 asset markets are competitive and complete and that there are no frictions in the markets for loans.[3] These assumptions allow us to compactly write the lifetime individual budget constraint:

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

where $\overline{S}_t$ is the total household labour-market income (or the individual's income if he/she is single), λ_t is the expected price of the contingent claim, and W is bequest net of inheritance.[4] We assume that λ_t is distributed according to the known distribution $F_\lambda(\lambda_{t+1} \mid \lambda_t)$. We also assume that the process of formation and dissolution of households is exogenous and is governed by $F_{i0}(x_{t+1} \mid z_t)$. Thus, expectations of the budget constraint of each individual within a household take into account the distribution of future household composition.

3. Other papers that use similar assumptions include Altug and Miller (1990, 1998), Card (1990), Mace (1991), Townsend (1994), and Altonji *et al.* (1996).

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

2.1.1. Frisch demand for consumption and the indirect utility. The additive and intertemporal separability of consumption along with the complete-asset-markets assumption allow us to characterize the Frisch demand for consumption separately from the labour-supply decisions. Let η denote the Lagrange multiplier associated with the budget constraint in equation (2) for each household member. The first-order conditions with respect to individual consumption c_{it} are

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

for all $t \in \{0, 1, 2, \ldots\}$ and $i \in \{\mathrm{w}, \mathrm{m}\}$. Thus, each individual within a household smooths consumption. Note that η is also the inverse of the Pareto weight of each individual in a household in the social planner's problem. Thus, the optimal consumption allocation is constrained Pareto efficient.[5]

Let $S_{i\tau t}(h_t, \cdot)$ be the salary paid to an individual that works h_t hours in occupation τ in period t. The second element of the salary schedule is the worker's characteristics as observed by the employer. We analyse the model under symmetric and asymmetric information. Under asymmetric information, the employers' information set is a subset of their information set under symmetric information. We can now write the current-period utility as

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

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

2.2. *Firms*

2.2.1. Technology. Each firm in each occupation produces output $(y_{\tau t})$ using labour (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 labour-market history and other variables that affect 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

5. The allocation is constrained Pareto efficient as the planner is constrained by the same information asymmetry the employers face.

productivity is greater than it is in the other occupation,

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

We also assume that $Y_\tau(h_t, z_t^{\mathcal{P}}, K_{\tau t})$ is twice continuously differentiable and that for every pair of occupations τ' and τ, the left-hand side and the right-hand side of equation (6), each representing the discounted sum of the expected lifetime net productivity in an occupation, cross at most once as a function of current hours worked. The strict inequality in equation (6) implies that, for every occupation τ and worker's characteristics, $z_t^{\mathcal{P}}$, an open set of hours is offered, $h \subset (\underline{h}_{\tau t}(z_t^{\mathcal{P}}), \overline{h}_{\tau t}(z_t^{\mathcal{P}}))$, and that the sets are disjoint across occupations. The other two assumptions imply that for a given occupation the set of hours offered forms an interval. Note that the expected lifetime net productivity of a worker in an occupation depends on experience and past hours worked. Thus, it is possible that in one occupation the current productivity is lower, however, the expected lifetime net productivity is greater compared to other occupations. In addition, because different hours are offered in different occupations, for given worker characteristics, changes in hours worked may involve changing occupations. This assumption captures the limited flexibility of hours offered within some occupations.

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

2.2.3. Timing. At the beginning of each period, the state variables ω_t are realized, including the worker's time-varying characteristics as well as the aggregate productivity innovations and prices and $(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})$. Observing these realizations, the workers then make participation, occupation, and hours choices. Observing the workers' choices, firms simultaneously offer a salary to each worker. Workers observe the offers and choose a firm. Finally, production occurs and the agents consume. This structure repeats itself every period. While the aggregate prices and productivity innovations are commonly observed, the workers' state variables observed by the firms will vary in the analysis of the symmetric- and asymmetric-information cases.

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:

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

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

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

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 with 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 period 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 labour markets are competitive, the worker's discounted expected lifetime earnings are equal to the worker's expected net output in an outside firm.

Equilibrium labour 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

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

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

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

The probability of participation conditional on ω_t is

$$p_i^{o}(\omega_t)=E[d_i^{o}\mid\omega_t]=\int_{-\infty}^{V_{1i}-V_{0i}}(\varepsilon_{0t}-\varepsilon_{1t})dF_1(\varepsilon_{0t},\varepsilon_{1t})\equiv Q(V_{1i}(\omega_t)-V_{0i}(\omega_t)). \tag{15}$$

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

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

and, for non-participation, it is

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

Next, we characterize the optimal hours and occupation decisions beginning with the optimal hours choice, $h^*_{i\tau t}$, in *each* occupation. The first-order condition with respect to hours is

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

To obtain the optimal hours in each occupation, $h^*_{i\tau t}$, the worker solves equation (18) for each occupation over the open interval $(\underline{h}_{\tau t}(z_t^{\mathcal{P}}),\overline{h}_{\tau t}(z_t^{\mathcal{P}}))$. The worker then chooses the hours in the occupation that yields the highest expected lifetime utility, h^*_{it}. Thus, the occupation choice is given by

$$I^0_{i\tau}(\omega_t)\equiv I\{\underline{h}_{\tau t}(z_t^{\mathcal{P}})<h^*_{it}(\omega_t)<\overline{h}_{\tau t}(z_t^{\mathcal{P}})\}. \quad (19)$$

The occupation choice affects the probability of working in the firm next period and therefore future salaries; this probability is given by

$$p_{i\tau t+1}(h_t,\omega_t)=\int Q(\omega_{t+1})I^0_{i\tau}(\omega_{t+1})f_{i0}(\omega_{t+1}\mid\omega_t,a_t)d\omega_{t+1}. \quad (20)$$

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 labour-supply decisions considered by Polachek (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 labour-market experience and non-market hours. Labour-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 observe ω_t^*, a subset of the worker's state variables, ω_t, and do not observe the i.i.d. 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 $\omega_t^* = (a_0, \dots, a_{t-1}, K_{Pt}, K_{NPt}, x_t^*)$. We assume that all production-relevant variables, $z_t^{\mathcal{P}}$, are observed. To make the complete-asset-markets assumption consistent with private information, we assume anonymous trades in the asset markets.

This information structure transforms the model into a signalling 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 labour-market actions, a_t, firms update their beliefs about each worker's type. We denote the *posterior* beliefs by $\widetilde{\mu}_{it}(\omega_t \mid \omega_t^*, a_t)$. Note 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 1 (Equilibrium). A perfect Bayesian equilibrium consists of labour-market strategies, σ_{it}, Frisch demand for consumption, c_{it}^o, the workers' contract-offer choices, $\{S_{i\tau t}\}_{\tau \in \{P,NP\}}$, 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,

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

and, for all histories, types, and actions,

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

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 non-trivial 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 non-existence 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 signalling game, and the non-existence result of the original Rothschild–Stiglitz model is broken. The proposition below establishes existence.

Proposition 1. *There exists a perfect Bayesian equilibrium characterized by the following strategies and beliefs.*

1. *Denote by $\widetilde{p}_{i\tau t+1}(h_t, \omega_t^*)$ the employers' beliefs about the probability of working in the firm next period. For all $h_t \in (\underline{h}_{\tau t}, \overline{h}_{\tau t})$, the optimal salary is given by*

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

2. *Under asymmetric information the workers' optimal consumption and labour-supply strategies, $\{a_t^0, c_{it}^o\}$, are as described in the symmetric-information case, except for the salaries in the state variables entering equation* (3) *and the valuation functions in equations* (13)–(19) *which are replaced by the salary in equation* (23).

3. The employers' beliefs are given by

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

The above proposition establishes that there exists an equilibrium, but there may be many equilibria that satisfy all the above conditions. In particular, note that the employers' beliefs described in part 3 are the implied beliefs about the probability of working in the firm in the following period and that the conditions in part 2 are necessary conditions for the optimality of the workers' strategies. Therefore, there may be other equilibria with different sets of prior beliefs, $\mu_{it}(\omega_t|\omega_t^*)$, and workers' equilibrium choices, implying different future probabilities of working in the occupation, which satisfy all the above conditions.

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),

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

If the information were 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 labour-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 analysed in Laffont and Tirole, 1988). Note that participation and occupation choices have similar signalling elements as they affect ω_{t+1}^* similarly.

The optimal contract provides insurance to the risk-adverse workers; hence it dominates contracts that 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, *e.g.* 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.

3.2.1. Labour-market-outcome gender gaps. In the model, gender differences in labour-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 labour-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.[6]

3.2.2. 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 men's 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 labour-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 cohorts.

4. EMPIRICAL IMPLEMENTATION

Panel data are needed for the empirical implementation of the model, so let $n = 1, \ldots, N$ index individuals and $t = 1, \ldots, T$ index time periods. Note that we do not use i to index individuals because there are a number of gender-specific functions (*e.g.* $F_{i0}(\cdot|\cdot)$ and $\widetilde{p}_{i\tau t+1}(\cdot,\cdot)$) that will be estimated for both genders for all observations in our data set. It is assumed that all labour-market decisions, $\{a_{nt}\}_{n=1,t=1}^{N,T} = \{(d_{nt-1}, \{I_{n\tau t-1}\}_{\tau\in\{\text{P,NP}\}}, h_{nt})\}_{n=1,t=1}^{N,T}$, are observed by the econometrician. We denote the gender of a given observation by the indicator i_n and assume that all the individual systematic time-varying determinants of the utility of non-market hours, $\{z_{nt}\}_{n=1,t=1}^{N,T} = \{(a_{n0}, \ldots, a_{nt-1}, x_{nt}, i_n)\}_{n=1,t=1}^{N,T}$, are observed by the econometrician. We also assume that our measure of salary, S_{nt}, is an error-ridden measure of the true salary, $S_{i\tau t}$. The error is assumed to be classical and therefore independent of all production-relevant variables, gender, and occupation. We further assume that all the individual production-relevant variables, $z_t^{\mathcal{P}}$, except for individual-specific time-invariant variables, denoted by ν_n, are observed by the econometrician. Consumption is assumed to be observed and is measured at the household level. To handle this issue, we follow the standard solutions in the literature on the estimation of the consumption equation. We provide details below.

6. 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.

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 employer's beliefs, $\widetilde{p}_{i\tau t+1}(h_t, \omega_t^*)$. These two functions depend on age and years of education and including year dummies makes them a function of the cohort as well. These functions are estimated non-parametrically using kernel-density estimation.

The private information we use in the estimation consists of data that the econometrician observes as well as persistent heterogeneity unobserved to the econometrician. To select which variables are private information, we first choose variables that are not observed by potential employers and the worker is not obligated to report truthfully by law.[7] 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) 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 a 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 details). Below, we describe the data and parametric assumptions we make about the model's underlying 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 non-food consumption and good job-to-job–transition data. As a result, job switches in the data are occupation changes and transitions in and out of the labour force.[8]

The data are 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 non-professional. 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 labour-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 7.

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 (non-professional) occupations over the period. Note that the earnings gap is normally 50% larger than the wage gap because women not only earn less but

7. Note 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.

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

also work fewer hours per year than men. We therefore focus on the earnings gap in order to capture both dimensions.

Table 1 contains summary statistics of our main labour-market and human-capital variables. The participation rate for men is relatively constant over the sample period with a slight decline towards 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 1400 hours per year in 1968 to 1868 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 working in the non-professional 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 the 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 are 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.

TABLE 1
Summary of labour-market and human-capital variables

	Participation		Hours		Earnings		Fraction of women		Education	
	Male	Female	Male	Female	Male	Female	Professional	Non-professional	Male	Female
1968	0·93	0·54	2244	1401	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	2240	1371	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	2216	1332	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	2175	1382	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	2155	1389	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	2188	1411	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	2130	1424	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	2230	1415	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	2092	1395	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	2119	1418	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	2115	1454	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	2141	1472	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	2112	1450	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	2199	1642	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	2166	1630	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	2136	1632	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	2142	1635	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	2188	1646	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	2192	1665	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	2215	1690	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	2230	1691	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	2221	1703	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	2251	1683	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	2259	1807	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)

(*continued*)

TABLE 1
Continued

	Participation		Hours		Earnings		Fraction of women		Education	
	Male	Female	Male	Female	Male	Female	Professional	Non-professional	Male	Female
1992	0·87	0·74	2221	1815	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	2205	1801	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	2250	1820	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	2260	1833	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	2276	1829	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	2280	1868	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)

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

4.3. *Production function*

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

$$Y_\tau(h_{nt}, z_{nt}^{\mathcal{P}}, K_{\tau t}) = K_{\tau t} + b_{\tau 1}h_{nt} + b_{\tau 2}h_{nt}^2 + \sum_{r=1}^{\rho} b_{\tau 3r}h_{nt-r} \sum_{r=1}^{\rho} b_{\tau 4r}d_{nt-r} + b_{\tau 5}\text{age}_{nt} + b_{\tau 6}\text{age}_{nt}^2 + b_{\tau 7}\text{age}_{nt} \times \text{education}_n + \nu_n, \tag{29}$$

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 non-linearity 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 $E_t[d_i^0 \mid \omega_t]$ is given by

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

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

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

and

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

TABLE 2
Summary of demographic and wealth variables

	Food consumption	Family size	Age	Number of kids ≤ 5 years old	Number of kids aged 6 to 17	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)

(*continued*)

TABLE 2
Continued

	Food consumption	Family size	Age	Number of kids: $\leq$ 5 years old	Number of kids: aged 6 to 17	Marital status
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)

Notes: Standard deviation in parentheses. Household income and food consumption in thousands of year 2000 US$.

where ξ is the Euler constant.

We parameterize 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

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

The specification of the utility from non-market hours follows the literature on non-separable utility (see, among others, Hotz *et al.*, 1988; Altug and Miller, 1990; Becker *et al.*, 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 signalling game and past decisions are central to modelling reputation. Specifically, in a perfect Bayesian equilibrium, the beliefs depend on all past decisions, not just on those relevant to the pay-offs (*i.e.* utility and output). In order to make the estimation feasible and at the same time capture the signalling content of past decisions, we restrict the beliefs to depend on the complete labour-market history in the past three periods and the total number of periods 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. Below we state formally the definition and show by construction that our model satisfies this condition.

Definition 2. Given any value of the initial state variable, ω_0, there exists a finite integer, $\rho(\omega_0)$, a value of the state variable, $\omega_{\rho(\omega_0)}$, and two sequences of choices over the next $\rho(\omega_0)$ periods, denoted by $d_{0t}^{\rho(\omega_0)}(\omega_t) = (d_{0t+1}^{\rho(\omega_0)}(\omega_{t+1}), \dots, d_{0t+\rho(\omega_0)}^{\rho(\omega_0)}(\omega_{t+\rho(\omega_0)}))$ and $d_{1t}^{\rho(\omega_0)}(\omega_t) = (d_{1t+1}^{\rho(\omega_0)}(\omega_{t+1}), \dots, d_{1t+\rho(\omega_0)}^{\rho(\omega_0)}(\omega_{t+\rho(\omega_0)}))$ such that irrespective if choice $k = 0$ or $k = 1$ is chosen today the state will be $\omega_{\rho(\omega_0)}$ at date $\rho(\omega_0)$.[9]

We first characterize all the different possible sequences of choices that can lead to the same labour-market history at a certain point in time due to finite-state dependence. Consider the hypothetical labour-market histories

$$z_{1nt}^{(s)} = (NY_{n1t-1+s}, NY_{n2t-1+s}, d_{nt-3+s}I_{n1t-3+s}, d_{nt-3+s}I_{n2t-3+s}, \dots, d_{nt-1}I_{n1t-1}, d_{nt-1}I_{n2t-1}, I_{n1t}, I_{n2t}, 0, \dots, 0, h_{nt-3+s}, \dots, h_{nt-1}, h_{nt}^*, 0, \dots, 0, x'_{nt+s}) \quad (34)$$

and

$$z_{0nt}^{(s)} = (NY_{n1t-1+s}, NY_{n2t-1+s}, d_{nt-3+s}I_{n1t-3+s}, d_{nt-3+s}I_{n2t-3+s}, \dots, d_{nt-1}I_{n1t-1}, d_{nt-1}I_{n2t-1}, 0, 0, I_{n1t}, I_{n2t}, 0, \dots, 0, h_{nt-3+s}, \dots, h_{nt-1}, 0, h_{nt}^*, 0, \dots, 0, x'_{nt+s}), \quad (35)$$

where

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

for $s = 1, 2, 3$. The vector $z_{1nt}^{(s)}$ 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 $z_{0nt}^{(s)}$ 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 he/she worked.* For the following $s - 2$ periods after period $t + 1$, this individual chooses not to work. Note that these two sequences of decisions will lead to the same labour-market history in period $t + 4$. Let us define the following participation indices that correspond to the decision sequences that lead to a state vector $z_{knt}^{(s)}$:

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

and

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

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 $z_{0nt}^{(s)}$. These are then the sequences as required by Definition 2. This definition can be generalized to account for stochastic transitions of the state variables. In that case, a decision process is said to exhibit *stochastic finite-state dependence* if

9. Unlike the optimal participation choice, $d_t^{\rho(\omega_0)}(\omega_t)$ does not depend on ε by definition. Instead, it is a deterministic function of the state variables.

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. The Markovian transition of the variables x_{nt} is then needed for the above decision sequences to satisfy the assumption of stochastic finite-state dependence.

5. ESTIMATION AND IDENTIFICATION

The possibility of multiple equilibria poses a problem in estimating game-theoretic models because it induces indeterminacy in standard estimation-criterion functions, such as the likelihood functions, that map 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 that 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.

In order to formally describe the distinction between multiple equilibria and identification, we first introduce the following notation. Define the mapping $\mathcal{F}$ from the space of structural characteristics $\Theta(M)$ to the space of conditional distributions Π such that $\mathcal{F}(\theta(M))$ contains all the conditional distributions predicted by the model when the structural characteristic is $\theta(M)$. Multiple equilibria means that $\mathcal{F}(\theta(M))$ is a correspondence, while identification means that the inverse mapping $\mathcal{F}^{-1}(\cdot)$, evaluated at the population distribution $F^{o}_{Y|X}$, is a function. Clearly, a mapping can be a correspondence and its inverse a function, which means that multiple equilibria may not imply nonidentification. Relying on Chesher (2007), we show that there exists a functional of the conditional distribution, $\mathcal{F}^{-1}(F_{Y|X})$, with the property that the functional returns the value θ^* in all the admissible structures of the model with $\theta(M)=\theta^*$ and hence our model is identified. This functional gives us a natural analogue estimator of the structural parameters of our model 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 labour supply (*i.e.* propensity to change occupation, labour-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 non-linear 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,[10] the data are generated by only one equilibrium (see Bajari *et al.*, 2007; Aguirregabiria and Mira, 2007; Pesendorfer and Schmidt, 2008). Under this assumption, the conditional choice probabilities are identified. Note, 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, labour-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—*i.e.* $d_t = k$—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.* $p_{kit}^{(s)} = E_t[d_{t+s} \mid \omega_{kt}^{(s)}]$. 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

$$m_{i2t} \equiv \eta\lambda_t \sum_{\tau\in\{\text{P,NP}\}} I_{\tau t}^{\text{o}} S_{i\tau t}^{\text{o}} + \sigma \sum_{s=1}^{3} \beta^s \ln\left(\frac{1-p_{1it}^{(s)}}{1-p_{0it}^{(s)}}\right) - \sigma \ln\left[\frac{p_{it}}{1-p_{it}}\right] + \zeta_t + \sum_{s=1}^{2} \kappa_{is} d_{t-s} + x_t' B_{i1} - x_t' h_t B_{i2} - \theta_{0i}(1-l_t^2) - \sum_{s=1}^{2} \theta_{si} h_t (l_{t-s} + \beta^s) \tag{36}$$

and

$$m_{i3t} \equiv \eta\lambda_t \sum_{\tau\in\{\text{P,NP}\}} I_{\tau t}^{\text{o}} \frac{\partial S_{i\tau t}^{\text{o}}}{\partial h_t} + \sigma \sum_{s=1}^{3} \beta^s (1-p_{1it}^{(s)})^{-1} \frac{\partial p_{1it}^{(s)}}{\partial h_t} - x_t' B_{i2} - 2\theta_{0i} l_t + \sum_{s=1}^{2} \theta_{si}(l_{t-s} + \beta^s), \tag{37}$$

where $E_t[m_{ijt}] = 0$ 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). The above equations are derived in Appendix B. By inspection, the model is identified once the consumption and salary equations, the beliefs, and the conditional choice probabilities are identified.

5.2. *Estimation*

5.2.1. 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

10. Observable characteristics are either characteristics directly observed in the data or characteristics that can be identified from the data separately, such as the individual fixed effects in production and consumption.

sample. Because 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.

5.2.2. Equilibrium beliefs, their derivatives, and conditional choice probabilities. The equilibrium beliefs for each occupation, $\widetilde{p}_{in\tau t+1}$, are computed as a non-linear 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 age 5 and younger, the number of children from age 6 to 17, age, years of completed education, marital status, spouse's years of education, and gender. The conditional choice probabilities, p_{int}, are computed using non-linear regressions of the participation index, d_{nt}, on the current state-variable vector, $\omega^{\mathcal{N}}_{nt} \equiv (z^{\mathcal{N}\prime}_{t}, K^{\mathcal{N}}_{\mathrm{P}t}, K^{\mathcal{N}}_{\mathrm{NP}t}, \eta^{N}_{n} \lambda^{\mathcal{N}}_{t})'$, where the superscript $\mathcal{N}$ denotes an estimated quantity.

5.2.3. Estimation of the finite-state path probabilities and their derivatives. Let $\omega^{(s)\mathcal{N}}_{knt} \equiv (z'_{nt+s}, K^{\mathcal{N}}_{\mathrm{P}t}, K^{\mathcal{N}}_{\mathrm{NP}t} \eta^{\mathcal{N}}_{n} \lambda^{\mathcal{N}}_{t+s})'$ for $k = \{0,1\}$ be the empirical counterpart of the hypothetical state. Recall that $p^{(s)}_{kit} = E_t[d_{t+s} \mid \omega^{(s)}_{kt}]$; hence, it can be estimated as non-linear regressions of the participation index, d_{nt}, on the hypothetical state, $\omega^{(s)\mathcal{N}}_{knt}$, conditional on $d^{(s)}_{knt} = 1$. We then evaluate the term $\partial p^{(s)}_{i1nt}/\partial h_{nt}$ using the standard non-parametric derivative kernel estimator (Pagan and Ullah, 1999).

5.2.4. 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 periods as well as the total number of periods 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:

$$m_{n\tau t}(\theta_{w\tau}) \equiv S_{in\tau t} - K_{\tau t} - b_{1\tau} h_{nt} - b_{2\tau} h^2_{nt} - \sum_{r=1}^{\rho} b_{3r\tau} h_{nt-r} - \sum_{r=1}^{\rho} b_{\tau 4r} d_{nt-r} - Z_{nt}{}' B_{5\tau} - \upsilon_n + \gamma_\tau - d_{nt+1} I_{n\tau t+1} \beta \gamma_\tau, \tag{38}$$

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 (2SLS).

5.2.5. Utility of non-market time. The remaining parameters are estimated by Generalized Method of Moments (GMM) using the empirical counterpart of equations (36) and (37). They are constructed by substituting the estimated quantities: β, $\eta_n\lambda_t$, $S^{o}_{i\tau t}()$, p_{int}, $p^{(s)}_{i0nt}$, and $\frac{\partial p^{(s)}_{i1nt}}{\partial h_{nt}}$. The remaining details of the implementation, the estimation of p_{int}, $p^{(s)}_{i0nt}$, and $\frac{\partial p^{(s)}_{i1nt}}{\partial h_{nt}}$, and the asymptotic properties of the estimator are in Section 2 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 non-professional 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 non-professional 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*

6.1.1. 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 $E_t[m_{n\tau t}(\theta_{w\tau}) \mid \omega_t, i, d_{nt} I_{n\tau t} = 1] = 0$, while under asymmetric information it implies that $E_t[m_{n\tau t}(\theta_{w\tau}) \mid \omega^*_t, i, d_{nt} I_{n\tau t} = 1] = 0$. Thus, under asymmetric information, the employers' information is a subset of the information available to them under symmetric information. Therefore, if the model is identified under asymmetric information, the variables assumed to be private information should also be orthogonal to the zero-profit condition under the null hypothesis that information is symmetric. Whether these variables are orthogonal or not serve as over-identifying restrictions. 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 over-identifying 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 Labour-Market Income*, and *Spouse Education* are included in the information set of the employer.

In the second step, we perform an over-identifying restrictions test for the earnings equation under the null that the information is asymmetric. There are different over-identifying 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 periods of labour-market histories, gender, and cohort, we were not able to reject the over-identifying 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 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 increased 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 that of men.

6.1.2. Parameter estimates. The main purpose of estimating the consumption equation is to obtain estimates of the marginal utility of wealth for our main estimation equations. Therefore, we do not focus the discussion on these results. Table 3 contains the results from this estimation. These results are standard and consistent with estimates of these parameters in the previous literature (see Altug and Miller, 1998; Gayle and Miller, 2004, for similar estimates).

TABLE 3
Panel estimates. Dependent variable: log household annual food consumption, 1969–2007

Age	0·00447
	(0·0164)
Age squared	0·000280***
	$(2{\cdot}15e-05)$
Number of individuals living in the household	0·226***
	(0·00367)
Number of kids between ages 6 and 17 in the household	−0·136***
	(0·00565)
Number of kids between ages 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	424,656
R^2	0·524
Number of id	15,702

Notes: Standard errors are in parentheses, $^{***}p < 0{\cdot}01$, $^{**}p < 0{\cdot}05$, $^{*}p < 0{\cdot}1$. Data from PSID, each regression also includes yearly dummies.

TABLE 4

2SLS estimates of earnings equation. Dependent variable: Average weekly earnings, 1970–1996

Variables	Professional	Non-professional
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 × 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 2 years ago	−50·66***	−15·84***
	(13·22)	(5·639)
Probability 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	37,618	81,272
R^2	0·258	0·284
Number of id	5600	10,271

Notes: Standard errors are in parentheses, $^{***}p < 0{\cdot}01$, $^{**}p < 0{\cdot}05$, $^{*}p < 0{\cdot}1$. Data from PSID, each regression also includes yearly dummies.

The estimation results in Table 4 reveal comparatively larger returns to experience in professional occupations, faster depreciation of human capital in non-professional occupations, and a larger hiring cost in professional occupations. For example, the return to working full time (40 hours per week) versus part time (20 hours per week) in the previous year is $2751 in professional occupations versus $2284 in non-professional 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 $4502.65 in professional occupations versus $3174.74 in non-professional occupations. The three factors we describe above affect occupational sorting. To see that, consider a man and a woman choosing an occupation, and suppose that the woman's probability of an employment spell interruption and her time spent on non-market activities are larger than the man'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 non-professional occupations. In addition, the depreciation rate of human capital is higher in the non-professional occupations. The returns to labour-market experience, however, are larger in professional occupations. Thus, women, in equilibrium, are more likely to sort into non-professional 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

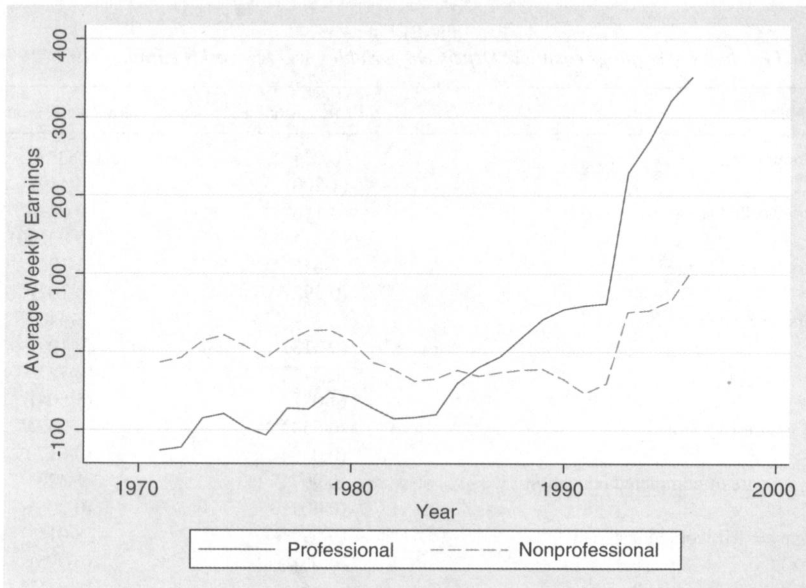

FIGURE 1
Aggregate occupation-specific productivity shocks

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 1980s and a smaller increase in the 1990s in non-professional occupations; these findings are consistent with the increase in the proportion of women working in professional occupations.

6.1.3. Other sources 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 men and women. This effect, however, is not large.

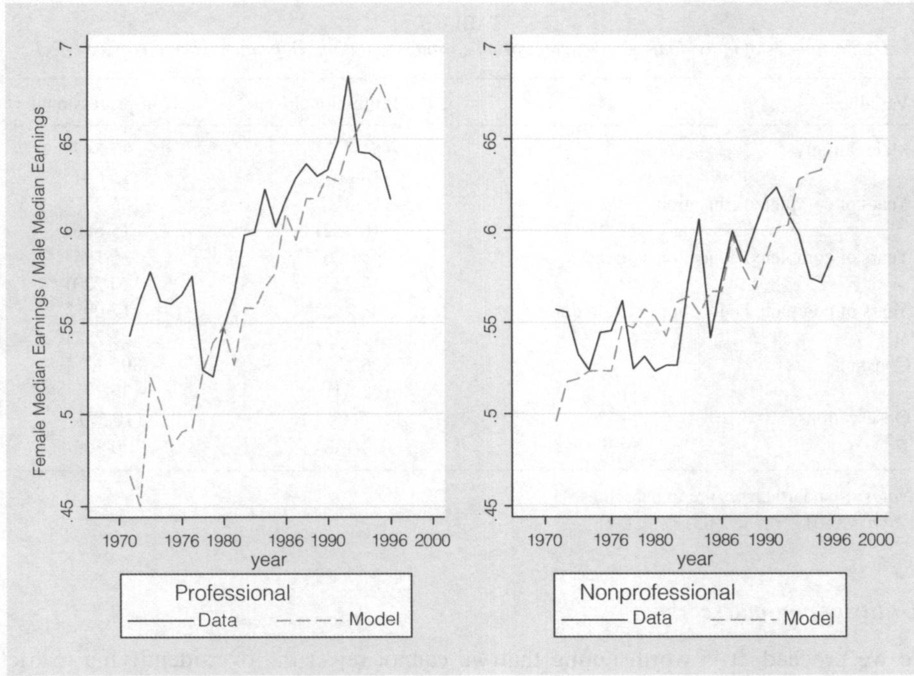

FIGURE 2
Gender earnings gap data and model predicted

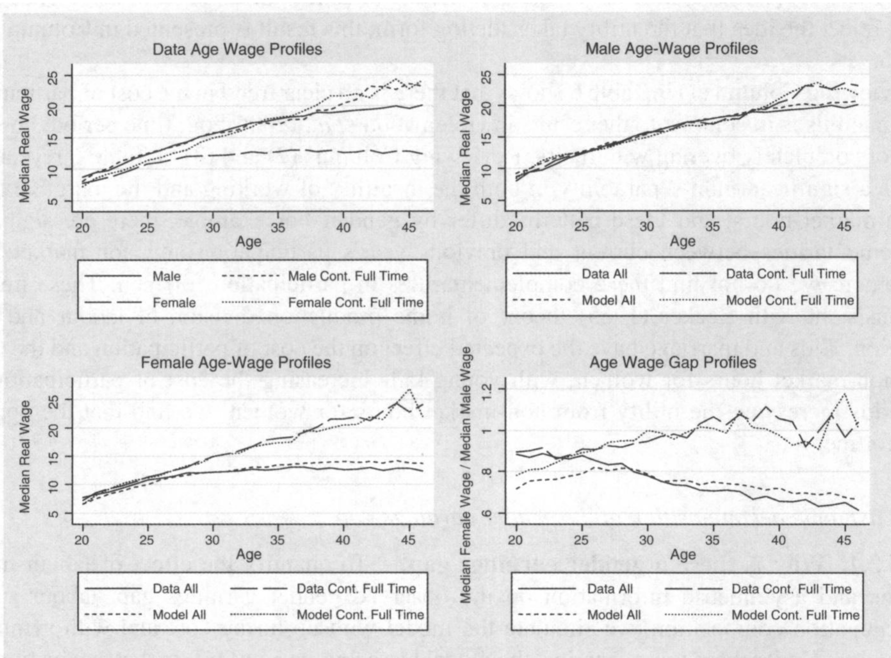

FIGURE 3
Wage gap and wage–age profiles

TABLE 5

OLS estimates of individual-specific effects from earnings equation. Dependent variable: fixed effect

Variables	Professional	Non-professional
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 × male dummy	25·26***	17·09***
	(1·316)	(1·251)
Constant	816·2***	805·1***
	(42·24)	(36·29)
Observations	5600	10,271
R^2	0·138	0·106

Notes: Standard errors are in parentheses.
***$p < 0{\cdot}01$, **$p < 0{\cdot}05$, *$p < 0{\cdot}1$.

6.2. *Utility of non-market time*

Before we proceed, it is worth noting that we cannot reject the over-identifying restrictions test for the participation and hours Euler equations. As in other studies that use the PSID and assume additive separability of consumption and non-market hours (see Zeldes, 1989; Lawrence, 1991; Shea, 1995, among others), our estimate of the elasticity of intertemporal substitution is -1.01 (*i.e.* $-1/(1-\alpha)$), but it is not statistically significantly different from zero. That is, we cannot reject the idea 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 *et al.* (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 (1965) theory of home production division of labour 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. *A dynamic decomposition of the gender earnings gap*

6.3.1. Why is there a gender earnings gap? To quantify the effect of labour-market frictions and asymmetric information on the observed gender earnings gap, labour 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 that 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 labour-market

TABLE 6

GMM estimates utility function. Moment conditions: hours Euler equation and labour-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	$6{\cdot}93\times10^{-05}$
	(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	× male dummy	0·128	× male dummy	−0·0126***
	(0·105)		(0·149)		(0·00228)
1978	−0·454***	Kids age 6 to 17	0·110	Kids age 6 to 17	0·00778***
	(0·104)		(0·0882)		(0·00116)
1979	−0·163*	× male dummy	0·682***	× 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	$-5{\cdot}50\times10^{-06}$*
	(0·0968)		(0·000196)		$(3{\cdot}15\times10^{-06})$
1982	0·0101	Years of education	0·163***	Years of education	−0·00792***
	(0·0971)		(0·0198)		(0·000386)
1983	0·0203	× male dummy	0·224***	× 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**	× male dummy	−15·77***	× 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	× male dummy	0·774***	× 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)		$(5{\cdot}62\times10^{-05})$
1989	−0·338***	× male dummy	−4·846***	× leisure last year	0·000935***
	(0·0906)		(0·167)		$(5{\cdot}67\times10^{-05})$
1990	−0·199**	Part. dummy 2 years lagged	0·0548	× male dummy	0·000971***
	(0·0994)		(0·0894)		$(5{\cdot}80\times10^{-05})$
1991	−0·0233	× male dummy	0·761***	× Leisure 2 years lagged	0·000811***
	(0·103)		(0·172)		$(4{\cdot}49\times10^{-05})$
1992	0·0384			× male dummy	−0·000505***
	(0·107)	Risk aversion	0·01		$(5{\cdot}84\times10^{-05})$
1993	0·281**		(0·012)	Inverse of variance of shock	0·000665***
	(0·121)	Observations	80,986		$(7{\cdot}69\times10^{-05})$

Notes: Standard errors are in parentheses.
***$p<0{\cdot}01$, **$p<0{\cdot}05$, *$p<0{\cdot}1$.
Part. = participation.

differences in the model with no hiring costs are driven by differences in preference and skills only; the difference in the labour-market-outcomes gender gap between the no-hiring-costs scenario and the observed data is therefore due to labour-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

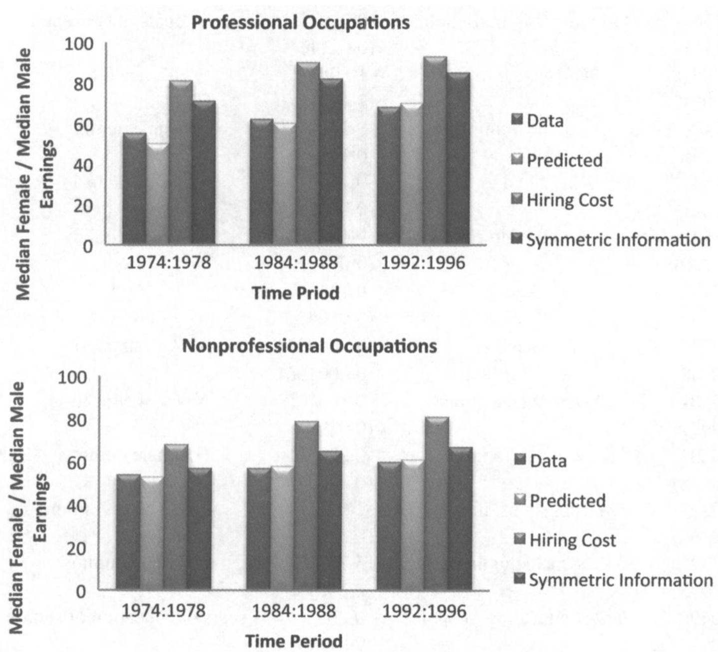

FIGURE 4
Counterfactual decomposition of gender earnings gap

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
Non-professional			
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	1640	2201	1904	2226	1902	2297
Hiring cost	1975	2017	2057	2098	2049	2087
Private information	1813	2088	1980	2090	1988	2120
	Non-professional					
	1974:1978		1984:1988		1984:1988	
	Women	Men	Women	Men	Women	Men
Raw	1424	1998	1635	2117	1773	2184
Hiring cost	1580	2000	1790	2060	1820	2076
Private information	1510	1970	1640	1930	1768	1897

have been 12% and 19%, respectively, instead of the 40% our model predicts.[11] Repeating the exercise for non-professional 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 labour force, but those who participate would have worked more hours. These results suggest that the asymmetric information affects the selection of women into the labour 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 labour-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 labour markets.

6.3.2. 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 in the gender earnings gap; the gender earnings gap would not have declined much if the human-capital gap had 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: 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 the model has a unique equilibrium when there are no hiring costs. When

11. Figure 4 shows that under the 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 labour-market frictions is 28/40. The other calculations are conducted in the same way.

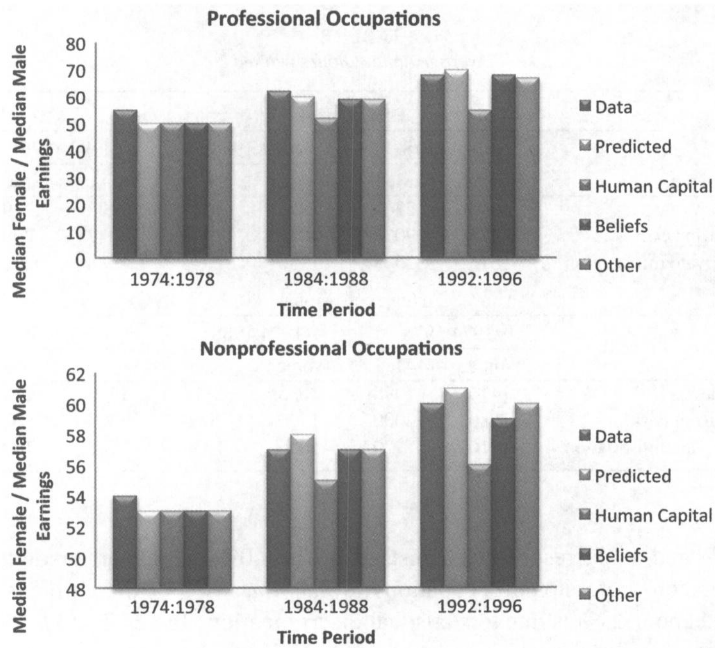

FIGURE 5
Decomposition of the change in gender earnings gap

there are no hiring costs, wages are the marginal productivity, and therefore all the changes in labour-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 labour-market participation; our analysis, therefore, complements the work by Fernandez (2007), Fogli and Vander (2007), and Fernandez and Fogli (2009), which models explicitly the causes of these changes. The results of this decomposition are presented in Table 9. They reveal that labour-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, accounting 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 non-professional occupations. Overall, we find that demand-side factors played a greater role than supply-side factors in the decline of the labour-market-outcomes gender gaps and greater than found in the literature (see Lee and Wolpin, 2010, *e.g.*); the main reason for this difference is that we account for the endogenous effect of discrimination.

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		Non-professional	
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

7. CONCLUSION

This paper finds that the most important factor affecting the gender earnings gap is the gender differences in human capital accumulated both in the labour 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 1970s and the 1990s 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 labour 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 men earn more than women after accounting for statistical discrimination and human-capital differences. This suggests that education choice and education signalling 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.

APPENDIX A: THEORETICAL RESULTS PROOFS

Proof of Proposition 1. We establish optimality 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 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* $h_t > \overline{h}_{t\tau}$.[12]

Lemma 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 1. An increase in salary enters the value function through the Frisch demand for consumption:

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

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 1 (1–2). We begin with the workers' equilibrium strategies. Note 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 behaviour 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 labour-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 labour 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 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 pay-offs beyond working $\overline{h}_\tau$ or $\underline{h}_\tau$, such deviations cannot strictly increase pay-offs.

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 1, workers accept the highest offer; thus, a deviation to a lower salary implies the worker rejects the offer and the pay-off is zero. Note 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, $Q(V_{1it+1}(\omega_{t+1}) - V_{0it+1}(\omega_{t+1}))$, $I_{i\tau t+1}^0(\omega_{t+1})$, and $h_{it+1}(\omega_{t+1})$ also remain the same. To see that the current salary does not change future labour-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-asset-markets assumption, and the assumption that the workers move first (announcing hours), the functions $V_{1it+1}()$ and $V_{0it+1}()$ remain the same.

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

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:

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

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

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

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

Proof of Proposition 1 (3). 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 are a function of the firms' beliefs about the next period's participation. To see this, consider the beliefs about period T's participation:

$$\widetilde{p}_{i\tau T} = \int Q(\omega_T) I^0_{i\tau}(\omega_T) f_{i0}(\omega_T \mid \omega_{T-1}, a_{T-1}) \widetilde{\mu}_{iT-1}(\omega_{T-1} \mid \omega^*_{T-1}, a_{T-1}) d\omega_T. \tag{A.2}$$

Recall that ω_{T-1}, and hence $\widetilde{p}_{i\tau T}$, is a function of the labour-market history $a_0, \ldots, a_{T-1}$. However, a^0_{T-1}, the equilibrium labour-market decisions, are a function of the sequence $\widetilde{p}_{i\tau T}, \ldots, \widetilde{p}_{i\tau 2}$. 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,

$$\begin{aligned}\widetilde{p}_{i\tau T} &= \Gamma_{iT}(\widetilde{p}_{i\tau T}, \ldots, \widetilde{p}_{i\tau 2}), \\ \widetilde{p}_{i\tau T-1} &= \Gamma_{iT-1}(\widetilde{p}_{i\tau T-1}, \ldots, \widetilde{p}_{i\tau 2}), \\ &\vdots \\ \widetilde{p}_{i\tau 2} &= \Gamma_{i2}(\widetilde{p}_{i\tau 2}),\end{aligned} \tag{A.3}$$

where Γ_{it} is the R.H.S. of equation (24). ||

Corollary 1. *A necessary and sufficient condition for existence of equilibrium is that there exists a fixed point in* $\{\widetilde{p}_{i\tau 2}, \ldots, \widetilde{p}_{i\tau T}\}^{\tau \in \Upsilon}_{i \in \{w,m\}}$ *of the system of equations in equation (A.3) for all* τ *and* i.

Proof of Corollary 1. In order to prove this, we first show necessity. Suppose there exists an equilibrium in which equation (A.3) 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 (A.3) 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 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 (A.3).

Proof of Existence. Given the triangular nature of the system of equations in (A.3), 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, $\tilde{p}_{i\tau t+1}$ is the solution to

$$\tilde{p}_{i\tau t+1} = \int_{\omega_t} \left\{ \int_{\omega_{t+1}} f_{i0}(\omega_{t+1} \mid \omega_{t+1}) Q(V_{1it+1}(\omega_{t+1}) - V_{0it+1}(\omega_{t+1})) \right.$$
$$\left. I_{i\tau t+1}(\underline{h}_{\tau t+1} < h^*_{it+1}(\omega_{t+1}) < \overline{h}_{\tau t+1}) d\omega_{t+1} \right\} \tilde{\mu}_{it}(\omega_t \mid \omega^*_t, a_t) d\omega_t. \quad \text{(A.4)}$$

Here, we only make explicit the arguments of interest. Note that $\tilde{p}_{i\tau t+1} : [0,1]$ 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 R.H.S. in $\tilde{p}_{i\tau t+1}$ 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 $I_{i\tau t+1}(\cdot)$ is an indictor function, we can rewrite the inner integral as

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

Since $h_t(\tilde{p}_{i\tau t+1})$ is continuous in $\tilde{p}_{i\tau t+1}$ and $Q(h_t(\tilde{p}_{i\tau t+1}), \cdot)$ 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 $\tilde{p}_{i\tau t+1}$. 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 $h_t(\tilde{p}_{i\tau t+1}, \cdot)$ is continuous in $\tilde{p}_{i\tau t+1}$. Hence, their inverses are continuous in $\tilde{p}_{i\tau t+1}$. 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 (A.4) 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. ||

APPENDIX B: IDENTIFICATION

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:

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

and

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

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

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

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 (36). Similarly, using above *ex ante* function representation, the necessary condition for equilibrium hours can be rewritten as

$$\begin{aligned}\eta^{o}\lambda_t^{o} \sum_{\tau\in\{\mathrm{P,NP}\}} I_{\tau t}\frac{\partial S^{o}_{i\tau t}}{\partial h_t} + \sigma E_t\left[\sum_{s=1}^{3}\beta^{s}(1-p^{(s)}_{i1t})^{-1}\frac{\partial p^{(s)}_{i1t}}{\partial h_t}\right] - z_t' B_{i2} \\ - 2\theta_{i0} l_t + E_t\left[\sum_{s=1}^{2}\theta_{si}(l_{t-s}+\beta^{s})\right] = 0. \qquad \text{(B.2)}\end{aligned}$$

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 (37).

Salary equation identification

Consider the optimal salary under asymmetric information.[13] Combining equations (23) and (29),

$$\begin{aligned}S_\tau(h_{nt},\omega^*_{nt}) &= K_{\tau t} + b_{\tau 1}h_{nt} + b_{\tau 2}h^2_{nt} + \sum_{r=1}^{\rho} b_{\tau 3r}h_{nt-r}\sum_{r=1}^{\rho} b_{\tau 4r}d_{nt-r} + b_{\tau 5}\mathrm{age}_{nt} \\ &+ b_{\tau 6}\mathrm{age}^2_{nt} + b_{\tau 7}\mathrm{age}_{nt}\times \mathrm{education}_n + \nu_n - \gamma_\tau + \beta\gamma_\tau \widetilde{p}_{i\tau t+1}(h_{nt},\omega^*_{nt}). \qquad \text{(B.3)}\end{aligned}$$

Integrating out ν_n gives us

$$\begin{aligned}E_t[S_{i\tau}(h_{nt},\omega^*_{nt}) \mid h_{nt},\{\omega^*_{nt}\backslash\nu_n\}] = K_{\tau t} + b_{\tau 1}h_{nt} + b_{\tau 2}h^2_{nt} + \sum_{r=1}^{\rho} b_{\tau 3r}h_{nt-r} + \sum_{r=1}^{\rho} b_{\tau 4r}d_{nt-r} \\ + b_{\tau 5}\mathrm{age}_{nt} + b_{\tau 6}\mathrm{age}^2_{nt} + b_{\tau 7}\mathrm{age}_{nt}\times\mathrm{education}_n \\ + E_t[\nu_n \mid i, \mathrm{education}_n] \\ - \gamma_\tau + \beta\gamma_\tau E_t[\widetilde{p}_{i\tau t+1}(h_{nt},\omega^*_{nt}) \mid h_{nt},\{\omega^*_{nt}\backslash\nu_n\}], \qquad \text{(B.4)}\end{aligned}$$

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

$$\begin{aligned}S_{i\tau}(h_{nt},\{\omega^*_{nt}\backslash\nu_n\}) = K_{\tau t} + b_{\tau 1}h_{nt} + b_{\tau 2}h^2_{nt} + \sum_{r=1}^{\rho} b_{\tau 3r}h_{nt-r}\sum_{r=1}^{\rho} b_{\tau 4r}d_{nt-r} + b_{\tau 5}\mathrm{age}_{nt} \\ + b_{\tau 6}\mathrm{age}^2_{nt} + b_{\tau 7}\mathrm{age}_{nt}\times\mathrm{education}_n + \overline{\nu}_n \\ - \gamma_\tau + \beta\gamma_\tau \widetilde{p}_{i\tau t+1}(h_{nt},\{\omega^*_{nt}\backslash\nu_n\}). \qquad \text{(B.5)}\end{aligned}$$

13. All the arguments remain the same in the symmetric-information case.

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 (B.5) is a standard panel-data model with unobserved effect $\overline{\nu}_n$. Thus, under standard assumptions, $(K_{\tau t},b_{\tau 1},\dots,b_{\tau 7},\gamma_\tau,\beta\gamma_\tau)$ are identified. Integrating equation (B.3) over all the information except ν_n gives

$$\begin{aligned}E_t[S_{in\tau}\mid \nu_n]=&\,E_t[K_{\tau t}]+b_{\tau 1}E_t[h_{nt}\mid \nu_n]+b_{\tau 2}E_t[h^2_{nt}\mid \nu_n]+\sum_{r=1}^{\rho}b_{\tau 3r}E_t[h_{nt-r}\mid \nu_n]\\&+\sum_{r=1}^{\rho}b_{\tau 4r}E_t[d_{nt-r}\mid \nu_n]+b_{\tau 5}E_t[\text{age}_{nt}\mid \nu_n]+b_{\tau 6}E_t[\text{age}^2_{nt}\mid \nu_n]\\&+b_{\tau 7}E_t[\text{age}_{nt}\times \text{education}_n\mid \nu_n]+\nu_n-\gamma_\tau+\beta\gamma_\tau E_t[\widetilde{p}_{i\tau t+1}\mid \nu_n].\end{aligned} \tag{B.6}$$

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

APPENDIX 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 labour 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 family unit. If an individual has never had any children, one record indicates that report. Note that *eligible* refers to 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 4 years or less. We then drop all individuals who were older than 65 years 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 4 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 \$9999.00. 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, and 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 Labourers 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 non-professional 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; Labourers, Except Farm; and Service Workers, Except Private Household.[14]

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.

Acknowledgements. 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, Stephen 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 Behaviour 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.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

REFERENCES

AGUIRREGABIRIA, V. and MIRA, P. (2007), "Sequential Estimation of Dynamic Discrete Games", *Econometrica*, **75**, 1–53.

ALBANESI, S. and OLIVETTI, C. (2008), "Gender Roles and Medical Progress" (Working Paper NBER WP 13179, Columbia University).

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

ALTONJI, J. (2005), "Employer Learning, Statistical Discrimination and Occupational Attainment", *American Economic Review*, **95**, 112–117.

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

ALTONJI, J. and BLANK, M. (1999), "Race and Gender in the Labor Market", in Ashenfelter, O. and Card, D. (eds) *Handbook of Labor Economics*, Vol. 3 (Amsterdam: Elsevier) 3143–3259.

ALTONJI, J., HAYASHI, F. and KOTLIKOFF, L. (1996), "Risk Sharing Between and Within Families", *Econometrica*, **64**, 261–294.

ALTUG, S. and LABADIE, P. (1994), *Dynamic Choice and Asset Markets* (San Diego, CA: Academic Press).

ALTUG, S. and MILLER, R. A. (1990), "Household Choices in Equilibrium", *Econometrica*, **58**, 543–570.

ALTUG, S. and MILLER, R. A. (1998), "The Effect of Work Experience on Female Wages and Labour Supply", *Review of Economic Studies*, **65**, 45–85.

ANTONOVICS, K. (2004), "Statistical Discrimination and Intergenerational Income Mobility" (Mimeo, University of California, San Diego).

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

ATTANASIO, O., LOW, H. and SANCHEZ-MARCOS, V. (2008), "Explaining Changes in Female Labor Supply in a Life-Cycle Model", *American Economic Review*, **98**, 1517–1552.

BAJARI, P., BENKARD, L. and LEVIN, J. (2007), "Estimating Dynamic Models of Imperfect Competition", *Econometrica*, **75**, 1331–1370.

BARRON, J., BLACK, D. and LOWENSTEIN, M. (1993), "Gender Differences in Training, Capital and Wages", *Journal of Human Resources*, **28**, 343–364.

BECKER, G. (1965), "A Theory of Allocation of Time", *Economic Journal*, **75**, 496–517.

BECKER, G. (1971), *The Economics of Discrimination*, 2nd edn (Chicago: The University of Chicago Press).

BECKER, G., GROSSMAN, M. and MURPHY, K. (1994), "An Empirical Analysis of Cigarette Addiction", *American Economic Review*, **84**, 396–418.

BOWLUS, A. J. (1997), "A Search Interpretation of Male–Female Wage Differentials", *Journal of Labor Economics*, **15**, 625–657.

BROWNING, M. and LUSARDI, A. (1996), "Household Saving: Micro Theories and Micro Facts", *Journal of Economic Literature*, **34**, 1797–1855.

CARD, D. (1990), "Labour Supply with a Minimum Threshold", *Carnegie–Rochester Series on Public Policy*, **33**, 137–168.

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

CHAMBERLAIN, G. (1984), "Panel Data", in Griliches, Z. and Intriligator, M. D. (eds) *Handbook of Econometrics*, Vol. 2 (Amsterdam: North-Holland) 1247–1318.

CHESHER, A. (2007), "Instrumental Values", *Journal of Econometrics*, **139**, 15–34.

COATE, S. and LOURY, G. (1993), "Will Affirmative Action Policies Eliminate Negative Stereotypes?", *American Economic Review*, **83**, 1220–1240.

DIONNE, G. and DOHERTY, N. A. (1994), "Adverse Selection, Commitment, and Renegotiation: Extension to and Evidence from Insurance Markets", *Journal of Political Economy*, **102**, 209–235.

ECKSTEIN, Z. and WOLPIN, K. (1989), "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience", *Review of Economic Studies*, **56**, 375–390.

EROSA, A., FUSTER, L. and RESTUCCIA, D. (2005), "A Quantitative Theory of the Gender Gap in Wages" (Working Paper No. 05-09, Federal Reserve Bank of Richmond, September).

FERNÁNDEZ, R. (2007), "Culture as Learning: The Evolution of Female Labor-Force Participation over a Century"(Working Paper NBER No. 13373, New York University).

FERNÁNDEZ, R. and FOGLI, A. (2009), "Culture: An Empirical Investigation of Beliefs, Work, and Fertility", *American Economic Journal: Macroeconomics*, **1**, 146–177.

FLABBI, L. (2009), "Prejudice and Gender Differentials in the U.S. Labor Market in the Last Twenty Years", *Journal of Econometrics*, **156**, 190–200.

FLABBI, L. (2010), "Gender Discrimination Estimation in a Search Model with Matching and Bargaining", *International Economic Review*, **51**, 745–783.

FOGLI, A. and VELDKAMP, L. (2007), "Nature or Nurture? Learning and the Geography of Female Labor-Force Participation"(Working Paper No. 14097, National Bureau of Economic Research).

GAYLE, G-L and MILLER, R.A. (2004), "Life-Cycle Fertility Behavior and Human Capital Accumulation", (Tepper School of Business working paper NO 2004-E16, Carnegie Mellon University).

GREENWOOD, J., SESHARDRI, A. and YORUKOGLU, M. (2005), "Engines of Liberation", *Review of Economic Studies*, **72**, 109–133.

HECKMAN, J. J. and MACURDY, T. E. (1980), "A Life-Cycle Model of Female Labour Supply", *Review of Economic Studies*, **47**, 47–74.

HELLWIG, M. (1987), "Some Recent Developments in the Theory of Competition in Markets with Adverse Selection", *European Economic Review*, **31**, 319–325.

HOTZ, J., KYDLAND, F. and SEDLACEK, G. (1988), "Intertemporal Preferences and Labor Supply", *Econometrica*, **56**, 335–360.

HOTZ, V. J. and MILLER, R. A. (1993), "Conditional Choice Probabilities and the Estimation of Dynamic Models", *Review of Economic Studies*, **60**, 497–529.

JONES, L., MANUELLI, R. and MCGRATTAN, E. (2003), "Why Are Married Women Working So Much?" (Report No. 317, Federal Reserve Bank, Minneapolis).

LAFFONT, J. J. and TIROLE, J. (1988), "The Dynamics of Incentive Contracts", *Econometrica*, **56**, 1153–1175.

LAWRENCE, E. (1991), "Poverty and the Rate of Time Preference", *Journal of Political Economy*, **99**, 54–77.

LEE, D. and WOLPIN, K. (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.

MACE, B. (1991), "Full Insurance in the Presence of Aggregate Uncertainty", *Journal of Political Economy*, **99**, 928–956.

MINCER, J. and POLACHEK, S. W. (1974), "Family Investments in Human Capital: Earnings of Women", *Journal of Political Economy*, **82**, 76–108.

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.

MORO, A. and NORMAN, P. (2004), "A General Equilibrium Model of Statistical Discrimination", *Journal of Economic Theory*, **114**, 1–30.

MUNDLAK, Y. (1961), "Empirical Production Function Free of Management Bias", *Journal of Farm Economics*, **43**, 44–56.

MUNDLAK, Y. (1978), "On the Pooling of Time Series and Cross-section Data", *Econometrica*, **46**, 69–85.

PAGAN, A. and ULLAH, P. (1999), *Nonparametric Econometrics* (London: Cambridge University Press).

PAKES, A., OSTROVSKY, M. and BERRY, S. (2007), "Simple Estimators for the Parameters of Discrete Dynamic Games", *The Rand Journal of Economics*, **38**, 373–399.

PESENDORFER, M. and SCHMIDT-DENGLER, P. (2003), "Identification and Estimation of Dynamic Games" (Mimeo, London School of Economics).

PESENDORFER, M. and SCHMIDT-DENGLER, P. (2008), "Asymptotic Least Squares Estimators for Dynamic Games", *Review of Economic Studies*, **75**, 901–928.

PHELPS, E. (1972), "The Statistical Theory of Racism and Sexism", *American Economic Review*, **62**, 659–661.

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

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

RUST, J. (1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher", *Econometrica*, **55**, 999–1033.

SHEA, J. (1995), "Union Contracts and the Life-Cycle Permanent Income Hypothesis", *American Economic Review*, **95**, 186–200.

TIROLE, J. (1996), "A Theory of Collective Reputation", *Review of Economic Studies*, **63**, 1–22.

TOWNSEND, R. (1994), "Risk and Insurance in Village India", *Econometrica*, **62**, 539–591.

ZELDES, S. P. (1989), "Consumption and Liquidity Constraints: An Empirical Investigation", *Journal of Political Economy*, **97**, 3005–3046.