Estimating a Dynamic Adverse-Selection Model: Labor-Force Experience and the Changing Gender Earnings Gap 1968-93.*

George-Levi Gayle and Limor Golan

Tepper School of Business, Carnegie Mellon University.

Previous Version: August 2006 Current Version: August 2007

Abstract

This paper formulates and estimates a dynamic model of labor supply, occupational sorting, human capital accumulation and discrimination to explain the narrowing gender earnings gap from 1968 to 1993. The paper proves the model is identified and develops a three-step estimation technique. Imperfect information significantly amplifies exogenous shocks: statistical discrimination accounts for 36 percent of the observed gender earnings gap in the mid-to-late 1970s, declining to 22 percent in the mid-to-late 1980s. Differences in preferences are comparatively less important: the gap would have been at least 56 percent smaller in the mid-to-late 1970s and would have nearly closed by the mid-to-late 1980s if it was driven only be preference. Increases in overall productivity and demographic changes account for a large percentage of the decline in the gender earnings gap and the increase in female labor market experience, while a relative increase in productivity raises women's representation in professional occupations.

*We thank Zvi Eckstein, Larry Jones, John Kennan, Kevin Lang, Robert A. Miller, Derek Neal, Andrew Postlewaite, Victor Rios-Rull, Robert Shimer, Steve Spear, Yoram Weiss, the participants of SOLE 2005, Tel-Aviv Summer Workshop 2005, SITE Theory-Based Micro-Econometric Modelling Workshop 2006, Econometric Society 2007, NBER Summer Workshop 2007, SITE Family Behavior and the Aggregate Economy Workshop 2007 and seminar participants at Purdue University, University of Minnesota, Carnegie Mellon, University of Pennsylvania, New York University, Boston University, Northwestern University and The Federal Reserve Bank of Chicago for comments and suggestions.

Keywords: Adverse selection, Dynamic general equilibrium, Occupation sorting, Labor market experience, Gender earning gap, Discrimination, Semi-parametric estimation, Structural estimation, Estimation of dynamic games, Kernel density.

1 Introduction

One of the most striking changes in the U.S. labor market over the last four decades is the significant reduction in the gender wage gap in the 1970's and 1980's. In 1968, the unconditional median gender wage differential was about 40%; this gap was reduced to around 28% by 1992. Accompanying this decline were significant changes in labor-market attachment of women. According to figures from the Michigan Panel Study of Income Dynamics (PSID), the participation rate of women has increased from 54% in 1968 to 74% in 1992. The annual hours worked by women also increased over the period, from 1400 hours in 1968 to 1800 in 1992. Concurrently, there were little or no changes in the figures for men. Furthermore, women were entering in greater numbers the traditionally male dominated occupations. In the PSID, the percentage of women holding professional jobs went from 28% in 1968 to 43% in 1992.[1]

These significant changes prompt the question: What are the main driving forces of these changes in gender patterns of labor-market outcomes? This paper seeks to answer this question. It investigates the roles of labor-market attachment, on-the-job human-capital accumulation, occupational sorting, preferences and discrimination in the above mentioned changes in gender labor-market outcomes. The main challenge in quantifying these effects is to account for the endogeneity of labor supply, discrimination, and earnings.[2] This paper contributes in three ways: First, it formulates a dynamic model of labor-supply, occupational sorting, and human-capital accumulation in which gender discrimination and an earnings gap arise endogenously. Second, it demonstrates identification of such models and develops a three-step estimation technique. Third, it uses the framework to quantify the main driving forces behind the changes in labor-market gender patterns.

[1]Many papers document the changes in the gender wage gap, occupational composition, and patterns of participation, including Blau and Kahn (1997), Lewis (1996), and Eckstein and Nagypal (2004), among others.

[2]By earnings we refer to annual earnings. The earnings gap is larger than the wage gap and the declining trend is roughly the same in both. We therefore, focus on the earnings gap instead of the traditional wage gap.

The model emphasizes the interaction of future expected length of employment spells with past labor market experience in the presence of private information and uncertainty; this interaction may give rise and amplify differences in life-cycle human capital investment patterns of women and men. It differs from the standard incomplete information discrimination models (see Arrow, 1972; Phelps, 1972; Coate and Loury, 1993; Moro, 2003; Antonovics, 2004; Altonji, 2005; among others) in that it focuses on the uncertainty about the turnover propensity of workers instead of uncertainty about productivity differences across groups. This element of our model is consistent with regularities found in several papers that empirically examine the relationship between the gender earnings gap and the length of employment spells. For example, Light and Ureta (1992) found that, controlling for all observable characteristics, women have a higher probability of changing employers than men. They found that 94% of women are movers for unobserved reasons while only 23% of men are movers for unobserved reasons (see Light and Ureta,1992, pp.168). Moreover, Light and Ureta (1995) find that 12% of the male-female wage gap is due to differences of timing of experience, and career interruptions. Finally, Altonji and Paxton (1992) found that job mobility is strongly linked to changes in hours; women who face changes in family responsibility adjust their hours, and this may lead to lower earnings.

In the model workers make participation, occupation, hours, and consumption decisions each period. They are heterogeneous with respect to characteristics affecting the disutility of working; these characteristics are stochastic, but persistent over time. The labor market consists of different occupations which are characterized by different returns to experience, depreciation rate of human capital, and costs of hiring a new worker. Because of the asymmetric information and incompleteness of labor market contracts, there are non-trivial dynamic effects on labor supply behavior which interact with the standard effect of on-the-job training ((Mincer and Polachek (1974)). Expectations of future family structure and market activities of women and men are therefore important determinants of the level and type of current labor supply (similar to Polachek (1981)). In addition, private information implies that the optimal contract induces signaling because individual labor-supply decisions provide information on the worker's type. In equilibrium, information on workers is revealed gradually over time, and employers update beliefs on workers' types based on observed differences in workers' labor-market history.

The equilibrium in our model may exhibit aspects of statistical discrimination. In the absence of private information, differences in disutility from working and skills can induce differences in men's and women's choices of participation, occupation, and hours worked hence giving rise to an earnings gap. However, even if there are no differences in preferences and skills across gender groups, different outcomes may arise due to the multiplicity of equilibria when there is private information. If employers have different beliefs about future participation of men and women, women may face lower earnings. Women may work less and sort into occupations with lower cost of intermittence, on average accumulating less labor-market experience than men. Because of the different channels that gender earnings gaps can arise in our model, we define discrimination as the difference between the labor-market outcomes of men and women under symmetric information and under private information. That is, gender differences in earnings that arise due to observed *group affiliation* are referred to as discrimination, as opposed to differences that arise due to differences in preferences and productive skills.

A recent paper by Erosa, Fuster and Restuccia (2005), also emphasizes one of the three motives for human capital investment in our model, that is, the effect of expectations of future family structure and labor-market activities. These two models have strikingly different predictions about the return to actual experience between men and women. Erosa et al.(2005)'s model predicts that women should have lower return to actual experience than men while our model predicts women should have higher return to actual experience than men. This prediction of our model is consistent with the empirical findings of Weinberger and Kuhn (2006), that for a given cohort of full-time workers the gender wage gap is U-shaped in potential experience. These differences may lead to considerable different conclusion in trying to ascertain the sources behind the narrowing of the gender earnings gap. Consider an increase in labor market productivity. In Erosa et al.(2005)'s model this would have a bigger impact on men than women since men's human capital has a higher return in the market and their value of home work is less than women. In our model this would have a bigger impact on women than men since they are working at the lower margin and this would lead to an increase in the employers beliefs. Hence, as will be seen in the results of this paper, this would lead one to attribute more of the decline in the earnings gap to fertility decline, for example, as opposed to an increase market productivity.

Another paper that has an endogenous gender wage gap is Albanesi and Olivetti (2005). It develops a one-period model of statistical gender discrimination in which effort in the labor market and hours worked at home are determined endogenously. The agency problem in the labor market can endogenously generate gender discrimination. In equilibrium, women may have higher costs of effort as it is optimal to work more hours at home.[3]

The literature focuses on several factors that may have caused the changes of earnings gap during 1980's. We explore which of these factors drove the changes in relative earnings of men and women, and quantify their relative importance. The first factor is technological changes in the labor market, which we model as occupation-specific changes in productivity. The second factor is a decline in costs of producing home goods. The third factor is changes in education, marriage, and fertility over time. The later may drive changes in labor supply behavior because they affect the disutility of working and skill in the labor market. All three may affect beliefs, if women's employment spells are shorter than men's, may driving a decline in the gender earnings gap.

One of the main difficulties in estimating game theoretic models is the possibility of multiple equilibria. We use a multistage estimation procedure—based on necessary conditions for all equilibria—to overcome the indeterminacy caused by multiplicity of equilibria. Assuming the data per cohort (or other data partition) is generated by a single equilibrium, this procedure is free of the indeterminacy problem. Other papers taking this approach are, among others, Tamer (2003); Aguirregabiria (2005); Aguirregabiria and Mira (2007); Pesendorfer and Schmidt-Dengler (2003, 2006); Bajari, Benkard, and Levin (forthcoming); Pakes, Ostrovsky, and Berry (forthcoming).

This paper then uses a constructive strategy to show that the model of discrete and continuous choice is semiparametrically identified (for examples of papers using this approach see Chesher (2003, 2005, 2007) and Pesendorfer and Schmidt-Dengler (2006)). That is, conditional on the distribution of the unobserved preference shocks and the class of risk aversion, our model is nonparametrically identified up to two additive normalizations. This result is stronger than that obtained in the standard discrete-choice model (see for example Magnac and Thesmar, 2002, and Pesendorfer

[3]Baron, Black, and Lowenstein (1993) developed a model in which employers expect women to have a higher turnover rate and give women lower training levels, explaining the lower wages in a non game theoretical setting.

and Schmidt-Dengler, 2006), but has parallels in the literature on continuous-choice models (see Jofre-Bonet and Pesendorfer, 2003, for example).

The estimation proceeds in three steps. The employer's problem is estimated in the first step, along with other inputs from the individual consumption Euler equation. The estimates from the first stage are used in the second step to nonparametrically estimate each individual's choice-specific probabilities and their derivatives, as well as the employer's beliefs about workers' future participation probabilities. In the final step, these estimates are combined with a tractable alternative representation of the agents' choice-specific valuation functions to form moment conditions to estimate the structural parameters of the agents' utility functions. Our estimator is akin to a number of estimators in the literature for the estimation of discrete games and single-agent models (Hotz and Miller, 1993; Altug and Miller, 1998; Pakes, Ostrovsky, and Berry, forthcoming; Pesendorfer and Schmidt-Dengler, 2003; Bajari, Benkard, and Levin, forthcoming); our estimator is different, however, in that we are estimating a dynamic adverse-selection model with both discrete and continuous controls without any simulation. To the best of our knowledge, this is the first paper to estimate a structural dynamic signaling model.

We find that increase in female labor market experience and hours worked accounts for 67% of the change in the earnings gap in professional occupations. The decomposition of the change in labor market experience of women in professional occupations reveals that 18% of the change in female labor market experience and hours worked was driven by technological changes (see Lee and Wolpin (forthcoming) for similar results), while demographic changes (mainly fertility decline) accounted for about 28% of the change. Similar results were found for the nonprofessional occupations. The estimation results do not support the hypothesis that changes in home-production technology explain the increase in women's labor-market experience and hours worked (see similar findings in Jones, Manuelli, and McGrattan, 2003).

Further analysis shows that market frictions significantly amplify exogenous changes in our model. Without labor-market frictions, the earnings gap is mainly driven by preferences differences; these differences drive relatively small portion of the observed gender earnings gap. In the mid-to late 70's the gap would have been at least 56% smaller if it was driven by preferences differences only, and 43% in the mid to late 80's. By the mid to late 80's, however, almost no gap would have remain, the median earnings of women over the median earnings of men would have been 96%. Discrimination accounts for 36% of the observed gender earnings gap in the mid-to late 70's declining to 22% in the mid-to late 80's. We find that the significantly higher level and growth rate of aggregate productivity in professional occupations account for much of the increased representation of women in these occupations.

This paper is organized as follows. Section 2 gives an overview of the data. Section 3 describes the model. Section 4 contains the equilibrium analysis and discusses the theoretical model's predictions for the labor market gender gaps. The identification results are presented in Section 5, while the empirical strategy is outlined in Section 6. Section 7 contains the estimation results, empirical analysis, and the gender–earnings-gap decomposition. Section 8 concludes. The appendices present proofs, the implementation and asymptotic properties of our estimator, and a detailed data description.

2 Data

The data for this study 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 nonprofessional. We only keep white individuals between the ages of 25 and 65 in our sample. After eliminating missing values, we are left with 5,978 individuals over the years 1968 to 1992 of which 46% are women. The construction of our sample and the definitions of the variables are described in greater detail in Appendix A

The average annual earnings for men increased by roughly 18% over the period, from US $40,000 per year in 2000-constant dollars in 1968 to US$47,000 in 1992. Meanwhile, the average annual earnings for women increased by around 49% over the same period, from US$16,200 in 1968 to US%24,100 in 1992. As Figure 1 shows, the wage gap declined by around 30% over the period, going from around 40% in 1968 to around 28% in 1992. At the same time the earnings gap declined by around 25%. The earning gap is 50% larger than the wage gap. Therefore, in this paper, we focus on the earnings gap.

Table 1 contains summary statistics of our main labor-market and human-capital variables. The participation rate for men is relatively constant over the sample period with a slight decline toward the end. In contrast, the participation rate for women increased significantly, starting at around 54% in 1968 and increasing to 74% by 1992. The average annual hours worked by men is also relatively constant, but the average annual hours worked by women increased by roughly 30%, from around 1,400 hours per year in 1968 to 1,800 per year in 1992. Although the hours-worked gap between women and men has narrowed significantly over the period, it remains large. The gap in the average years of completed education between men and women was almost completely closed by the early 1990s, yet the participation and hours gaps remain.

The percentage of women in the professional occupations increased by roughly 54% over the sample period, going from 28% of the occupation in 1968 to around 43% of the occupation by 1992. At the same time the fraction of women in the nonprofessional occupations increased by around 10% over the period, going from 45% in 1968 to around 50% in 1992. In addition, the earning gaps are significantly different in the two occupations. In the professional occupations, the earnings gap is much smaller than in the nonprofessional occupations. Also the earnings gap is narrowing at a faster rate in the professional than the nonprofessional occupations.[4]

Table 2 contains summary statistics of our main demographic and wealth variables. The sample includes age distribution of the number of children and household size which have both declined; the decline is most pronounced amongst young children. Roughly 80% of our final sample was married throughout the period. Therefore, the major demographic change that occurred over the period is a decline in the fertility rate.

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

[4]With finer occupation classification, Dey and Hill (2007) found that in some professional occupations women earn more than men.

3 Theoretical Model

The economy consists of infinitely lived firms and finitely lived workers. There exists a continuum of workers on the unit interval $[0, 1]$ in each age–education cohort. These workers are divided into two observed gender groups, $i \in \{w, m\}$, women and men, respectively. For notational ease, we will denote age and calendar year for each cohort by t $(t = 0, \ldots, T)$. The theoretical model is written and solved for a given cohort. Labor markets are competitive with free entry. There are Υ occupations, $\tau = 1, \ldots, \Upsilon$, each consists a continuum of identical firms. Firms offer jobs to maximize lifetime expected discounted profits; each firm offers one job each period. There is a homogeneous product with price normalized to 1.

3.1 Workers' Preferences and Choice

Workers have preferences over non-market hours, l_t, and consumption, c_t. Let h_t denote the time spent working at t. There is a fixed amount of time in each period available for working which we normalized to one, therefore $0 \leq h_t \leq 1$, thus $l_t = 1 - h_t$. Two additional indicators are defined: a work force participation indicator d_t, where $d_t = 1$ if $h_t > 0$ and 0 otherwise, and an occupation participation indicator $I_{\tau t}$, where $I_{\tau t} = 1$ if the worker is employed in occupation τ and 0 otherwise. Let a_t be the worker's set of labor-market actions: $a_t \equiv (d_t, \{I_{\tau t}\}_{\tau=1}^{\Upsilon}, h_t)$. Define the employment history of a worker by $H_{t-1} = (a_0, a_1, \ldots, a_{t-1})$. Therefore, the worker's next employment history, H_t, is $H_t = (H_{t-1}, a_t)$.

Each worker's preferences are additively separable in consumption and leisure contemporaneously. Consumption is additively separable over time, whereas leisure is not. It is assumed that there are two time-varying vectors of characteristics that determine the utility associated with alternative consumption and leisure allocations. Denote the first by the $K \times 1$ vector z_t which includes gender and the second by the 3×1 vector $(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})'$. It is assumed that z_t is independently distributed over the population. It evolves according to a known group-specific transition distribution function, $F_{i0}(z_{t+1} \mid z_t, H_t)$. The vector $(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})'$ is assumed to be independent across the population and time and drawn from a population with a common distribution function $F_1(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})$. The distribution function $F_{i0}(z_{t+1} \mid z_t, H_t)$ and $F_1(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})$ are absolutely continuous with continuously differentiable density $f_{i0}(z_{t+1} \mid z_t, H_t)$ and $f_1(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})$, respectively. There are several things to notice: First, $F_{i0}(z_{t+1} \mid z_t, H_t)$

depends on H_{t-1} and a_t but not on c_t. This property follows directly from the definition of time-additive separability of consumption. Second, although we assume that characteristics in z_t are contemporaneously not a function of the work decision, $F_{i0}(z_{t+1} \mid z_t, H_t)$ is a function of the worker's labor-market actions a_t. Therefore, those characteristics are endogenous in a predetermined sense.

The current-period utility function at date t for an individual i is defined as

$$U_{it} \equiv d_t u_{0it}(z_t) + u_{1i}(l_t, z_t, H_{t-1}) + u_{2i}(z_t, c_t, \varepsilon_{2t}) + (1 - d_t)\varepsilon_{0t} + d_t \varepsilon_{1t}, \tag{1}$$

where u_{it0} represents the fixed utility costs of working. The utility from non-market hours, u_{1i}, is additive and nonseparable over time. We assume the utility function is concave, continuous, and twice differentiable everywhere in l_t and c_t.

Let $\beta \in (0, 1)$ denote the common subjective discount factor, and $E_t(.)$ denotes the expectation conditional on information available to the individual at period t. The expected lifetime utility of the individual is then

$$E_t \left[\sum_{r=t}^{T} \beta^{r-t} U_{it} \right]. \tag{2}$$

3.2 Budget Constraint and Consumption

To provide a tractable solution to the model, we assume that asset markets are competitive. Here the word *competitive* is synonymous with price-taking behavior. In addition, we assume that there are no frictions in the markets for loans and borrowers and lenders face a common interest rate. A rich set of financial securities exists to hedge against uncertainty in a market with anonymous trades.

The above assumptions allow us to compactly write the lifetime individual budget constraint. These assumptions imply that individuals can condition their choices at time t on information that is publicly available at that time, and can purchase contingent claims to consumption that pay off in each state of the world. Therefore the workers' budget constraint in each period can be written as

$$E_0 \left\{ \sum_{t=0}^{T} \beta^t \lambda_t \left[c_{wt} + c_{mt} - \overline{S}_t \right] \right\} \leq W, \tag{3}$$

where $\overline{S}_t$ is the total household labor-market income (if the individual is single, then it is only one income, but if the individual is married, it is the sum of two incomes), λ_t is the expected price of the contingent claim, and W is an exogenously determined quantity denoting bequests net of inheritances. See Altug and Labadie (1994) pp. 305 for a formal derivation of the budget constraint under these assumptions.

Let η_i denote the Lagrange multiplier associated with the budget constraint in equation (3) for each of the household member. The first order conditions with respect to individual consumption c_{it} are,

$$\frac{\partial u_{2i}(c_{it}, z_t, \varepsilon_{2t})}{\partial c_{it}} = \eta_i \lambda_t, \tag{4}$$

for all $t \in \{0, 1, 2 \ldots\}$ and $i \in \{w, m\}$.

With the contemporaneous separability of the consumption from leisure, the condition in (4) can be used to solve for the individual Frisch demand functions, which determine consumption in terms of the time-varying characteristics, z_t, the idiosyncratic shocks to preferences, ε_{2t}, and the shadow value of consumption, $\eta\lambda_t$. Formally, the optimal consumption allocation of individual i at each date t can be expressed as $c_{it}^o = c_i^o(z_t, \varepsilon_{2t}, \eta\lambda_t)$. Therefore, the Frisch demand for consumption is a sufficient statistics for the optimal consumption behavior.[5]

3.3 Technology

Let $z_t^p \subset z_t$ be a vector of individual characteristics that affect productivity. Output in period t and in occupation τ is denoted by $y_{\tau t} = y_{\tau t}(h_t, H_{t-1}, z_t^p)$. The production function is identical within an occupation but varies across occupations. Output is a function of current hours input, the worker's past labor-market experience, and the worker's other production-relevant characteristics. Assume that gender does not affect output, that is, a woman and a man with the same labor-market experience and other production characteristics produce the same level of output if they supply the same number of hours. Past labor-market experience augments the output produced per unit of hours input. Labor-market experience is general but the returns to experience vary across occupations.

There are costs to the employer when a new worker is hired. These costs are

[5] Other papers that use similar assumptions include Altug and Miller (1990, 1998); Card (1990); Mace (1991); Townsend (1994); Altonji, Hayashi, and Kotlikoff (1996); Miller and Sieg (1997); and Gayle and Miller (2004).

specific to the employers. Within occupations, the costs are the same for all employers. We denote employer's cost of hiring a new employee (hiring costs) in occupation τ by γ_τ. Hiring cost captures all possible training, administrative and other net costs that accrued in the first period of hiring a new worker.

Workers and firms can only commit to (noncontingent) spot contracts.[6] Firms hire a worker for a job (a job is equivalent to fraction of time worked) only if, given the worker's characteristics, the output he/she produces net of the costs of hiring within the occupation is the highest among all occupations. This is formalized in the assumption below.

Assumption 3.1 *A firm in occupation τ offers a contract for hours h to a worker with experience H_{t-1} and characteristics z_t^p if there exists no other occupation τ' such that*

$$y_{\tau' t}(h_t \mid H_{t-1}, z_t^p) - \gamma_{\tau'} > y_{\tau t}(h_t \mid H_{t-1}, z_t^p) - \gamma_\tau. \tag{5}$$

If in addition to Assumption 3.1, the production function satisfies a single-crossing condition, then occupations would be segmented into jobs (hours) offered to workers conditional on their characteristics affecting production. This property is formalized in Assumption 3.2.

Assumption 3.2 (Single crossing) *For any two occupations τ' and τ, $y_{\tau' t}(h_t \mid H_{t-1}, z_t^p)$ and $y_{\tau t}(h_t \mid H_{t-1}, z_t^p)$ may not cross more than once. Also, $y_{\tau t}(h_t \mid H_{t-1}, z_t^p)$ is twice continuously differentiable.*

Let $\tau - 1$ (or $\tau + 1$) denote the occupation of the worker if he/she chooses hours below (or above) $\overline{h}_{\tau t}(H_{t-1}, z_t^p)$ (or $\underline{h}_{\tau t}(H_{t-1}, z_t^p)$). These hours are determined by the following conditions,

$$\begin{aligned} (y_{\tau t}(\underline{h}_{\tau t}(H_{t-1}, z_t^p)) - y_{\tau-1t}(\underline{h}_{\tau t}(H_{t-1}, z_t^p)) - \gamma_\tau + \gamma_{\tau-1} &= 0 \\ y_{\tau t}(\overline{h}_{\tau t}(H_{t-1}, z_t^p)) - y_{\tau+1t}(\overline{h}_{\tau t}(H_{t-1}, z_t^p)) - \gamma_\tau + \gamma_{\tau+1} &= 0 \end{aligned}$$

[6] A more realistic assumption is that firms can commit to long-term contracts, but workers cannot. The main feature, that contracts do not fully separate workers in such a framework, can be maintained. See Dionne and Doherty (1994) for a derivation of an optimal renegotiation-proof contract with semi-commitment in a dynamic adverse-selection model with partial commitment.

The single-crossing condition in assumption 3.2 ensures a connected set of occupation segments while the continuous differentiability of the production function ensures that $\overline{h}_{\tau t}(H_{t-1}, z_t^p)$ (or $\underline{h}_{\tau t}(H_{t-1}, z_t^p)$) exists.

Assumptions 3.1 and 3.2 imply that for every occupation τ and worker with characteristics H_{t-1}, z_t^p, there is a range of hours in occupation τ, $h \subset (\underline{h}_{\tau t}(H_{t-1}, z_t^p), \overline{h}_{\tau t}(H_{t-1}, z_t^p))$, where $\underline{h}_{\tau t}(H_{t-1}, z_t^p) \leq \overline{h}_{\tau t}(H_{t-1}, z_t^p)$, for which employers will offer jobs. These assumptions simplify the proof of existence by ensuring continuity at the boundaries. Moreover, they imply that competition over workers takes place within each occupation once the worker chooses hours, restricting on-equilibrium-path contracts for currently employed workers. This can also be achieved by assuming that workers commit to occupations before offers are made.

3.4 The Structure of the Game: Information, Timing and Strategies, and Beliefs

Our model is a reputation game in which workers have private information and select how much hours to work and employers offer contracts for those hours in each period. Observing workers' choices, employers update their beliefs. Below, we describe the information, timing and strategies, and beliefs.

3.4.1 Information

All utility- and production-function parameters, hiring costs, labor-market histories, workers' gender and production-relevant characteristics, z_t^p, are common knowledge. The worker's private information, z_t^*, includes consumption, a subset of the time-varying characteristics that determine the utility associated with alternative consumption and leisure allocations, and ε_t. The distribution functions of z_t and ε_t, however, are common knowledge. Let $\omega_t \equiv (H_{t-1}, z_t, \eta\lambda_t,)$ denote the worker's nonidiosyncratic state variables. The worker's complete state vector is $(\omega_t, \varepsilon_t)$. We denote the worker's type by $z_t^* \subset (\omega_t, \varepsilon_t)$. Note types are persistent over time, therefore workers have better information about the probability of remaining in the firm in the future. Let $F_{0i}^*(z_{t+1}^* \mid z_t^*, H_t)$ denote the derived transition distribution of z_t^*.

3.4.2 Timeline and Strategies

At the beginning of each period, the growth rates of the aggregate utility costs and of the aggregate permanent occupation-specific productivity shocks between periods t and $t+1$ are observed by all agents in the economy, and workers privately observe their type, z_t^*. Given this information, workers make participation, occupation, and hours decisions, a_t. Observing workers' choices, firms simultaneously offer salaries for each worker. Workers observe the offers and choose a firm.

Thus, we can summarize a worker's strategy as $\{\sigma_i(a_t \mid z_t^*, H_{t-1}),\, c_t,$ offer choice$\}$. A firm's strategy is denoted by $S_{i\tau t}(h_t, H_{t-1}, z_t^p)$. Figure 2 summarizes the exact timeline.

3.4.3 Beliefs

At the beginning of each period, firms form a (common) set of prior beliefs on each individual worker's type, $\mu_{it}(z_t^* \mid H_{t-1}, z_t^p)$. Upon observing the worker's labor-market actions, a_t, firms update their beliefs about each worker's type. We denote *posterior* beliefs—which are formed upon observation of a_t by $\widetilde{\mu_{it}}(z_t^* \mid H_{t-1}, z_t^p, a_t)$. Notice that $\widetilde{\mu_{it}}(z_t^* \mid H_{t-1}, z_t^p, a_t)$ is used to form the prior beliefs in period $t+1$, as the types are persistent over time and evolve according to the Markov process specified above.

3.4.4 Equilibrium Definition

This paper uses the Perfect Bayesian equilibrium concept.

Definition 3.1 (Equilibrium) *A Perfect Bayesian Equilibrium consists of labor-market strategies, σ_{it}, Frisch demand for consumption, c_{it}^o, the worker's offer choice, $\{S_{i\tau t}\}_{\tau=1}^{\Upsilon}$, and a common belief system, such that*

1. *Each player's strategy is optimal given beliefs and other players' strategies*

2. *The posterior beliefs, $\widetilde{\mu}$, satisfy Bayes' rule when possible.*[7]

$$\widetilde{\mu}_{it}(z_t^* \mid z_t^p, H_{t-1}, a_t) = \frac{\mu_{it}(z_t^* \mid z_t^p, H_{t-1})\sigma_{it}(a_t \mid z_t^*, H_{t-1})}{\int \mu_{it}(z_t^* \mid .)\sigma_{it}(a_t \mid .)\,\mathrm{d}z_t^*} \tag{6}$$

[7] See Definition 8.2 of Fudenberg and Tirole (1996) for the formal description of the equilibrium conditions.

and for all histories, types, and actions

$$\widetilde{\mu}_{it}(z_t^* \mid z_t^P, H_{t-1}, a_t) = \widetilde{\mu}_{it}(z_t^* \mid z_t^p, H_{t-1}, \widehat{a}_t) \text{ if } a_t = \widehat{a}_t.$$

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

$$\mu_{it+1}(z_{t+1}^*|H_t, z_t^p) = f_{i0}^*(z_{t+1}^* \mid z_t^*, H_t)\widetilde{\mu}_{it}(z_t^* \mid .) \tag{7}$$

The equilibrium conditions state that given the (common) beliefs of firms, each employer offers a salary which maximizes payoffs. Workers choose participation, hours, occupation, and consumption optimally given the firms' offer strategies. Firms observe workers' hours and participation decisions and form beliefs. These beliefs follow Bayes' rule and are consistent with the workers' strategies on equilibrium path. At the beginning of each period, firms' beliefs about worker's types are updated according to the distribution of the i.i.d. shocks and the transition densities of workers' types.

4 Equilibrium Analysis

The equilibrium contracts will now be determined as a function of the employers' information, (z_t^p, H_{t-1}), worker's labor-market actions, a_t, and the firms' beliefs. The contract depends on the perceived probability of remaining in the job, which will satisfy conditions 6 and 7 above. Workers' expected lifetime utility will be maximized by labor supply decisions, given the worker's observed state variables and employers' optimal contracts. The derived participation, hours and occupation decisions will be driven by three effects on current and future market productivity, employers' beliefs and utility. First, the returns to past non-market hours, and the returns to experience in the labor market. Second, the differences in returns to experience and the depreciation of human capital in the different occupations. Third, the "reputation" effect of labor supply on employers' beliefs about worker's types. The latter is a result of the effect of past labor supply on beliefs and the property of Bayesian updating in conditions 6 and 7. We then prove that an equilibrium with the above properties exists. Lastly we derive the main results: the equilibrium labor-market gender gaps and their evolution over time.

4.1 Employers' Equilibrium Strategy

The employers' equilibrium strategy consists of a salary schedule for hours and observable characteristics. Proposition 4.1 states the optimal salary schedule.

Proposition 4.1 (Competitive Salary Schedule) *Suppose that given any choice of hours, a worker accepts the contract with the highest salary. Then for any hours and firms' beliefs, the competitive salary is*

$$S_{i\tau t}(h_t, H_{t-1}, z_t^p) = y_{\tau t}(h_t, H_{t-1}, z_t^p) - \gamma_\tau + \beta\gamma_\tau \widetilde{p}_{i\tau t+1}(H_t, z_t^p) \tag{8}$$

for all $h_t \in (\underline{h}_{\tau t}, \overline{h}_{\tau t})$*, where* $\widetilde{p}_{i\tau t+1}(H_t, z_t^p)$ *is the employers' perceived probability that the worker will work in the firm the proceeding period.*

See proof in Appendix B.

First notice that we now express the employers' strategy as a function of probability of staying in the job and not the generic beliefs. This can be done because there exists a one-to-one mapping between the posterior beliefs and the implied participation probability.

The salary schedule is equal to the worker's productivity plus a function of the hiring costs and the probability the worker stays with the firm the next period. It is the highest outside offer a worker can receive, and is offered so an outside employer makes expected profits of zero over the worker's years in the job. Furthermore, no employer can profitably unilaterally deviate from this salary schedule (see Proposition 4.3 below). A proof and formal statement of off-equilibrium beliefs is in Appendix B.

Under symmetric information, each worker earns the expected productivity over time net of hiring costs. However, under asymmetric information, each worker earns the expected productivity (net of costs) over the expected employment spell of all workers with the same observable characteristics. Therefore, the only difference in the optimal salary under symmetric versus asymmetric information is the perceived probability of remaining in the job. Thus, consider two workers with the same publicly observable characteristics, z_t^p, who choose the same number of hours worked in equilibrium under asymmetric information. The worker whose z_t^* implies that she/he is less attached to the firm earns more than his/her salary under symmetric information, and the worker whose z_t^* implies she/he is more attached earns less.

4.2 Workers' Equilibrium Strategies

Conditional on the optimal salary schedule, we derive the worker's participation, hours, and occupation decisions. Define the ex-ante conditional valuation functions associated with the decisions to work and not to work at time t as V_{1it} and V_{0it}, respectively. The ex-ante conditional valuation function is the discounted sum of future payoffs before the choice-specific idiosyncratic utility shocks are realized and actions taken. Formally,

$$\begin{aligned} V_{kit}(\omega_t) \quad \equiv \quad & \max_{\{h_s;\{I_{\tau s}\}_{\tau=1}^{\Upsilon}\}_{s=t}^{T}} E_t \left\{ \sum_{s=t}^{T} \beta^{s-t} \left[d_s u_{0is}(z_s) + u_{1i}(l_s, z_s, , H_{s-1}) + d_s \varepsilon_{1s} \right.\right. \\ (9) \qquad & \left.\left. + \ (1-d_s)\varepsilon_{0s} + \eta\lambda_s \sum_{\tau=1}^{\Upsilon} I_{\tau s} S_{i\tau s}(h_s, z_s^p, H_s) \right] \mid d_t = k \right\}, \end{aligned}$$

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

$$(10) \qquad d_{it}^o(\omega_t, \varepsilon_{0t}, \varepsilon_{1t}) = \begin{cases} 1 & \text{if } V_{1it}(\omega_t) + \varepsilon_{1t} \geq V_{0it}(\omega_t) + \varepsilon_{0t} \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 4.1 below summarizes the optimal offer choice of the workers.

Lemma 4.1 *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.*

See proof in Appendix B.

Higher salary, given the number of hours worked, increases utility. The only reason for choosing a lower salary is if it affects beliefs and therefore future earnings. Future earnings depend on experience through the production function and beliefs. Since past salaries are not observed, future salary choice cannot be altered by choosing a lower salary today. Since by assumptions (3.1) and (3.2) only one occupation offers hours for a given set of characteristics, there is no profitable deviation for the worker.

Define by h_{it}^{o} the optimal labor-supply decision in period t and by $h_{it}^{*} \in (0,1)$ the optimal interior solution of the labor-supply decision in period t. Using the above notation, the workers' optimal strategies are summarized in Proposition 4.2.

Proposition 4.2 *Given the firms' strategies, equation (4) describes the optimal consumption behavior, a worker's best response is characterized by* p_{it}^{o}*, the conditional probability of participation,* $h_{it}^{*}(\omega_t)$*, the number of hours worked, and* $I_{i\tau t}^{0}(\omega_t)$*, the optimal occupation choice. That is,* $\sigma_{it} \equiv \{p_{it}^{o}, h_{it}^{*}, I_{i\tau t}^{0}\}$*, where*

(11)

$$p_{it}^{o}(\omega_t) = E[d_{it}^{o} \mid \omega_t] = \int_{-\infty}^{V_{1it}-V_{0it}} (\varepsilon_{0t} - \varepsilon_{1t})\, \mathrm{d}F_1(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t}) = Q(V_{1it}(\omega_t) - V_{0it}(\omega_t)),$$

$$\begin{aligned} h_{it}^{*}(\omega_t) \;=\; & \underset{\{I_{\tau t}\}_{\tau=1}^{\Upsilon}}{\arg\max} \Bigg\{ \underset{h_t \in (\underline{h}_{\tau t}, \overline{h}_{\tau t})}{\arg\max} \Big[u_{0it}(z_t) + u_{1i}(l_t, H_{t-1}, z_t) + \eta \lambda_t S_{i\tau t}(h_t, z_t^{p}, H_{t-1}) \\ & + \beta E_t\{p_{it+1} V_{1it+1}(\omega_{t+1}) \mid \omega_t, d_t = 1, I_{\tau t} = 1\} \\ & + \beta E_t\{(1 - p_{it+1}) V_{0it+1}(\omega_{t+1}) \mid \omega_t, d_t = 1, I_{\tau t} = 1\} \Big] \Bigg\}, \end{aligned} \tag{12}$$

and

$$I_{i\tau t}^{0}(\omega_t) \equiv I\{\underline{h}_{\tau t} < h_{it}^{*}(\omega_t) < \overline{h}_{\tau t}\}. \tag{13}$$

See proof in Appendix B.

A useful insight from the seminal work of Hotz and Miller (1993) applies to our model: There is a one-to-one relationship between the equilibrium choice probabilities and the difference between the ex-ante value functions, $V_{1it}(\omega_t) - V_{0it}(\omega_t)$. Let $Q : R \to (0,1)$ denote the mapping from the choice-specific value function to the conditional choice probabilities. Then we obtain equation (11) of Proposition 4.2.

Equation (12) is obtained by maximizing separately over each open interval $(\underline{h}_{\tau t}, \overline{h}_{\tau t})$ and then choosing the occupation that yields the highest expected life-time utility. Under Assumptions 3.1 and 3.2 any choice of hours, h_t, conditional on (H_{t-1}, z_t^{p}) *maps* into a unique occupation choice in period t implying equation (13).

Notice that if an interior solution to equation (12) exists the first order condition is

$$
\begin{aligned}
& \frac{\partial u_{1i}(l_t, H_{t-1}, z_t)}{\partial h_t} + \eta\lambda_t \frac{\partial y_{\tau t}(h_t, H_{t-1}, z_t^p)}{\partial h_t} + \eta\lambda_t\beta\gamma_\tau \frac{\partial \widetilde{p}_{i\tau t+1}(H_t, z_t^p)}{\partial h_t} \\
= & -\beta E_t \left\{ \frac{\partial V_{0it+1}(\omega_{t+1})}{\partial h_t} + p_{it+1} \frac{\partial [V_{1it+1}(\omega_{t+1}) - V_{0it+1}(\omega_{t+1})]}{\partial h_t} \right. \\
& \left. + [V_{1it+1}(\omega_{t+1}) - V_{0it+1}(\omega_{t+1})] \frac{\partial p_{it+1}}{\partial h_t} | h_t \in (\underline{h}_{\tau t}, \overline{h}_{\tau t}), I_{\tau t} = 1 \right\}, \qquad (14)
\end{aligned}
$$

for each occupation τ. From equation (9) the derivatives of the continuation valuation functions with respective to current hours are

$$
\begin{aligned}
\frac{\partial V_{kit+1}(\omega_{t+1})}{\partial h_t} = & E_{t+1} \left\{ \sum_{s=t+1}^{T} \beta^{s-t-1} \left[\frac{\partial u_{1i}(l_s, z_s, H_{s-1})}{\partial h_t} \right. \right. \\
& + \eta\lambda_s \sum_{\tau=1}^{\Upsilon} I_{\tau s} \frac{\partial y_{\tau s}(h_s, H_{s-1}, z_s^p)}{\partial h_t} \\
& \left. \left. + \eta\lambda_s \sum_{\tau=1}^{\Upsilon} I_{\tau s}\beta\gamma_\tau \frac{\partial \widetilde{p}_{i\tau s+1}(H_s, z_s^p)}{\partial h_t} \right] \mid \omega_t, d_{t+1} = k \} \right\}. \qquad (15)
\end{aligned}
$$

Equation (14) can be interpreted in terms of the standard dynamic marginality conditions. It highlights the trade off that a worker faces when making hours and occupation choices in our model. First note that the standard (Mincer and Polachek (1974)) on-the-job-training motive to invest in both non-market and market human capital is capture by the first and second terms on the RHS of equation (15). Thus, the expectations of future family structure and market activities of women and men are important determinants of the level and type of current labor supply. These terms are non zero because of the existence of intertemporal nonseparabilities in preferences and market production.

The second term on the LHS of equation (14) captures the varying returns to current hours across occupation. Occupations may have different returns to working part time versus full time. This, however, has to be balanced against the varying returns to experience across occupations in the future, captured by the second term on the RHS of equation (15). Thus depending on an individual level of hours this may generate different lifetime profile of occupation selection. This feature of our model is similar to that of Polachek (1981).

The third terms on the LHS of equation (14) and RHS of equation (15) capture the signalling and reputation effects of hours choice. In choosing hours and occupation, workers consider their effects on the employers current and future beliefs.

Finally, note that workers take into account the value that an extra hour of work today creates through its effect on the probability of future participation (the last term on the RHS of equation (14)), highlighting that a worker's future type is endogenous in our model.

4.3 Existence of Equilibria

In order to establish the existence of an equilibrium, we first show that given Proposition 4.1, Lemma 4.1 and Proposition 4.2, no firm can profitably unilaterally deviate from the competitive salary schedule. Using this result we establish the existence of an equilibrium.

Proposition 4.3 *Given that firms are following the competitive salary schedule in Proposition 4.1, workers are following the strategies specified in Proposition 4.2, and the beliefs system which is defined in Definition 3.1 and Assumption B.1:*

1. *No single firm can profitably unilaterally deviate from the competitive salary schedule.*

2. *There exists an equilibrium in which the perceived probability a worker will work in the firm in next period is correct conditional on employers' information.*

See proof in Appendix B.

The first part of Proposition 4.3 stands in contrast to the well-known nonexistence result in Rothschild and Stiglitz's (1976) model. However, as pointed out in Hellwig (1987), this nonexistence result is very 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 with the workers finally choosing which offer to accept. This difference in timing transforms the within period game into a signaling game and the nonexistence result of the original Rothschild–Stiglitz model is broken. Using the Walrasian Equilibrium concept instead of a Perfect Bayesian Equilibrium would give existence of equilibrium with the same optimal salary schedule. For papers that have demonstrated existence of equilibrium in Rothschild–Stiglitz type models with

Walrasian equilibrium, see Riley (1979), Prescott and Townsend (1984), Gale (1992), and Bisin and Gottardi (2006), among others.

4.4 The Gender Labor-Market Gaps

Under symmetric information, if men and women were identical in all respects, there would be no labor market outcomes differences. Suppose instead, that there are some z_t^* for which the cost of participation and the disutility of hours worked are larger for women. Then a woman with the same characteristics as a man may earn a lower salary because transitions into a state in which the costs of participation are higher may imply a lower probability of future participation. Facing states which can be reached, and in which the likelihood of participation is lower, may cause the current value of participation to be lower and the value of working long hours to be lower as well. Women would participate less, work fewer hours and sort into different occupations. Thus, the earnings gap and different labor-market histories can be generated under symmetric information.

Discrimination in this paper refers to the difference between the earnings of men and women under symmetric and asymmetric information. Under symmetric information individual workers' future probability of remaining in the job is known to employers, whereas under asymmetric information gender is used to infer this probability. Suppose that the distributions of skills, characteristics, and preferences are ex-ante identical for men and women. Then our model may give rise to *discriminatory equilibria*.[8] Suppose employers believe that women have a lower likelihood of future employment than men have. Then women would face lower earnings than men with identical characteristics. These beliefs are self-fulfilling in equilibrium and induce women and men to make different participation and labor-supply decisions.

Moreover, if women work less in equilibrium, they will sort into occupations with lower returns to labor-market experience and lower costs of hiring new workers. Occupations with lower costs of hiring new employees will have smaller differences in earnings for men and women with the same observable characteristics; the gender earnings gap in these occupations may be smaller.

[8] See Tirole (1996) for a discussion of dynamic adverse selection and statistical discrimination. The difference between this model and Tirole's arises because the matching in Tirole's model between firms and workers is random. In our framework, workers post hours and firms offer contracts of earnings for those hours.

Whereas discrimination may be a result of pure coordination failure, our model may exhibit discriminatory equilibrium due to cross-group (gender) effects (see, for example, Moro and Norman, 2004). Note that there are complementaries in the utility function, i.e., consumption depends on the household budget constraint, hence there is complementarity between women's and men's hours and participation. A discriminatory (asymmetric) equilibrium may then be established, even if there is no coordination failure.

According to our model, the following exogenous changes could account for the narrowing in the observed gender earnings gap over time: occupation-specific aggregate productivity, demographic characteristics, fixed costs of participation, education levels across cohorts and beliefs across cohorts. Suppose that there is an increase in overall productivity within an occupation. Such an increase affects the earnings of all workers because $y_{\tau t}(h_t, H_{t-1}, z_t^p)$ increases, but if men's participation rate is high, beliefs about women's participation may increase women's earnings relative to men's earnings. This increase in earning will result in a bigger increase in labor supply and participation of women. A similar effect can be driven by changes in demographics (such as a decline in fertility), and decline in the cost of participating in the labor market.

Note that the equilibrium characterization is for each cohort separately. In the data, we observe several overlapping cohorts. There are differences is educational attainment across cohorts which may affect both productivity and beliefs. A worker's cohort is an observable characteristic, therefore, it is possible that employers' initial beliefs will be different for workers who are identical in all observable characteristics, except for the cohort they belong to. The theory imposes no restrictions on how initial beliefs are formed. If there were social and cultural changes over cohorts, the beliefs about future participation would capture that.[9]

5 Multiplicity of Equilibria and Identification

One of the main problems in estimating game theoretic models is the possibility of multiple equilibria. Before discussing different solutions to this problem, it is important to clarify that equilibrium uniqueness is neither a *necessary* nor *sufficient*

[9] See Fernandez (2007) and Fogli and Veldkamp (2007) for models of social learning and female labor force participation.

condition for the identification of a model.

Formally, define the mapping $\mathcal{F}$ from the space of structural characteristics $\Theta(\mathcal{M})$ to the space of conditional distributions Π such that $\mathcal{F}(\theta(\mathcal{M}))$ contains all the conditional distributions predicted by the model when the structural characteristic is $\theta(\mathcal{M})$. Multiple equilibria means that $\mathcal{F}(\theta(\mathcal{M}))$ is a correspondence. Identification means that the inverse mapping $\mathcal{F}^{-1}(.)$, evaluated at the population distribution $F^o_{Y|X}$, is a function. Obviously a mapping can be a correspondence and its inverse a function.[10]

The solution to the multiple-equilibrium problem is based on the following intuition. The players' equilibrium strategies can be recovered from the data under the assumptions that the data was generated by a single equilibrium (within observable groups of agents), and that the econometrician observes all the nonidiosyncratic state variables. 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 in all equilibria, hence, an estimator of the structural parameters can be obtained from a criterion function based on these best responses. This criterion function will be well defined once the model is identified.

The approach we take in this paper to show identification of the single-agent problem is constructive in nature. We rely on Chesher (2007), which shows that if 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 $\theta(\mathcal{M}) = \theta^*$, then the value of a structural characteristic $\theta(\mathcal{M})$ is identified by the model when $\theta(\mathcal{M}) = \theta^*$. The advantage of this approach to identification is that there is a natural analog estimator of the value of the structural characteristic. Given panel data, $\{a_{nt}, H_{nt-1}, z_{nt}, c_{nt}, S_{n\tau t}\}_{n=1,\tau=1,t=1}^{N_i,\Upsilon,T}$, where n indexes individual, t indexes the year, we outline our identification strategy below.

1. We show that under standard regularity conditions on $u_{2i}(z_t, c_t, \varepsilon_{2t})$, if it is multiplicatively separate in ε_{2t}, then the standard independence assumption between z_t and ε_{2t} allows us to identify $\eta\lambda_t$ up to a proportionality constant.

2. Assuming that z_t^p is fully observed by the econometrician, we show that, under standard regularity conditions, the equilibrium salary schedule and β are identified.

[10] For more a detailed discussion of this problem, see Jovanovic (1989) and Aguirregabiria (2005b).

3. Let $F_1(\varepsilon_{0t}, \varepsilon_{1t})$ be the marginal of $F_1(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{1t})$. Given that $S_{i\tau t}(h_t, H_{t-1}, z_t^p)$, β, and $\eta\lambda_t$ are identified, we show that $u_{0it}(z_t)$ and $u_{1i}(l_t, z_t, H_{t-1})$ are identified up to $F_1(\varepsilon_{0t}, \varepsilon_{1t})$ and two additive constants.

Putting 1, 2, and 3 together, we conclude that our model is identified up to $F_1(\varepsilon_{0t}, \varepsilon_{1t})$ and two additive constants.

5.1 Identification of the Marginal Utility of Wealth

It is well known in the literature on the estimation of consumption functions that the general form of utility with risk aversion is not identified without quantity and price data, which we do not have. Therefore, we follow the literature and state sufficient conditions to obtain identification of the marginal utility of wealth. Below, we state a set of such conditions.

Assumption 5.1 *The marginal utility of consumption has the following form.*

$$\frac{\partial u_{2i}(c_{nt}, z_{nt}, \varepsilon_{2nt})}{\partial c_t} = u_{2c}(c_{nt})u_{2z}(z_{nt})\varepsilon_{2nt}, \qquad (16)$$

where $u_{2c}(c_t) > 0$*,* $u_{2z}(z_t) > 0$*, and* $\varepsilon_{2t} > 0$.

Assumption 5.2 *1)* $E[\log \varepsilon_{2nt} | z_{nt}] = 0$ *for all* n *and* t*. 2)* $E_n[\log(\eta_n) | z_{nt}] = 0$.

Assumption 5.3 *1)* z_{nt} *has a continuous element* z_{cnt} *with continuous variation on its support* $[\underline{z_c}, \overline{z_c}]$*. 2)* $u_{2z}(\underline{z}_c, .) = 0$.

Assumption (5.1) states that the marginal utility of consumption is multiplicatively separable. For example, both the class of constant absolute risk aversion and the class of constant relative risk aversion satisfy this assumption. Assumption 5.2(1) formally states that the error is mean independent of z_{nt} with expectation zero. Assumption 5.2(2) is the standard normalization needed in a panel data model in order to recover the level of the time component. Finally, Assumption 5.3(1) states that at least one variable with continuous variation on its support is required, and Assumption 5.3(2) is a boundary condition. Assumption 5.3 can be replaced with a parametric assumption on the function $u_{2z}(z_{nt})$.

Lemma 5.1 *If* $u_{2c}(c_{nt})$ *is known, and assumptions 5.1–5.3 are satisfied. Then* $\eta_n\lambda_t$ *is identified.*

See proof in Appendix C.

5.2 Identification of the Equilibrium Salary Schedule

Next, we establish the identification of the equilibrium salary schedule.

Assumption 5.4 *1. Suppose there exist observable characteristics, H_{t-1}, z_t^p, on a set of positive measure, such that*

$$\triangle\, \widetilde{p}_{\tau t+1}(H_t, z_t^p) = \widetilde{p}_{m\tau t+1}(H_t, z_t^p) - \widetilde{p}_{w\tau t+1}(H_t, z_t^p) \neq 0 \quad \vee\, \tau$$

2. $y_{t\tau}(0, H_{t-1}, z_t^p) = 0 \;\vee \tau, H_{t-1}, z_t^p$

Assumption 5.4(1) states that in each occupation there is a range of hours, labor market experience and individual characteristic for which the employers hold different belief about the two genders, Assumption 5.4(2) states that an input of zero hours produces zero output.

Lemma 5.2 *Under Assumptions (3.1), (3.2), and (5.4), $y_{\tau t}(h_t, H_{t-1}, z_t^p)$, β, and γ_τ are identified, and there are at least two overidentifying restrictions.*

See proof in Appendix C.

The two overidentifying restrictions allow to test the restriction that there is no difference in productivity between men and women. The main identifying assumption in Lemma (5.2) is that there is a difference in future participation probability between men and women. An alternative assumption is the there is a difference in future participation probability of different cohorts for a given gender.

5.3 Identification of the Utility of Non-market Time

In the literature on the estimation of dynamic Markovian games, it is standard to use time-series data to estimate and identify models.[11] We extend this approach to a panel data setting, considering age–education cohort partition of data generated by a single path of play, exploiting, therefore, the information contained in the repeated observation of the same players in the cohort partition along the path of play. Because different cohorts may be playing different equilibria, we also have variation across cohort partitions. Below, we formalize the equilibrium selection discussed in the previous paragraph.

[11] See Pesendorfer and Schmidt-Dengler (2006) for an example of this approach.

Assumption 5.5 (Equilibrium Selection) *Conditional on the time invariant component of z_t^p, the data for each age–education cohort is generated by only one equilibrium.*

This assumption rules out the possibility that for any given age–education cohort and time invariant component of z_t^p, the time series data is generated by a mixture of two or more equilibria.

Assumption 5.6 *The econometrician observes all the private information of the worker except for the idiosyncratic components ε_t.*

This assumption formalizes the notion that retrospective survey data often allows the econometrician to obtain information that is not publicly observed at the time of play. It is similar to the assumption made by Altonji and Pierret (2001), which tests for employers' learning in a model of statistical discrimination.

Let us redefine the primitives of our problem as follows. The per-period utility is

$$U_{kit}(\omega_{nt}) = \begin{cases} u_{1i}(1, z_{nt}, H_{nt-1}) & \text{for } k = 0 \\ u_{0it}(z_{nt}) + u_{i1}(l_{nt}, z_{nt}, H_{nt-1}) & \\ \qquad + \eta_n \lambda_t \sum_{\tau=1}^{\Upsilon} I_{n\tau t} S_{i\tau t}(h_{nt}, z_{nt}^p, H_{nt-1}) & \text{for } k = 1. \end{cases} \tag{17}$$

We can write the ex-ante value functions more concisely as

$$V_{kit}(\omega_{nt}) \equiv \max_{\{h_s; \{I_{\tau s}\}_{\tau=1}^{\Upsilon}\}_{s=t}^{T}} E_t \left[\sum_{s=t}^{T} \beta^{s-t} \{ d_s [U_{1is}(\omega_{ns}) + \varepsilon_{1ns}] + (1 - d_s)[U_{0is}(\omega_{ns}) + \varepsilon_{0ns}] \} \mid d_t = k \right] \tag{18}$$

Our game differs in two important dimensions from the games typically estimated. First, in our model, employers learn and update beliefs based on the complete history of workers' behavior. Second, the state variables do not have discrete support because labor-market experience has continuous components.[12] In order to generalize results from that literature so we can apply them to our model, we borrow the concept of finite state dependence from Altug and Miller (1998).

[12] This point is less critical. See Bajari and Hong (2006) for an extension of the standard results to continuous state variables.

Definition 5.1 *Given any value of the initial state variable ω_0, there exists a finite integer $\rho(\omega_0)$, a value of the state variable ω_ρ, and a sequence of choices, over the next $\rho(\omega_0)$ periods, denoted by $d_{kt}^{\rho(\omega_0)}(\omega_t) = (d_{kt+1}^{\rho(\omega_0)}(\omega_{t+1}), \ldots, d_{kt+\rho(\omega_0)}^{\rho(\omega_0)}(\omega_{t+\rho(\omega_0)}))$, such that the state will be $\omega_{\rho(\omega_0)}$ at date $\rho(\omega_0)$ for $d_t = k$ for all k.*[13]

The finite-dependence assumption is a generalization of the Markovian assumption to a larger class of models. It breaks the connection between state variables which evolve stochastically, and individual choice prior to the ρ periods.

Assumption 5.7 (Finite State Dependence) *The nonidiosyncratic state variable ω_t has the property of finite state dependence.*

Next, we characterize the necessary condition for equilibrium, namely condition (1) of Definition 3.1. Equation(11) of Proposition 4.2 and Lemma 1 of Hotz and Miller (1993) imply that[14]:

$$V_{1it}(\omega_{nt}) - V_{0it}(\omega_{nt}) = Q^{-1}(p_{it}(\omega_{nt})). \tag{19}$$

Proposition 1 of Hotz and Miller (1993) also states that there exists a mapping $\varphi_k : [0,1] \to R$, that measures the expected value of the unobservable in the current utility, conditional on action $k \in \{0,1\}$. That is,

$$\varphi_k(p_{it}(\omega_{nt})) \equiv E[\varepsilon_{knt} \mid \omega_{nt}, d_{nt}^o = k]. \tag{20}$$

To set some notation, let $\omega_{kt}^{(s)}$ denote the state in period $t+s$ if, at time t, the k^{th} option is taken—that is, $d_t = k$—and the sequence of decisions for the next s periods is $d_{kt+1}^{\rho(\omega_0)}(\omega_{t+1}), \ldots, d_{kt+s}^{\rho(\omega_0)}(\omega_{t+s})$. Denote by $p_{kit}^{(s)}$, the probability that $d_{t+s} = 1$ conditional on $\omega_{kt}^{(s)}$, i.e. $p_{kit}^{(s)} = E\left[d_{t+s} | \omega_{kt}^{(s)}\right]$.

Combining equation (11) of Proposition 4.2, equation (19), and equation (20) with the ex-ante value function (18) allows us to write the ex-ante equilibrium value function for any initial state ω_0:

[13] 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.

[14] This equation is central to estimation in a number of papers, including Hotz et al. (1994); Altug and Miller (1998); Aguirregabiria and Mira (2002); Gayle and Miller (2004); Pesendorfer and Schmidt-Dengler (2003); Bajari, Benkard, and Levin (forthcoming); Pakes, Ostrovsky, and Berry (forthcoming); and Bajari and Hong (2005).

$$\begin{aligned}
V_{kit}(\omega_0) &= U_{kit}(\omega_0) + E_t \left\{ \sum_{s=1}^{\rho(\omega_0)} \beta^s \left[U_{0it+s}\left(\omega_{kt}^{(s)}\right) + \varphi_0\left(p_{kit}^{(s)}\right) \right.\right. \\
&\quad + p_{kit}^{(s)} \left\{ Q^{-1}\left(p_{kit}^{(s)}\right) + \varphi_1\left(p_{kit}^{(s)}\right) - \varphi_0\left(p_{kit}^{(s)}\right) \right\} \Big] \\
&\quad + \beta^{\rho(\omega_0)+1} \Big[V_{0it+\rho(\omega_0)+1}(\omega_{\rho(\omega_0)+1}) + \varphi_0\left(p_{kit}^{(\rho(\omega_0)+1)}\right) \\
&\quad \left.\left. + p_{kit}^{(\rho(\omega_0)+1)} \left\{ Q^{-1}\left(p_{kit}^{(\rho(\omega_0)+1)}\right) + \varphi_1\left(p_{kit}^{(\rho(\omega_0)+1)}\right) - \varphi_0\left(p_{kit}^{(\rho(\omega_0)+1)}\right) \right\} \right] \right\} \qquad (21)
\end{aligned}$$

This value function representation can be thought of as the continuous state/finite dependent analogue of equation (6) of Pesendorfer and Schmidt-Dengler (2006), a proof of this representation can be found in Altug and Miller (1998).

Next, we characterize, using 19 and 21, the necessary conditions for equilibrium in Proposition 4.2. First, we characterize the equilibrium relationship from (11). Substituting (21) into (19) gives equilibrium:

$$\begin{aligned}
Q^{-1}(p_{it}(\omega_{nt})) &= U_{1it}(\omega_{nt}) - U_{0it}(\omega_{nt}) + E_t \left\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[U_{0it+s}\left(\omega_{0t}^{(s)}\right) - U_{0it+s}\left(\omega_{1t}^{(s)}\right) \right.\right. \\
&\quad \left. + \varphi_0\left(p_{0it}^{(s)}\right) - \varphi_0\left(p_{1it}^{(s)}\right) \right] + p_{0it}^{(s)} \left[Q^{-1}\left(p_{0it}^{(s)}\right) + \varphi_1\left(p_{0it}^{(s)}\right) - \varphi_0\left(p_{0it}^{(s)}\right) \right] \\
&\quad \left. - p_{1it}^{(s)} \left[Q^{-1}\left(p_{1it}^{(s)}\right) + \varphi_1\left(p_{1it}^{(s)}\right) - \varphi_0\left(p_{1it}^{(s)}\right) \right] \right\} \qquad (22)
\end{aligned}$$

Note that all the elements from period $\rho(\omega_{nt})+1$ onward are the same irrespective of whether action 1 or 0 is taken today by Assumption 5.7. Hence, they fall out of the above equation. Similarly, using equation(21), the necessary condition for equilibrium hours (combining equations (14) and (15)) can be rewritten as

$$\begin{aligned}
-\frac{\partial U_{1it}(\omega_{nt})}{\partial h_t} &= E_t \left\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{\partial U_{0it+s}\left(\omega_{1t}^{(s)}\right)}{\partial h_t} + \frac{\partial \varphi_0\left(p_{1it}^{(s)}\right)}{\partial h_t} \right] \right. \\
&\quad + \frac{\partial p_{1it}^{(s)}}{\partial h_t} \left[Q^{-1}\left(p_{1it}^{(s)}\right) + \varphi_1\left(p_{1it}^{(s)}\right) - \varphi_0\left(p_{1it}^{(s)}\right) \right] \\
&\quad \left. + p_{1it}^{(s)} \frac{\partial \left[Q^{-1}\left(p_{1it}^{(s)}\right) + \varphi_1\left(p_{1it}^{(s)}\right) - \varphi_0\left(p_{1it}^{(s)}\right) \right]}{\partial h_t} \right\} \qquad (23)
\end{aligned}$$

Note that all the elements from period $\rho(\omega_{nt})+1$ onward are independent of current participation choice and choice of hours by Assumption (5.7). Hence, they fall out of the above equations.

By Proposition (4.2), equations (22) and (23) are necessary conditions and therefore must hold in all equilibria. $\left\{p_{it}(\omega_{nt}), \left\{p_{0it}^{(s)}, p_{1it}^{(s)}\right\}_{s=1}^{\rho(\omega_{nt})}\right\}_{i=w,m}$, and the distribution over which E_t is taken, are the only elements which differ across the different equilibria. These elements are conditional expectation functions and, given Assumption (5.5), can be recovered from the data; therefore, they are identified.[15] For the purpose of the identification of the utility of non-market hours, we can treat $\left\{p_{it}(\omega_{nt}), \left\{p_{0it}^{(s)}, p_{1it}^{(s)}\right\}_{s=1}^{\rho(\omega_{nt})}\right\}_{i=w,m}$ as known. Denote by $\eta_n^o\lambda_t^o$ and $S_{i\tau t}^o(h_{nt}, z_{nt}^p, H_{nt-1})$ the shadow prices and salary schedule under the true equilibrium in the data respectively. The true probabilities under the true equilibrium in the data are denoted by $\left\{p_{it}^o(\omega_{nt}), \left\{p_{0it}^{o(s)}, p_{1it}^{o(s)}\right\}_{s=1}^{\rho(\omega_{nt})}\right\}_{i=w,m}$, and define

$$\begin{aligned} Y_{i1nt} \equiv\ & \eta_n^o\lambda_t^o \sum_{\tau=1}^{\Upsilon} I_{n\tau t} S_{i\tau t}^o(h_{nt}, z_{nt}^p, H_{nt-1}) + \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left\{ \varphi_0\left(p_{0it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right) \right. \\ & + p_{0it}^{o(s)}\left[Q^{-1}\left(p_{0it}^{o(s)}\right) + \varphi_1\left(p_{0it}^{o(s)}\right) - \varphi_0\left(p_{0it}^{o(s)}\right)\right] \\ & \left. - p_{1it}^{o(s)}\left\{Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right\}\right\} - Q^{-1}(p_{it}^0(\omega_{nt})) \end{aligned} \tag{24}$$

and

$$\begin{aligned} Y_{i2nt} \equiv\ & \eta_n^o\lambda_t^o \sum_{\tau=1}^{\Upsilon} I_{in\tau t} \frac{\partial S_{i\tau t}^o(h_{nt}, z_{nt}^p, H_{nt-1})}{\partial h_t} - \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left\{ \frac{\partial \varphi_0\left(p_{1it}^{o(s)}\right)}{\partial h_t} \right. \\ & + \frac{\partial p_{1it}^{o(s)}}{\partial h_t}\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right] \\ & \left. + p_{1it}^{o(s)} \frac{\partial\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right]}{\partial h_t} \right\} \end{aligned} \tag{25}$$

Because $\eta_n^o\lambda_t^o$, $S_{i\tau t}^o(h_{nt}, z_{nt}^p, H_{nt-1})$, $F(\varepsilon_{0t}, \varepsilon_{1t})$, and β are treated as known, Y_{i1nt} and Y_{i2nt} can be treated as observed data.

[15] See Haavelmo (1944) for detail.

Lemma 5.3 *In all the equilibria, the following system of equations holds.*

$$
\begin{aligned}
Y_{i1nt} &= u_{1i}(1, z_{nt}, H_{nt-1}) - u_{0it}(z_{nt}) - u_{1i}(l_{nt}, z_{nt}, H_{nt-1}) \\
&\quad + \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[u_{1i}\left(1, z_{1nt}^{(s)}, H_{1nt-1}^{(s)}\right) - u_{1i}\left(1, z_{0nt}^{(s)}, H_{0nt-1}^{(s)}\right) \right] + \xi_{i1nt} \qquad (26)
\end{aligned}
$$

$$
(27) \quad Y_{i2nt} = -\frac{\partial u_{1i}(l_{nt}, z_{nt}, H_{nt-1})}{\partial h_t} - \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{u_{1i}\left(1, z_{1nt}^{(s)}, H_{1nt-1}^{(s)}\right)}{\partial h_t} \right] + \xi_{i2nt}
$$

where $z_{knt}^{(s)}$ *and* $H_{knt-1}^{(s)}$ *denote the worker's type and labor experience in period* $t+s$*, if, at time* t*,* $d_t = k$*, and the sequence of decisions for the next* s *periods is* $d_{kt+1}^{\rho(\omega_0)}(\omega_{t+1})$*, ...,* $d_{kt+s}^{\rho(\omega_0)}(\omega_{t+s})$*:* $E_t^o[\xi_{i1nt}|\omega_{nt}] = 0$ *and* $E_t^o[\xi_{i2nt}|\omega_{nt}, d_{nt} = 1] = 0$ *for* $i = \{w, m\}$*, and* E_t^o *is taken over the actual equilibrium played. The formal definition of the residual* ξ_{i1nt} *and* ξ_{i2nt} *are in the proof in Appendix C.*

See proof in Appendix C.

Note that the data is informative about E_t^o under Assumption (5.5). The identificability of the structural functions depends on whether these functions can be deduced from knowledge of $E_t^o[Y_{i1nt}|\omega_{nt}]$ and $E_t^o[Y_{i2nt}|\omega_{nt}, d_{nt} = 1]$ The following lemma establishes the main identification result.

Lemma 5.4 (Nonmarket Utility Identification) *Under Assumptions (3.1)–(5.7),* $u_{0it}(z_{nt})$ *is identified up to an additive constant, and* $u_{1i}(l_{nt}, z_{nt}, H_{nt-1})$ *is identified up to an additive function of* z_{nt} *and* H_{nt-1}*.*

See proof in Appendix C.

The above Lemma implies that the levels of non-market-hours utility are not identified, but the marginal utility of non-market hours are identified.[16] This result is stronger than identification results in discrete choice models with discrete state variables in which only the difference between the utilities, conditional on the choices, is identified (see Magnac and Thesmar, 2002, and Pesendorfer and Schmidt-Dengler, 2006). In this way, our identification result is similar to that found in Jofre-Bonet and Pesendorfer (2003). The continuous choice of hours in our model increases the model's identification power: if the level of utility from not working is normalized, then the rest is fully identified.

[16] Because of the ordinality of the utility function, this result is expected.

6 Empirical Strategy

The estimation of the model follows the outline of the identification strategy. First, the marginal utility of wealth is estimated from the consumption Euler equation. Second, the earnings equation is estimated from the zero-profit condition. Third, the conditional choice probabilities and the firms' equilibrium beliefs are estimated nonparametrically. Last, we form the empirical analogue of equations (26)–(27). Using a method-of-moment estimation procedure, we recover the parameters of the utility from the non-market work and the risk-aversion parameter.

The production function and the utility from non-market hours are nonparametrically identified, conditional on the distribution of ε_{nt} and the consumption-function class. In the estimation stage, however, we specify parametric functional forms for two reasons. First, in order to conduct the decomposition of the change in the earnings gap-the focus of this paper-we need to solve for counterfactual versions of the model, and that can only be done with parametric functions. Second, given the nature of the three-step estimation, a nonparametric final step would have undesirable sampling properties.

In order to implement the model, we need to further address the following issues. First, we need to specify which variables the worker observes privately and which variables are observed by employers. Second, we need to decide which restrictions to place on the employment history that employers observe. Third, we will incorporate unobserved productivity in the production function.

To address the first issue, we assume consumption is private information, and hence the marginal utility from wealth is private information. All spouse related variables (such as education), marital status, and the number and age distribution of children are assumed to be private information. We use the following criteria to determine which variables is private information. Notice that it is crucial that potential employers do not observe this information rather then current employers. First, is the variable privately observed in nature. Second, if workers do not have incentives to report truthfully, is this information easily verifiable.

We assumed that employers observe the actual hours worked in each occupation for the most recent three years, and the total number of years worked in each occupation. This assumption is made in order to reduce the dimension of the conditioning variables in the nonparametric estimation of beliefs; this restriction *is not* imposed, however, by the finite-state dependence assumption. Finally, given the functional-form restrictions on the production function, we can account for individual-specific productivity differences. These differences are assumed to be observed by the firms and workers, but not by the econometrician. The following sections provide a brief overview of each stage of the estimation and a description of the functional form restrictions imposed at each stage. A detailed description of the estimation procedure can be found in Appendix D.

6.1 Consumption Equation

We assume the following functional form for the utility from consumption.

$$u_{2i}(c_{nt}, z_{nt}, \varepsilon_{2nt}) = \exp(z'_{nt}B_4 + \varepsilon_{2nt})c_{nt}^{\alpha}/\alpha. \tag{28}$$

The Euler equation of consumption becomes

$$\exp(z'_{nt}B_4 + \varepsilon_{2nt})c_{nt}^{\alpha-1} = \eta_n\lambda_t. \tag{29}$$

Under Assumption (5.2), it is possible to estimate $(1-\alpha)^{-1}B_4$, $(1-\alpha)^{-1}\eta_n$, and $(1-\alpha)^{-1}\lambda_t$ using panel data on individual consumption and characteristics z_{nt}. See the discussion in Heckman and MaCurdy (1980), MaCurdy (1981), and Altug and Miller (1990, 1998), among others. There are several issues to address when implementing this estimation strategy.

First, the only panel data set that includes reliable labor-market variables and consumption related variable in the United States is the PSID; the PSID, however, includes data on food consumption only.[17] Therefore, most papers estimating labor supply and consumption behavior jointly use food consumption as the basis to estimate consumption equations (see Blundell and MaCurdy, 2007, for a survey of this literature). This can be justified under the assumption that food consumption is additively separable from nonfood consumption with a quasi-linear utility function. Second, food consumption is only observed at the household level, whereas, in the model, consumption is chosen at the individual level. To address this issue, we divide consumption among the household members by allowing the data to implicitly adjust

[17]See Blundell, Pistaferri, and Preston (2004) for discussion of this problem and alternative solutions.

the weight given to each household member, accounting for the number of individuals in the household and the age distribution of members of the household. Third, following standard practice, we include regional dummies in z_{nt} to account for differences in consumption prices across regions.

Finally, $(1-\alpha)^{-1}\eta_n$ is estimated for each individual in our sample. Therefore, the traditional fixed-effect estimates (as used in Heckman and MaCurdy, 1980) would be biased for small T. As shown in MaCurdy (1981), $(1-\alpha)^{-1}\eta_n$ can be written as a function of individual-specific variables. We follow the Altug and Miller (1998) implementation of this idea and construct a nonparametric estimator that is consistent as the number of individuals goes to infinity; we use years of completed education, gender, marital status by age 30, children's age distribution at age 35, home ownership at 35 and geographical location at age 30 in this estimation.

6.2 Earning Equation

We assume that the production function has the following functional form.

$$(30) \qquad y_{\tau t}(h_{nt}, H_{nt-1}, z_{nt}^p) = b_{0\tau t} + b_{\tau 1}h_{nt} + b_{\tau 2}h_{nt}^2 + \sum_{r=1}^{\rho} b_{\tau 3r}h_{nt-r} + z_{nt}^{p'}B_{\tau 5} + \nu_n$$

The production function is nonparametrically identified, as discussed in the identification section. We, however, choose this functional form for the estimation; this form captures five main features of our model: First, occupation-specific aggregate changes in productivity are captured by $b_{0\tau t}$. Second, we include the quadratic term in current hours worked, h_{nt}, to capture possibility nonlinearity (i.e. differences in returns to part-time versus full-time). Third, we include finite lags of hours worked in the most recent ρ periods to capture return and depreciation rate of past labor-market experience, an important mechanism accounting for endogenous difference in productivity between men and women. This specification was chosen for its strong support in the empirical literature (see Eckstein and Wolpin, 1989, and Altug and Miller, 1998, among others). Note that we are assuming that human capital is general in nature, but that its rate of return and depreciation are different across occupations. Fourth, we included a quadratic term of age and an interaction term of age and education as elements of z_{nt}^p. These terms capture the potential effects of education and experience on productivity. Finally, we allow for a general individual-specific additive effect, ν_n. This component typically captures the individual's unobserved skill or ability. In the context of our paper, however, it may also capture other sources of discrimination that we do not model explicitly. Thus, it gives us a natural way of ascertaining the importance of other sources of discrimination that have a time-invariant effect on earnings differences between men and women.

Given this specification of the production function and the assumptions that employers observed the occupation worked in, hours worked for the past three years, and the total number of years worked in each occupation, we estimate the parameters of the earnings equation based on the zero-profit condition. A detailed description is in the estimation appendix (D).

6.3 Conditional Choice Probabilities and Beliefs

There are five inputs of equations (24) and (25) to be estimated before we can form the empirical counterparts of Y_{i1nt} and Y_{i2nt}. First, Y_{i1nt} is a function of the equilibrium salary schedule, which is a function of the employers' beliefs (equation (4.1)). These beliefs will be estimated nonparametrically. Second, Y_{i2nt} is a function of the derivative of the equilibrium salary schedule with respect to current hours; we estimate this derivative nonparametrically. Third, Y_{i1nt} is a function of the current conditional choice probabilities, $p^0_{int}(\omega_{nt})$, which we will also estimate nonparametrically. Finally, Y_{i1nt} and Y_{i2nt} are functions of $p^{o(s)}_{kint}$ and their derivatives, respectively, which will also be estimated nonparametrically. The following subsections discuss the estimation of all these elements.

6.3.1 Estimation of the Equilibrium Beliefs and their Derivatives

The equilibrium beliefs for each occupation, $\widetilde{p}_{in\tau t+1}$, are computed as a nonlinear regression of the product of next-period participation and occupation choice index, $d_{nt+1} \times I_{n\tau t+1}$ on today's public information variables, z^p_{nt}, work histories, H_{nt-1}, and hours worked, h_{nt}, conditional on working today in occupation τ. Let $X_{nt} = (z^p_{nt}, H_{nt-1}, h_{nt}, \nu_n, Gender_n)$ and $NY_{n\tau t-1}$ be the total number of years worked in occupation τ up to period $t-1$. Only two occupations are used in the estimation, so $\tau \in \{1, 2\}$. The labor-market history used in this paper is defined as

$$(31) \quad H_{nt-1} = (NY_{n1t-1}, NY_{n2t-1}, d_{nt-3}I_{n1t-3}, d_{nt-3}I_{n2t-3}, \\ \ldots, d_{nt-1}I_{n1t-1}, d_{nt-1}I_{n2t-1}, h_{nt-3}, \ldots, h_{nt-1})$$

Let $J_1[\delta_{1N}^{-1}(X_{nt} - X_{n's})]$ denote a kernel where δ_N is the bandwidth associated with each argument. The nonparametric estimate of $\widetilde{p}_{in\tau t+1}$, denoted $\widetilde{p}_{in\tau t+1}^N$, is computed using the kernel estimator:

$$(32) \qquad \widetilde{p}_{in\tau t+1}^N = \frac{\sum_{n'=1,n'\neq n}^{N} \sum_{s=1}^{T-1} d_{n's+1} I_{n'\tau s+1} d_{ns} I_{n'\tau s} J_1[\delta_{1N}^{-1}(X_{nt} - X_{n's})]}{\sum_{n'=1,n'\neq n}^{N} \sum_{s=1}^{T-1} d_{n's} I_{n'\tau s} J_1[\delta_{1N}^{-1}(X_{nt} - X_{n's})]}.$$

The derivative is then estimated using the standard nonparametric derivative kernel estimator (see Pagan and Ullah, 1999).

6.3.2 Estimation of the Conditional Choice Probabilities

The estimate of the conditional choice probabilities requires us to be more specific about the state variables. In contrast to the beliefs, the conditional choice probabilities are defined from the workers' perspective and not the firms perspective. From the estimation of the consumption equation, $\eta_n \lambda_t$ is known up to a proportionality constant. The elements included in z_{nt} are *number of individuals in the family unit*, *number of children younger than three*, *number of children between three and fourteen*, *age*, *years of completed education*, *marital status*, *spouse's years of education (if married)*, and *gender*.

The conditional choice probabilities, p_{int}, are computed as nonlinear regressions of the participation index, d_{nt}, on the current state, $\omega_{nt}^N \equiv (z'_{nt,} H_{nt-1}, \eta_n^N \lambda_t^N)'$, where the N superscript denotes an estimated quantity. We denote by $J\left[\delta_N\left(\omega_{nt}^N - \omega_{n's}^N\right)\right]$ the kernel and by δ_N the bandwidth associated with each argument. The nonparametric estimate of p_{int}, denoted by p_{int}^N, is computed using the kernel estimator:

$$(33) \qquad p_{int}^N = \frac{\sum_{n'=1,n'\neq n}^{N} \sum_{s=1}^{T} d_{n's} J\left[\delta_N^{-1}\left(\omega_{nt}^N - \omega_{n's}^N\right)\right]}{\sum_{n'=1,n'\neq n}^{N} \sum_{s=1}^{T} J\left[\delta_N^{-1}\left(\omega_{nt}^N - \omega_{n's}^N\right)\right]}.$$

6.3.3 Estimation of the Finite-State Path Probabilities and their Derivatives

We begin by characterizing the different possible sequences of choices that can lead to the same labor-market history at a certain point in time due to the assumption of finite state dependence.

The hypothetical labor-market history is defined as

$$(34)\quad H_{1nt}^{(s)} = (NY_{n1t-1+s}, NY_{n2t-1+s}, d_{nt-3+s}I_{n1t-3+s}, d_{nt-3+s}I_{n2t-3+s}, \ldots, d_{nt-1}I_{n1t-1}, \\ d_{nt-1}I_{n2t-1}, I_{n1t}, I_{n2t,}, 0, \ldots, 0, h_{nt-3+s}, \ldots, h_{nt-1}, h_{nt}^{*}, 0, \ldots, 0)$$

and

$$(35)\quad H_{0nt}^{(s)} = (NY_{n1t-1+s}, NY_{n2t-1+s}, d_{nt-3+s}I_{n1t-3+s}, d_{nt-3+s}I_{n2t-3+s}, \ldots, d_{nt-1}I_{n1t-1}, \\ d_{nt-1}I_{n2t-1}, 0, 0, I_{n1t}, I_{n2t,}, 0, \ldots, 0, h_{nt-3+s}, \ldots, h_{nt-1}, 0, h_{nt}^{*}, 0, \ldots, 0)$$

where

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

for $s = 1, 2, 3$. The vectors $H_{1nt}^{(s)}$ would be the work history of an individual entering period $t + s$ who had accumulated the work history H_{nt-1}, then chose to work in period t the optimal hours in the optimal occupation. However, for the following $s - 1$ periods between t and $t + s$, this individual chose not to work. Conversely, the vectors $H_{0nt}^{(s)}$ would be the work history of an individual entering period $t + s$ who had accumulated the work history H_{nt-1}, then chose *not* to work in period t. Then in period $t + 1$ he/she chooses to work the optimal hours and occupation as if he/she had worked in period t. However, for the following $s - 2$ periods between $t + 1$ and $t + s$, this individual chose not to work. Notice that these two sequences of decisions will lead to the same labor-market history in period $t + 4$:

$$H_{0nt}^{(4)} = H_{1nt}^{(4)} = (NY_{n1t-1+s}, NY_{n2t-1+s}, 0, 0, \ldots, 0)$$

Thus, assuming a Markovian transition of the variables z_{nt}, we have found the decision sequence that satisfies the assumption of finite state dependence. Let us define the following participation indices that correspond to the decision sequence that gets one to $H_{knt}^{(s)}$.

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

and

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

Note that $d_{1nt}^{(s)}$ and $d_{0nt}^{(s)}$ are equal to one if the individual entering period t has followed a decision path identical to the path hypothesized under $H_{1nt}^{(s)}$ and $H_{0nt}^{(s)}$. Let $\omega_{knt}^{(s)N} \equiv (z'_{nt+s}, H_{knt}^{(s)}, \eta_n^N \lambda_{t+s}^N)'$ for $k = \{0, 1\}$ be the empirical counterpart of the hypothetical state. Recall that $p_{kit}^{(s)} = E[d_{t+s}|\omega_{kt}^{(s)}]$, hence it can be estimated as nonlinear regressions of the participation index, d_{nt}, on the hypothetical state, $\omega_{knt}^{(s)N} \equiv (z'_{nt+s,} H_{knt}^{(s)}, \eta_n^N \lambda_{t+s}^N)$, conditional on $d_{knt}^{(s)} = 1$. Specifically,

$$p_{iknt}^{(s,N)} = \frac{\sum_{n'=1,n'\neq n}^{N} \sum_{r=1}^{T} d_{n'r} d_{kn'rs}^{(s)} J\left[\delta_N^{-1}\left(\omega_{knt}^{(s)N} - \omega_{n'r}^{N}\right)\right]}{\sum_{n'=1,n'\neq n}^{N} \sum_{r=1}^{T} d_{kn'r}^{(s)} J\left[\delta_N^{-1}\left(\omega_{knt}^{(s)N} - \omega_{n'r}^{N}\right)\right]}, \tag{36}$$

To evaluate the term $\partial p_{i1nt}^{(s)}/\partial h_{nt}$, which appears in the definition of Y_{i2nt}, define

$$f_{i1nt}^{(s)} \equiv f_{i1}\left(\omega_{1nt}^{(s)} \mid d_{nt+s} = 1\right) \tag{37}$$

to be the probability density function for $\omega_{1nt}^{(s)}$, conditional on participating at date $t + s$. Likewise, let $f_{int}^{(s)} \equiv f_i\left(\omega_{1nt}^{(s)}\right)$ be the related probability density that is not conditioned on participating in period $t + s$ for $s = 1, \ldots, 3$. Denote their derivatives with respect to h_{nt}^* by $f_{i1nt}^{'(s)}$ and $f_{int}^{'(s)}$, respectively. We can then show that

$$\frac{\partial p_{i1nt}^{(s)}}{\partial h_{nt}} = \left[\frac{f_{i1nt}^{'(s)}}{f_{i1nt}^{(s)}} - \frac{f_{int}^{'(s)}}{f_{int}^{(s)}}\right] p_{1nt}^{(s)}, \quad s = 1, \ldots, 3. \tag{38}$$

We derive this expression using the representation of $p_{i1nt}^{(s)}$ as $p_{i1nt}^{(s)} = \Pr\left(d_{nt+s} = 1 \mid \omega_{1nt}^{(s)}\right) = \Pr(d_{nt+s} = 1) f_{i1nt}^{(s)} / f_{int}^{(s)}$. Differentiating this expression with respect to h_{nt} yields the above expression. The nonparametric estimates of $f_{i1nt}^{(s)}$ and $f_{int}^{(s)}$ are defined, respectively, as the numerators and denominators of $p_{iknt}^{(s)N}$ in equation (38). The estimates of $f_{i1nt}^{'(s)}$ and $f_{int}^{'(s)}$ are obtained from the derivatives of the estimates, $f_{i1nt}^{(s)N}$ and $f_{int}^{(s),N}$, with respect to h_{nt} (Pagan and Ullah, 1999).

6.4 Utility of Non-market Time

The identification result in Proposition (5.4) assumes that the distribution of $(\varepsilon_{0nt}, \varepsilon_{1nt})$ is known. Because we are not estimating the utility of non-market labor nonparametrically, we can relax that assumption, and allow for the variance to be unknown; specifically, we assume $(\varepsilon_{0nt}, \varepsilon_{1nt})$ is distributed as a Type I extreme value with variance parameter σ^2 and mean zero. This distributional assumption for the preference shocks implies that $Q^{-1}(p) = \sigma \ln[p/(1-p)]$, $\varphi_0(p) = \frac{\zeta}{\sigma} - \sigma \ln[(1-p)]$, and $\varphi_1(p) = \frac{\zeta}{\sigma} - \sigma \ln[p]$, where ζ is the Euler constant. Under this distribution assumption, Y_{i1nt} and Y_{i2nt} simplify to

(39)

$$Y_{i1nt} \equiv \eta_n^o \lambda_t^o \sum_{\tau=1}^{\Upsilon} I_{n\tau t} S_{i\tau t}^o(h_{nt}^*, z_{nt}^p, H_{nt-1}) + \sigma \sum_{s=1}^{3} \beta^s \ln\left(\frac{1-p_{i1nt}^{(s)}}{1-p_{i0nt}^{(s)}}\right) - \sigma \ln[p_{int}/(1-p_{int})]$$

and

$$Y_{i2nt} \equiv \eta_n^o \lambda_t^o \sum_{\tau=1}^{\Upsilon} I_{n\tau t} \frac{\partial S_{i\tau t}^o(h_{nt}^*, z_{nt}^p, H_{nt-1})}{\partial h_{nt}} + \sigma \sum_{s=1}^{3} \beta^s \left(1-p_{i1nt}^{(s)}\right)^{-1} \frac{\partial p_{i1nt}^{(s)}}{\partial h_{nt}} \tag{40}$$

since $Q^{-1}\left(p_{kit}^{(s)}\right) + \varphi_1\left(p_{kit}^{(s)}\right) - \varphi_0\left(p_{kit}^{(s)}\right) = 0$. Note that based on the three previous stages of estimation, the only unknown parameter in Y_{i1nt} and Y_{i2nt} is σ. Since it only affects the probabilities, it is trivially identified and estimable.

Finally, we specify the form of the non-market-hours utility function. We allow the fixed cost of participating to change over time in order to capture possible changes in home-production technology. It takes the form

$$u_{it0}(z_{nt}) = B_{ot} + z_{nt}' B_{i1}.$$

We assume the following functional form for the utility of non-market hours.

$$u_{i1}(z_{nt}, H_{nt-1}, l_{nt}) = z_{nt}' l_{nt} B_{i2} + \theta_{i0} l_{nt}^2 + \sum_{s=1}^{2} \theta_{is} l_{nt} l_{nt-s}$$

Under the above specifications, Lemma (5.3) yields

$$\begin{aligned} Y_{i1nt} &= -B_{0t} - z_{nt}' B_{i1} + z_{nt}' h_{nt} B_{i2} + \theta_{0i}\left(1 - l_{nt}^2\right) + \sum_{s=1}^{2} \theta_{si} h_{nt}(l_{nt-s} + \beta^s) + \xi_{i1nt} \\ Y_{i2nt} &= z_{nt}' B_{i2} + 2\theta_{i0} l_{nt} + \sum_{s=1}^{2} \theta_{si}(l_{nt-s} + \beta^s) + \xi_{i2nt} \end{aligned} \tag{41}$$

We then construct the empirical counterpart of that system by substituting for the estimated quantities: β, $\eta_n^o \lambda_t^o$, $S_{i\tau t}^o(h_{nt}, z_{nt}^p, H_{nt-1})$, p_{int}, $p_{i0nt}^{(s)}$, and $\frac{\partial p_{i1nt}^{(s)}}{\partial h_{nt}}$. We then base our estimation on that system of equations for the remaining parameters using a GMM estimator. The remaining details of the implementation and asymptotic properties of the estimator are in Appendix D.

7 Empirical Results

Table 3 contains the results from this estimation. These results are standard and consistent with estimates of these parameters in previous literature (see Gayle and Miller, 2004, and Altug and Miller, 1998, for similar estimates).

Figure 3 shows a significant increase in aggregate productivity in both occupations. This increase, however, was much larger in the professional occupations than in the nonprofessional occupations; productivity increased by 80% from the mid-1970s to the late 1980s in the professional occupations and by only 33% in the nonprofessional occupations.

The estimation results of the earnings equation are reported in Table 4. The overidentifying restrictions cannot be rejected at the 5% level of significance. The coefficient on current hours worked is larger in the professional occupations (183,392 versus 100,688).[18] The professional occupations have significantly higher returns to labor-market experience than the nonprofessional occupations. However, the returns to working fewer hours (part time) in the nonprofessional occupations is higher; this can be seen by comparing the linear and quadratic terms in current hours. The professional occupation has a lower depreciation rate of human capital; this can be seen by comparing the coefficients on h_{nt-1} versus h_{nt-2}. There is a larger cost of hiring a new worker (3032 versus 875). These estimates are between 2% and 5% of average annual earnings in our sample. This is within the range of vacancy cost estimates used in the search literature (see for example Hagedorn and Manovskii, 2005). These results are consistent with the empirical fact that women's representation in nonprofessional occupations is higher. The increase in productivity in the professional occupations relative to the nonprofessional occupations is consistent with the increase in women's representation in the professional occupations.

Table 5 contains the estimates of the fixed cost of labor-force participation. There is no significant difference in the cost of participation for men and women with the same years of completed education (notice that a positive coefficient implies higher disutility). A larger number years of completed education reduces the cost of working

[18]Notice that the coefficients on the hours and experience are large because hours are between zero and one.

for men and women equally. The effect of marital status is highly nonlinear and depends on the education level of the spouse. A married individual has lower costs of participating in the labor force. A married women who is married to a more educated man, however, has higher cost of participation. In contrast, a man who is married to a more educated women has lower cost of participation in the labor force.

Table 6 contains the results of our estimates on the utility of non-market hours. Again there are gender differences; whereas women with kids have higher costs of participating, conditional on working, they have lower disutility of working more hours; the opposite is true for men. The opposite is also true for education. That is, educated women have lower costs of participating, but education has no significant effect on the disutility of working more hours conditional on working in the labor market. Education, on the other hand, does not have any significant effect on men's costs of participation, but it does increase men's disutility of working more hours conditional on working in the labor market. Lastly, conditional on working in the labor market, marriage increases the disutility of working more hours. Conditional on working, having a more educated spouse decreases the disutility of working more hours for women, but not for men.

Table 7 contains the estimates for the time nonseparability in non-market hours. They show that there are significant complementaries between non-market hours across time for women. The results on complementaries between non-market hours across time for men are mixed. In particular, non-market hours for men are compliments one period back. However, they become substitutes two periods back.

The difference in the estimated coefficients for number of kids (both young and old kids) for men and women is suggestive evidence that women *specialize* in non-market work relative to men. This is consistent with Becker's (1965) theory of home-production division of labor. It should be noted that this evidence is only suggestive because we only use time spent working and do not distinguish between leisure and hours worked at home. However, our results, using the number of kids as a proxy for home-production hours, support this theory. These results are also supportive of the idea of cross-group complementaries in the utility function (asymmetric equilibrium).

7.1 Decomposing the Gender Earnings Gap

Next we decompose the earnings and wage gaps into four components: human capital (current and past hours worked in the market), firms' beliefs, the fixed effects, and other (education and age composition). The results are reported in Table 8, which has the median wage of a woman over the median wage of a man.[19] The wage gap for our sample is 87% and 76% for professionals and nonprofessionals, respectively. Our model predicts a wage gap of 92% and 81% for professionals and nonprofessionals, respectively. Of the 92% predicted wage gap in the professional occupations, 69% is due to the difference in human capital, 14% is due to differences in firms' beliefs, 6% is due to differences in the estimated fixed effects, and 11% is due to gender differences in education and age composition. In the nonprofessional occupation, of the 81% wage gap predicted by our model, 74% is due to differences in human capital, 12% is due to differences in employers' beliefs, 8% is due to differences in the estimated fixed effect, and 6% is due to differences in the education and age composition between men and women. The fixed effects may be capturing implicit discrimination which is not in our model. We find, however, that they account, for only a small fraction of the predicted wage gap, suggesting these other possible sources of discrimination may not be empirically very important.

We further analyze the sources of the change in the earnings gap over two disjoint time periods: 1974–1978 and 1984–1988. To do this, we calculate the median earnings gap in both time periods and express the difference as a percentage of the median earnings gap in the first period. The results are reported in Table 9. The raw changes in the earnings gap over the period are 30% and 24% for professionals and nonprofessionals, respectively. Our model predicts changes of 29% and 22% for the two occupations, respectively. We then decompose the predicted changes into changes due to differences in human capital, firms' beliefs, and education and age composition (labeled *Other* in Table 9). Changes in the differences in human capital over the two periods account for 67% and 65% of the changes in our two occupations, respectively (see Weinberger and Kuhn (2006) for similar results). Changes in firms' beliefs account for 8% and 6% of the changes in professional and nonprofessional, respectively, whereas education and age composition account for 25% and 29% of the changes in professionals and nonprofessionals, respectively.

[19] Where wage is computed as earnings divided by hours worked.

Human capital is the most important factor in explaining both the wage gap and the changes in the earnings gap over time. Human-capital accumulation, however, is endogenous to our model. Hence, the effect of private information is compounded into the human-capital effect. The most straight forward way to disentangled these different effects on human-capital accumulation is to solve the model. The problem with solving the model is that with private information there is the possibility of multiple equilibria. Therefore, there are no guarantees the equilibrium we solve for is the one actually played in the data, the parameter values of our model may be consistent with many different equilibria. Fortunately, we can get around this problem by solving the model under two different environments in which there is a unique equilibrium. The equilibrium solution in both environments is obtained by using backwards induction. We then compare the results from our solution with the actual data to obtain a measure of the effect of private information on the effect of human-capital accumulation on the earnings gap. The first environment we consider is one in which there are no hiring costs, and the second is one in which there are hiring costs but the information is symmetric.

7.1.1 Preferences and Labor Market Frictions

To quantify the effect of the gender preferences difference on the observed earnings gap, labor supply, and occupation composition we simulate the model under the assumption that there are no labor-market frictions. Without costs of hiring new workers, earnings equal the worker's productivity. Therefore, the main cause for differences between men and women's behavior and earnings are *preference and skills differences.* The rows labeled hiring cost in Tables 10 through 13 contain the results from the simulation without hiring cost.

To calculate the effect of hiring costs on the gender earnings gap, we take the two disjoint time periods and calculate the average of all the inputs to our model for each period. These inputs include the demographic characteristics, aggregate shocks, the marginal utility of wealth, the fixed effects, and the estimated transition probabilities of marital status and number of kids. We then solve the model in both time periods, setting γ_τ equal to zero. Then we calculate the implied changes in the earnings gap. Without hiring costs, the changes in earnings and human capital over the years occur because of the aggregate productivity shocks, demographic characteristics, and costs of participation in the labor market. We decompose these changes in the same way they are calculated in Table 9. Then we express the amount of the change attributed to human capital in the same way as in Table 9. We find that 38% and 32% of this change in human capital in the professional and nonprofessional occupations, respectively, are due to labor market frictions.

In Table 11, the predicted earnings gap in the professional occupations under this simulation would have been 81% in 1974–1978 and 96% in 1984–1988. Thus, the gap would have been substantially smaller with no labor market frictions, and would have almost disappeared in the late 80's. In order to understand this, we look at the effect of the frictions in the labor market on labor supply. Table 11 shows that with no labor-market frictions women's participation rate in the 1970s and 1980s would have been *lower* (56% versus 62% and 62% versus 70%, respectively). Table 11 shows that the hours worked by professional women who participate in an economy with no labor-market frictions is larger than it is in the benchmark model (1980 versus 1690 and 2050 versus 1904, respectively). In fact, in the 1980s women who participate in an economy with no frictions are working almost as much as men are (2050 versus 2100). A similar pattern is observed in the nonprofessional occupations.

This exercise demonstrates that the effect of preferences differences on the gender earnings gap is smaller than the effect of labor market frictions. In the 80's the earnings gap would have almost vanished in an economy with no labor market frictions.

7.1.2 Discrimination

We simulate the model under symmetric information using the same inputs as in the frictionless economy. We solve the model backward, calculating the actual probability of working during the next period in the same occupation conditional on working today in that occupation. This calculation is conditioned on all the information known by the worker today. We then compute the earnings gap, participation rates, hours worked, and occupational composition for the two periods. The difference between the simulated outcomes and outcomes in the benchmark economy are due to the effect of discrimination.

The second entry in Table 10 presents the effect of the asymmetric information on the change in human capital between the two periods. We find that the change in human capital would have been smaller by 12% in professional occupations and 13% in non-professional occupations, under symmetric information. Therefore, the effect of discrimination on the change in the earnings gap in professional occupations is 12% from the change in human capital 67% together with the 8% of the direct effect on earnings amounts to 16% in total. Similar calculation reveals that the effect of asymmetric information on the change in the earnings gap is 14.5% in nonprofessional occupations.

Table 11 shows that under symmetric information, the earnings gap in professional occupations would have been substantially smaller in both periods: 71% instead of 48% in the 1970s and 82% instead of 62% in the 1980s. Thus, 56% and 43%, respectively of the observed earnings gap can be attributed to discrimination. Table 12 demonstrates that participation would have been *smaller* under symmetric information: 51% instead of 62% in the 70's, 57% instead of 70 % in the 80's in professional occupations; women would have represented 28% of the professional occupations instead of 34% (working 1820 hours instead of 1640 in the 70's and 1990 instead of 1904 in the 80's), similar patterns are found in non-professional occupations.

The above results demonstrate two important points. First, private information has a significant effect on labor market behavior. In an economy with private information, participation in the labor market has *larger* effects on future earnings because of the signalling effect. Lack of attachment is "penalized," and may affect beliefs and future earnings. Second, the informational frictions may also affect selection of women who participate; on the margin, women with higher costs of participation are participating more. This is because participation is providing a signal that she is more "attached". This is also consistent with female representation in professional occupations being lower under symmetric information. That is, under asymmetric information rewards to greater attachment may be larger in professional occupations. Moreover, women work *less hours* in the benchmark economy (with private information). This may be partly due to the effect of the private information on the *selection* of women who participate in the labor market. Differences in the selection of the "types" of women who participate due to the effect of the signalling, and the extra reward of women on the margin who have higher disutility from working.

7.1.3 Demographics, Participation Costs and Aggregate Productivity

To compute the effect of changes in demographic characteristics (marital status, number of kids, years of completed education, and spouse education) on the change in human capital, we set γ_τ equal to zero. We then solve the model for the period 1984-

1988 holding all the demographic characteristics and their transition probabilities at their 1974-1978 levels. By doing so, we shut down the effect of change of these characteristics. This accounts for 28% and 34% of the changes in human capital in the two occupations, respectively, over and above those found in the model with no hiring costs. We then repeat the exercise above to obtain the effect of home-production shocks and aggregate productivity shocks on the change in the earnings gap.

To calculate the effect of an aggregate shock to the utility of participating in the labor market, we hold it fixed at its 1974–1978 level in both time periods, while allowing all other input to vary across the to periods. This only accounted for 2% and 3% of the changes over and above the 28% and 34% of the changes accounted for by the measure in the model in which there are no hiring costs.

In order calculate the effect of aggregate shock to market productivity, we hold it constant across the two time periods while allowing all other input to vary. This accounted for 18% and 11% of the changes over and above the 28% and 34% of changes accounted for in the frictionless model.

8 Conclusion

This paper analyzes the sources of male and female labor market outcomes gap, and the main driving forces in its narrowing from 1968 to 1992. To do that it develops a model in which men and women may differ in preferences and skill, but may give rise to endogenous gap even if there are no ex-ante differences between men and women. The model emphasizes the interaction of future expected length of employment spells with past labor market experience in the presence of private information and uncertainty; this interaction may give rise and amplify differences in life-cycle human capital investment pattern of women and men. It differs from the standard incomplete information discrimination models in that it focuses on the uncertainty about the turnover propensity of workers instead of uncertainty about productivity differences across groups.

We prove that the model is semiparametrically identified and develop a three-step estimation technique. The identification results is stronger than what is normally obtain in the standard discrete models because of the mixture of continuous and discrete choices. The presence of the continuous choice increases the identifying power of the model. The estimates of the structural parameters are $\sqrt{N}$ consistent and asymptotically normal, although the second step is estimated nonparametrically. The estimator demonstrates that under the assumption that the researcher observes retrospectively part of the private information of the agent, the scope of game theoretic models that can be structurally estimated is increased.

The empirical investigation finds that while women may have higher disutility of working, the differences in observed labor market outcomes which can be attributed to preferences and skill differences are comparatively small. Private information and the existence of hiring costs amplify these differences and account for much of the observed labor market outcomes gender gap. A decomposition exercise show that increase in human capital of women is the main driving force behind the narrowing of the earnings gap. However, the increase human capital is driven by increase in market productivity and demographic changes (mainly fertility decline). The estimation results do not support the hypothesis that changes in home-production technology explain the increase in women's human capital.

A Data Description

We used data from the Family File, the Childbirth and Adoption History File, the Retrospective Occupation File, and the Marriage History File of the PSID. The Family File contains a separate record for each member of each household included in the survey in a given year, but includes labor income, hours worked, and years of completed education only for Heads and Wives. The Childbirth and Adoption History File contains information collected in the 1985–1992 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 1992. Each set of records for a specified individual contains all known cumulative data about the timing and circumstances of his/her childbirth and adoption experience up to and including 1992, 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* here means individuals of childbearing age in responding families. Similarly, the 1985–1992 Marriage History file contains retrospective histories of marriages for individuals of marriage-eligible age living in PSID families between 1985 and 1992. 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 1992, 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, which contains 24,762 individuals. We then drop any individual who was in the survey for four years or less, this selection criteria eliminated 4,300 individuals from our sample. We then drop all individuals who were older than 65 in 1967, this eliminated a further 3,331 individuals. We then drop all individuals that were less than 25 years old in 1991, this eliminated an additional 2,385 individuals. We then drop all individuals who were neither Head nor Wife in our sample for at least 4 years. this eliminated a further 4,567 individuals from our sample.

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. The second step was taken to account 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 from above in 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, while the relevant value for this variable in the subsequent years and for the value of food consumed at home and eating out is \$9,999.00. Taken by itself, the truncation of different consumption variables resulted in a loss of 467 person-years. 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. Issues with the age variable resulted in a loss of 462 individuals.

The race of the individual and the region where they were residing at the time of the interview were obtained from the Family portion of the data record. We defined the region variable to be 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 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, 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 Spanish American. After 1984, the variable was coded in such a way that 1–6 correspond to the categories White, Black, Hispanic, American Indian, Aleutian or Eskimo, and Asian or Pacific Islander, respectively. A value of 7 denotes the other category, a value of 9 denotes missing data. We used all available information for all the years to assign the race of the individual for years in the sample when that information was available. We then drop all individuals that were not coded as White.

The marital status of a women in our subsample was determined from the Marriage History File. The number of individuals in the household and the total number of children within that household were also determined from the family-level variables of the same name. In 1968, a code for missing data (equal to 99) was allowed for the first variable, but in the other years, missing data were assigned. The second variable was truncated above the value of 9 for the interviewing years 1968 and 1971. After 1975, this variable denotes the actual number of children in the family unit.

Household income was measured from the PSID variable, *Total Family Money Income*, which included taxable income of Head and Wife, total transfers of Head and Wife, taxable income of others in the family units and their total transfer payments.

We used the PSID Retrospective Occupation File to obtain a consistent three-digit occupational code for our sample. First we eliminated all self-employed, dual-employed, government workers, Farmers and Farm Managers, Farm Laborers and Farm Foremen, Armed Forces, and Private Household workers. The professional occupation is made up of the following classifications: Professional, Technical, and Kindred Worker; Managers and Administrators, Except Farm Managers; and some categories of Sales Workers. The Sales Workers included in professionals are Advertising and Salesmen; Insurance Agents; Brokers and Underwriters: Stock and Bond

Salesmen. The nonprofessional occupation is made up of the following classifications: Sales Workers (not included in Professional); Clerical and Kindred Workers; Craftsmen and Kindred workers; Operatives, Except Transport; Transport Equipment Operatives; Laborers, Except Farm; and Service Workers, Except Private Household.[20]

We used two different deflators to convert such nominal quantities 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.

B Theoretical Result Proofs

Let $\overline{z}^*_{\tau t}(H_{t-1}, z^p_t)$ be the type with the lowest costs of working and highest returns (both current and future expected) ; thus, this type will work the most hours in occupation τ. Similarly, define $\underline{z}^*_{\tau t}(H_{t-1}, z^p_t)$ to be the type with the highest disutility from hours worked and lowest returns (this type also works the smallest fraction of time in the occupation). Let $\overline{\mu}_\tau = \mu_{it}(\overline{z}^*_{\tau t} \mid H_{t-1}, z^P_t)$ be the beliefs about the type $\overline{z}^*_{\tau t}$, and $\underline{\mu}_\tau = \mu_{it}(\underline{z}^*_{\tau t} \mid H_{t-1}, z^p_t)$ be the beliefs about the type $\underline{z}^*_{\tau t}$.

Assumption B.1 (Off-equilibrium path) *if* $\forall z^*_t$, $\sigma_{it}(a_t \mid H_{t-1}, z^*_t) = 0$, *then* $\mu_{it}(z^*_t \mid H_{t-1}, z^p_t) = \underline{\mu}_\tau$ *if* $h_t < \underline{h}_{\tau t}$ *and* $\mu_{it}(z^*_t \mid H_{t-1}, z^P_t) = \overline{\mu}_\tau$ *if* $h_t > \overline{h}_{t\tau}$.

Proof of Proposition 4.1.

The free-entry condition implies that in equilibrium, the expected value of a vacancy in each occupation at any period, $\pi_{\tau t}$, is zero. Thus, $\pi_{\tau t}$ is the continuation value of hiring a new worker in occupation τ in period t. Define $\pi^e_{\tau t}$ to be the continuation value to the current employer in occupation τ in period t. That is, $\pi^e_{\tau t}$ is the expected profit from employing a worker who is employed in the firm for more than one period. We use this to derive the optimal contract by solving backwards:

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

At time $t = T$ (the worker's final year), the free-entry condition implies that a new employer expects no profit from offering the worker a contract. The expected profit from offering a contract, $S_{i\tau t}(h_t, H_{t-1,} z_t^p)$, i.e.

(42)
$$\pi_{\tau T} = y_{\tau T}(h_T, H_{T-1}, z_T^p) - S_{i\tau T}(h_T, H_{T-1}, z_T^p) - \gamma_\tau = 0.$$

Therefore,

$$S_{i\tau T}(h_T, H_{T-1}, z_T^p) = y_{\tau T}\,(h_T, H_{T-1}, z_T^p) - \gamma_\tau.$$

The current employer's profit, substituting the earnings, is

(43)
$$\pi_{\tau T}^e = y_{\tau T}(h_T, H_{T-1}, z_T^p) - S_{i\tau T}(h_T, H_{T-1}, z_T^p) = \gamma_\tau$$

Consider a potential employer making an offer at time $t = T - 1$:

$$\begin{aligned}\pi_{\tau T-1} = y_{\tau T-1}(h_{T-1}, H_{T-2}, z_{T-1}^p) \quad &- \quad \gamma_\tau - S_{i\tau T-1}(h_{T-1}, H_{T-1}, z_{T-1}^p) \\ &+ \quad \beta \widetilde{p}_{i\tau T}(z_{T-1}^p, H_{T-1})\pi_{\tau T}^e = 0.\end{aligned}$$

Thus,

$$S_{i\tau T-1}(h_{T-1}, H_{T-2}, z_{T-1}^p) = y_{\tau T-1}(h_{T-1}, H_{T-2}, z_{T-1}^p) - \gamma_\tau(1 - \beta \widetilde{p}_{i\tau T}(z_{T-1}^p, H_{T-1})).$$

The current employer's profit in period $T-1$ is γ_τ. Solving backwards, at any period $s < T$, we obtain

(44)
$$S_{i\tau T-s}(h_{T-s}, H_{T-s-1}, z_{T-s}^p) = y_{\tau T-s}(h_{T-s}, H_{T-s-1}, z_{T-s}^p) - \gamma_\tau(1 - \beta \widetilde{p}_{i\tau T-s+1}(z_{T-s}^p, H_{T-s})).$$

Given the beliefs and worker's strategy to accept the highest offer by Lemma 4.1, and other firm's strategies, equation (8) is the competitive salary schedule. ■

Proof of Lemma 4.1. We show that it is optimal to accept the highest salary offer. First, given any choice of h_t, the current utility is increasing in salary, and second, given any beliefs and any h_t, the continuation value of the worker is non-decreasing if she/he chooses a higher salary. The increase in salary enters the value function through the Frisch demand; see equation 9. Since $\eta\lambda_t > 0$, the current utility is increasing in $S_{i\tau t}(h_t, H_{t-1}, z_t^p)$. Next, we show that given choice of h, the continuation value is unchanged by a higher salary. We begin by showing that a higher salary does not affect tomorrow's beliefs. First, given the hours choice, one occupation is chosen by Assumption 3.1. Changing employers within occupation with the same hours worked does not change the beliefs. Second, we assume salary is not observed by outside employers, hence it is not part of employment history and does not affect beliefs. Therefore, accepting the highest salary given hours is optimal. ■

Proof of Proposition 4.2. Equation (4) describes the optimal consumption behavior on and off the equilibrium path. Off-equilibrium path is when salaries for given hours differ from the optimal salary schedule in Proposition 4.1. They occur with probability zero. An optimal consumption strategy response to a one period unanticipated salary shock is also optimal response to a single deviation by employer. Thus, the optimal consumption plan in equation (4) is optimal.

Using the Bellman principle, the ex-ante value function for an individual who chooses to participate in the labor force in period t and to behave optimally thereafter is

$$\begin{aligned}(45) \qquad & V_{1it}(\omega_t) \\ = \quad & \max_{h_t \in (0,1); \{I_{\tau t}\}_{\tau=1}^{\Upsilon}} \Bigg[u_{0it}(z_t) + u_{1i}(l_t, H_{t-1}, z_t) + \eta \lambda_t \sum_{\tau=1}^{\Upsilon} I_{\tau t} S_{i\tau t}(h_t, z_t^p, H_{t-1}) \\ & + \beta E_t \{ p_{it+1} V_{1it+1}(\omega_{t+1}) \mid \omega_t, h_t > 0 \} \\ & + \beta E_t \{ (1 - p_{it+1}) V_{0it+1}(\omega_{t+1}) \mid \omega_t, h_t > 0 \} \Bigg] .\end{aligned}$$

Then $Q(V_{1it}(\omega_t) - V_{0it}(\omega_t))$ and $h_{it}^*(\omega_t)$ follow directly from equations (4)–(10) along with the above equation and Lemma 1 of Hotz and Miller (1993). $I_{i\tau t}^0(\omega_t)$ follows directly from Assumptions 3.1 and 3.2. Last, we need to show that given off-equilibrium path beliefs, by Assumption B.1, workers who work fewer hours than the minimal (optimal) hours receive the productivity plus the beliefs component attached to the marginal type who works the least hours $(\underline{z}_{\tau t}^* \mid H_{t-1}, z_t^P)$. Because of this, the future returns of all workers are weakly larger and the costs weakly higher. Therefore the deviation is not profitable. The same argument applies to working more than the maximum optimal hours, as beliefs are not adjusted to be higher than the beliefs for hours worked more than $\overline{z}_{\tau t}^*$. Thus, a deviation implies that the worker has higher costs of working and lower returns. ■

Proof of Proposition 4.3(1). Now we show that the contract that satisfies the zero-profit condition is optimal. In order to establish that, we need to show that given other firms offering the competitive rate, the worker's strategy, and the firm's beliefs, no firm can deviate from the competitive rate and strictly increase its expected profits. First, we show that by offering a lower salary, the firm cannot increase profit. From Lemma 4.1, workers accept the highest offer, thus a deviation to a lower salary implies the worker rejects the offer and the payoff is zero.

Consider a firm offering a salary $\widetilde{s}_{i\tau t}$ for h_t such that $\widetilde{s}_{i\tau t} > s^0_{i\tau t}(h_t, H_{t-1}, z^p_t)$. The workers' state variables are not a function of past salaries. Therefore, at $t+1$, the worker's state variable remains ω_{t+1} and competing firms' offers are unchanged $s^0_{i\tau t+1}(h_{t+1}, H_t, z^p_{t+1})$ as past salary histories are unobserved.

Note that $Q(V_{1it+1}(\omega_{t+1}) - V_{0it+1}(\omega_{t+1}))$, $I^0_{i\tau t+1}(\omega_{t+1})$ and $h_{it+1}(\omega_{t+1})$ remain the same. First the function $Q()$ is only a function of $(\varepsilon_{0t}, \varepsilon_{1t})$ by Lemma 1 of Hotz and Miller (1993), hence is not affected by salary. Second, because of the additively separability of leisure and consumption, the asset market assumptions and that the workers move first, the functions $V_{1it+1}()$ and $V_{0it+1}()$ remain the same.

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

$$\begin{aligned}\widetilde{p}_{i\tau t+1}(H_t, z^p_t, \widetilde{s}_{\tau t}) &= \int Q(V_{1it+1}(\omega_{t+1}) - V_{0it+1}(\omega_{t+1}))I_{i\tau t+1}(\omega_{t+1})\mu_{it+1}(z^*_{t+1}|H_t, z^p_t)\,\mathrm{d}z^*_{t+1} \\ &= \widetilde{p}_{i\tau t+1}(H_t, z^p_t).\end{aligned}$$

Therefore, as established in equation (44), the continuation expected profit can be written as

$$\begin{aligned}\pi(\widetilde{s}_{\tau t}, .) &= y_{\tau t}(h_t, H_{t-1}, z^p_t) - \gamma_\tau - \widetilde{s}_{i\tau t} + \beta\gamma_\tau\widetilde{p}_{i\tau t+1}(H_t, z^p_t) \\ &< y_{\tau t}(h_t, H_{t-1}, z^p_t) - \gamma\tau - s^0_{i\tau t} + \beta\gamma_\tau\widetilde{p}_{i\tau t+1}(H_t, z^p_t) = 0.\end{aligned}$$

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

Before proceeding to the proof of part two of Proposition 4.3, we first establish a set of necessary and sufficient conditions for existence of equilibrium in our model. Note that the optimal hours worked and participation is a function of the firms' beliefs about the next period's participation. To see this, consider the beliefs about period

T's participation:

$$\widetilde{p}_{i\tau T} = \int Q(\omega_T) I_{i\tau T}(\omega_T) \mu_{iT} \left(z_T^* \mid H_T, z_T^P \right) \mathrm{d}z_T^*. \tag{46}$$

Recall that $\omega_T = (z_T, H_{T-1}, \eta\lambda_T)$ hence $\widetilde{p}_{i\tau T}$ is a function of H_T. However H_T is a function of the sequence $\widetilde{p}_{i\tau T}, \ldots, \widetilde{p}_{i\tau 2}$ in equilibrium. Therefore $\widetilde{p}_{i\tau T}$ is defined as an implicit function of itself. In fact, there is a triangular system of implicit equilibrium beliefs of the following form.

$$\begin{aligned}
\widetilde{p}_{i\tau T} &= \Gamma_{iT}(\widetilde{p}_{i\tau T}, \ldots, \widetilde{p}_{i\tau 2}) \\
\widetilde{p}_{i\tau T-1} &= \Gamma_{iT-1}(\widetilde{p}_{i\tau T-1}, \ldots, \widetilde{p}_{i\tau 2}) \\
&\vdots \\
\widetilde{p}_{i\tau 2} &= \Gamma_{i2}(\widetilde{p}_{i\tau 2}),
\end{aligned} \tag{47}$$

where Γ_{it} is the RHS of equation (46).

Corollary B.1 *A necessary and sufficient condition for existence of equilibrium in our model is that there exist a fixed point in* $\{\widetilde{p}_{i\tau 2}, \ldots, \widetilde{p}_{i\tau T}\}_{i \in \{w,m\}}^{\tau \in \Upsilon}$ *of the system of equations in (47) for all* τ *and* i

Proof of Corollary B.1. In order to prove this, we first show necessity. Suppose there exists an equilibrium in which equation (47) 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 (44) 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 (47) has a fixed point. Then, for any t, τ and i, by Lemma 4.1 the competitive salary schedule exists. Given the competitive salary schedule, by Proposition 4.2 the worker's strategies for hours, participation, occupation, and consumption exist and are unique. Hence conditions 1, 2, and 3 of Definition 3.1 exist (mutual best responses by construction, the beliefs satisfy Bayes' law). ■

Existence of equilibrium in our model is established by showing there exists a fixed point to the system of equations in (47).

Proof of Proposition 4.3(2). Given the triangular nature of the system of equations in (47), 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 any period t, occupation τ, and gender i, $\widetilde{p}_{i\tau t+1}$ is the solution to

$$\begin{aligned} \widetilde{p}_{i\tau t+1} &= \int_{z_t^*} \Big\{ \int_{z_{t+1}^*} f_{i0}^*(z_{t+1}^* \mid z_t^*) Q(V_{1it+1}(\omega_{t+1}) - V_{0it+1}(\omega_{t+1})) \\ & I_{i\tau t+1}(\underline{h}_{\tau t+1} < h_{it+1}^*(\omega_{t+1}) < \overline{h}_{\tau t+1}) \, \mathrm{d}z_{t+1}^* \Big\} \, \widetilde{\mu}_{it}(z_t^* \mid H_t, z_t^p) \, \mathrm{d}z_t^*. \end{aligned} \tag{48}$$

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

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

$$\int_{\underline{z}_{\tau t+1}^*}^{\overline{z}_{\tau t+1}^*} f_{i0}^*(z_{t+1}^* \mid z_t^*) Q_{t+1}() \, \mathrm{d}z_{nt+1}^*.$$

Since $h_t(\widetilde{p}_{i\tau t+1})$ is continuous in $\widetilde{p}_{i\tau t+1}$ and $Q(h_t(\widetilde{p}_{i\tau t+1}), .)$ is continuous in h_t, we only need to show that the functions $\underline{h}_{\tau t+1}^{-1}$ and $\overline{h}_{\tau t+1}^{-1}$ are continuous in $\widetilde{p}_{i\tau t+1}$. From the continuity of the production function in each occupation in all factors of production(Assumption 3.2), $\underline{h}_{\tau t+1}$ and $\overline{h}_{\tau t+1}$ are continuous in h_t and $h_t(\widetilde{p}_{i\tau t+1}, .)$ is continuous in $\widetilde{p}_{i\tau t+1}$. Hence, their inverses are continuous in $\widetilde{p}_{i\tau t+1}$. Therefore, there exists a solution to every period's beliefs separately.

The fact that there exists a one-to-one mapping between the posterior beliefs and the implied participation probability in equation (48) comes directly from the requirement in Bayesian games that beliefs are consistent with player's strategies and Bayes' rule be satisfied (when possible); this condition holds by construction. Therefore, the expected profit condition on salary is also correct on the equilibrium path. ■

C Identification Proofs

Proof of Lemma 5.1. Without loss of generality, assume that

$$\frac{\partial u_2 i(c_{nt}, z_{nt}, \varepsilon_{2t})}{\partial c_{nt}} = \exp(u_{2c}(c_{nt}))\exp(-u_{2z}(z_{nt}))\exp(-\varepsilon_{2nt}).$$

The above equation satisfies Assumption (5.1). The explicit functional form simplifies the exposition. Equation (4) implies that the Euler for consumption is

$$\exp(u_{2c}(c_{nt}))\exp(-u_{2z}(z_{nt}))\exp(-\varepsilon_{2nt}) = \eta_n\lambda_t. \tag{49}$$

Taking the log and then first derivative of equation (49) and rearranging gives us

$$\Delta u_{2c}(c_{nt}) = \Delta u_{2z}(z_{nt}) + \Delta\log(\lambda_t) + \Delta\varepsilon_{2nt} \tag{50}$$

By assumption 5.2(1), then,

$$E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}] = \Delta u_{2z}(z_{nt}) + \log(\lambda_t) \tag{51}$$

Taking the derivative of $E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}]$ with respect to z_{cnt} and z_{cnt-1}, respectively, and integrating back up to z_{cnt} and z_{cnt-1}, respectively, gives

$$u_{2zi}(z_{nt}) = u_{2z}(\underline{z_c}, z_{c'nt}) + \int_{\underline{z_c}}^{z_{cnt}} \left\{\frac{\partial E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}]}{\partial z_c}\right\} \mathrm{d}z_c \tag{52}$$

$$u_{2zi}(z_{nt-1}) = u_{2z}(\underline{z_c}, z_{c'nt-1}) + \int_{\underline{z_c}}^{z_{cnt-1}} \left\{\frac{\partial E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}]}{\partial z_{c-1}}\right\} \mathrm{d}z_{c-1}, \tag{53}$$

which by Assumption 5.3(2) and from Chesher's (2007) results is identified. Therefore

$$\begin{aligned}\Delta\log(\lambda_t) &= E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}] - \int_{\underline{z_c}}^{z_{cnt}} \left\{\frac{\partial E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}]}{\partial z_c}\right\} \mathrm{d}z_c \\ &\quad + \int_{\underline{z_c}}^{z_{cnt-1}} \left\{\frac{\partial E[\Delta u_{2c}(c_{nt})|z_{nt}, z_{nt-1}]}{\partial z_{c-1}}\right\} \mathrm{d}z_{c-1}\end{aligned} \tag{54}$$

and, by Assumption 5.2(1),

$$\log(\lambda_1) = E[\Delta u_{2c}(c_{n1})|z_{n1}] - \int_{z_c}^{z_{cn1}} \left\{ \frac{\partial E[\Delta u_{2c}(c_{nt})|z_{n1}]}{\partial z_c} \right\} \mathrm{d}z_c. \tag{55}$$

Hence, λ_t is identified. Finally, by Assumption 5.2(2), we have

$$\log(\eta_n) = E_t\{u_{2c}(c_{nt}) - \log(\lambda_t) - u_{2z}(z_{nt})|z_{nt}\}. \tag{56}$$

Using Chesher's (2007) result and the fact that $u_{2c}()$ is assumed known, we use the results from equations (54), (55), and (56). ■

Proof of Lemma 5.2. This result is show by proving that

$$\beta\gamma_\tau = \frac{\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z_{nt}^p]}{\triangle\ \widetilde{p}_{\tau t+1}(H_{nt}, z_{nt}^p)}, \tag{57}$$

$$\begin{aligned} y_{\tau t}(h_{nt}, H_{nt-1}, z_{nt}^p) \; &= \; \int_0^{h_{nt}} \left\{ \partial \left[\frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|H_{nt}, z_{nt}^p]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z_{nt}^p, i]} \right] \backslash \partial h \right\} \mathrm{d}h \\ &\quad - \int_0^{h_{nt}} \left\{ \partial \left[\frac{\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z_{nt}^p]}{\triangle\ \widetilde{p}_{\tau t+1}(H_{nt}, z_{nt}^p)} \right] \backslash \partial h \right\} \mathrm{d}h \end{aligned} \tag{58}$$

for $i \in \{m, w\}$ and

$$\begin{aligned} \gamma_\tau \; = \; & \left\{ \int_0^{h_{nt}} \left\{ \partial \left[\frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|H_{nt}, z_{nt}^p]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z_{nt}^p, i]} \right] \backslash \partial h \right\} \mathrm{d}h - \frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|H_{nt}, z_{nt}^p]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z_{nt}^p, i]} \right\} \\ & - \left\{ \int_0^{h_{nt}} \left\{ \partial \left[\frac{\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z_{nt}^p]}{\triangle\ \widetilde{p}_{\tau,t+1}(H_{nt}, z_{nt}^p)} \right] \backslash \partial h \right\} \mathrm{d}h - \frac{\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z_{nt}^p]}{\triangle\ \widetilde{p}_{\tau,t+1}(H_{nt}, z_{nt}^p)} \right\} \end{aligned} \tag{59}$$

for $i \in \{m, w\}$and where

$$\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z_{nt}^p] = E_t \left[\frac{E_t[d_{nt}I_{n\tau t}S_{mn\tau t}|H_{nt}, z_{nt}^p]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z_{nt}^p, i = m]} - \frac{E_t[d_{nt}I_{n\tau t}S_{fn\tau t}|H_{nt}, z_{nt}^p]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z_{nt}^p, i = f]} \right].$$

Applying the results from Chesher (2007), all these parameters are identified because data is informative about $\widetilde{p}_{i\tau,t+1}(H_t, z_t^p)$ by part two of Proposition (4.3). *Note that the two overidentifying restrictions come from equations(58) and(59). There are each one parameter to identify and there each two equations, one for each gender.*

From Proposition (4.1), the zero-profit condition implies that

$$E_t[d_{nt}I_{n\tau t}(S_{in\tau t} - y_{\tau t}(h_{nt}, H_{nt-1}, z^p_{nt}) + \gamma_\tau - d_{nt+1}I_{n\tau t+1}\beta\gamma_\tau)|H_{nt}, z^p_{nt}, i] = 0. \tag{60}$$

Rearranging and noting that

$$\widetilde{p}_{i\tau,t+1}(H_{nt}, z^p_{nt}) = \frac{E_t[d_{nt}I_{n\tau t}d_{nt+1}I_{n\tau t+1}|H_{nt}, z^p_{nt}, i]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z^p_{nt}, i]}, \tag{61}$$

we can write the zero-profit condition as

$$\frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|H_{nt}, z^p_{nt}]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z^p_{nt}, i]} = y_{\tau t}(h_{nt}, H_{nt-1}, z^p_{nt}) - \gamma_\tau + \beta\gamma_\tau\widetilde{p}_{i\tau t+1}(H_{nt}, z^p_{nt}). \tag{62}$$

Taking the difference between equation (62) for men and women and rearranging gives

$$\beta\gamma_\tau = \frac{\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z^p_{nt}]}{\triangle\ \widetilde{p}_{\tau t+1}(H_{nt}, z^p_{nt})}, \tag{63}$$

which is well defined by Assumption 5.4(1). Substituting equation (63) into (62) for men and women gives the following system of equations.

$$\begin{aligned} y_{\tau t}(h_{nt}, H_{nt-1}, z^p_{nt}) - \gamma_\tau = & \frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|H_{nt}, z^p_{nt}]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z^p_{nt}, i]} \\ & - \frac{\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z^p_{nt}] \times \widetilde{p}_{i\tau t+1}(H_{nt}, z^p_{nt})}{\triangle\ \widetilde{p}_{\tau t+1}(H_{nt}, z^p_{nt})}, \ i = \{m, w\} \end{aligned} \tag{64}$$

Differentiating equation (64) with respect to hours and then integrating,

$$\begin{aligned} y_{\tau t}(h_{nt}, H_{nt-1}, z^p_{nt}) \quad = \quad & y_{\tau t}(0, H_{nt-1}, z^p_{nt}) + \int_0^{h_{nt}} \left\{ \partial \left[\frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|H_{nt}, z^p_{nt}]}{E_t[d_{nt}I_{n\tau t}|H_{nt}, z^p_{nt}, i]} \right] \backslash \partial h \right\} \mathrm{d}h \\ & - \int_0^{h_{nt}} \left\{ \partial \left[\frac{\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|H_{nt}, z^p_{nt}]}{\triangle\ \widetilde{p}_{\tau t+1}(H_{nt}, z^p_{nt})} \right] \backslash \partial h \right\} \mathrm{d}h \end{aligned} \tag{65}$$

and by Assumption 5.4(2), we have the lemma's results. By substituting (65) into (64) and rearranging, we obtain the final equation in the lemma. ■

Proof of Lemma (5.3). Define the errors as

$$
\begin{aligned}
\xi_{i1nt} = & \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[U_{0it+s}\left(\omega_{0t}^{(s)}\right) - U_{0it+s}\left(\omega_{1t}^{(s)}\right)\right] \\
& + p_{0it}^{o(s)}\left[Q^{-1}\left(p_{0it}^{o(s)}\right) + \varphi_1\left(p_{0it}^{o(s)}\right) - \varphi_0\left(p_{0it}^{o(s)}\right)\right] \\
& - p_{1it}^{o(s)}\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right] \\
& - E_t^0\left\{\sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[U_{0it+s}\left(\omega_{0t}^{(s)}\right) - U_{0it+1s}\left(\omega_{1t}^{(s)}\right)\right]\right. \\
& + p_{0it}^{o(s)}\left[Q^{-1}\left(p_{0it}^{o(s)}\right) + \varphi_1\left(p_{0it}^{o(s)}\right) - \varphi_0\left(p_{0it}^{o(s)}\right)\right] \\
& \left. - p_{1it}^{o(s)}\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right]\right\}
\end{aligned} \tag{66}
$$

and

$$
\begin{aligned}
\xi_{i2nt} = & \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{\partial U_{0it+s}\left(\omega_{1t}^{(s)}\right)}{\partial h_t} + \frac{\partial \varphi_0\left(p_{1it}^{o(s)}\right)}{\partial h_t}\right] \\
& + \frac{\partial p_{1it}^{o(s)}}{\partial h_t}\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right] \\
& + p_{1it}^{(s)} \frac{\partial\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right]}{\partial h_t} \\
& - E_t^o\left\{\sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{\partial U_{0it+s}\left(\omega_{1t}^{(s)}\right)}{\partial h_t} + \frac{\partial \varphi_0\left(p_{1it}^{(s)}\right)}{\partial h_t}\right]\right. \\
& + \frac{\partial p_{1it}^{o(s)}}{\partial h_t}\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right] \\
& \left. + p_{1it}^{o(s)} \frac{\partial\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right]}{\partial h_t}\right\}.
\end{aligned} \tag{67}
$$

Given these definitions, the result follows immediately. ■

Proof of Lemma 5.4.

To establish the results, we prove that

$$
u_{0it}(z_{nt}) = C_{1it} - E_t^o[Y_{i1nt}|\omega_{nt}] + \frac{1}{2}\int_0^{h_{nt}} \left\{E_t^o[Y_{i2nt}|\omega_{nt}] + \frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h}\right\} \mathrm{d}h \tag{68}
$$

and

(69)
$$u_{1i}(l_{nt}, z_{nt}, H_{nt-1}) = C_{2i}(z_{nt}, H_{nt-1}) + \frac{1}{2}\int_0^{h_{nt}} \left\{ E_t^o[Y_{i2nt}|\omega_{nt}] + \frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h} \right\} \mathrm{d}h,$$

where

$$C_{1it} = E_t^o \left\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[u_{i1}\left(1, z_{1nt}^{(s)}, H_{1nt-1}^{(s)}\right) - u_{i1}\left(1, z_{0nt}^{(s)}, H_{0nt-1}^{(s)}\right)\right]\right\}$$

and $C_{2i}(z_{nt}, H_{nt-1}) = u_{1i}(1, z_{nt}, H_{nt-1})$. By applying the results from Chesher (2007) and using the above results, we obtain our functional $\mathcal{F}^{-1}(F_{Y|X})$.

Taking expectations of equations (26) and (27) gives

$$\begin{aligned} E_t^o[Y_{i1nt}|\omega_{nt}] &= u_{1i}(1, z_{nt}, H_{nt-1}) - u_{it0}(z_{nt}) - u_{1i}(l_{nt}^*, z_{nt}, H_{nt-1}) \\ &\quad + E_t^o \left\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[u_{1i}\left(1, z_{1nt}^{(s)}, H_{1ns-1}^{(s)}\right) - u_{1i}\left(1, z_{0nt}^{(s)}, H_{0ns-1}^{(s)}\right)\right]\right\} \end{aligned} \tag{70}$$

$$E_t^o[Y_{i2nt}|\omega_{nt}] = -\frac{\partial u_{1i}(l_{nt}^*, z_{nt}, H_{nt-1})}{\partial h_t} - E_t^o \left\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{u_{i1}(1, z_{1nt}^{(s)}, H_{1ns-1}^{(s)})}{\partial h_t} \right]\right\}. \tag{71}$$

Note that $H_{1ns-1}^{(s)}$ is a function of h_{nt} while $H_{0ns-1}^{(s)}$ is not. Hence, taking the derivative of (70) with respect to h_{nt} gives

$$\frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h_{nt}} = -\frac{\partial u_{1i}(l_{nt}^*, z_{nt}, H_{nt-1})}{\partial h_t} + E_t^o \left\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{u_{1i}(1, z_{1nt}^{(s)}, H_{1ns-1}^{(s)})}{\partial h_t} \right]\right\}, \tag{72}$$

which implies that

$$E_t^o \left\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{u_{1i}(1, z_{1nt}^{(s)}, H_{1ns-1}^{(s)})}{h_t} \right]\right\} = \frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h_{nt}} + \frac{\partial u_{1i}(l_{nt}^*, z_{nt}, H_{nt-1})}{h_t}. \tag{73}$$

Substituting (73) into (27) gives

$$E_t^o[Y_{i2nt}|\omega_{nt}] = -2\frac{\partial u_{1i}(l_{nt}^*, z_{nt}, H_{nt-1})}{\partial h_t} - \frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h_{nt}}. \tag{74}$$

Rearranging, we get

$$\frac{\partial u_{1i}(l^*_{nt}, z_{nt}, H_{nt-1})}{\partial h_t} = -\frac{1}{2}\left\{\frac{\partial E^o_t[Y_{i1nt}|\omega_{nt}]}{\partial h_{nt}} + E^o_t[Y_{i2nt}|\omega_{nt}]\right\}. \tag{75}$$

Integrating up to h_{nt} gives

$$u_{1i}(l^*_{nt}, z_{nt}, H_{nt-1}) = u_{1i}(1, z_{nt}, H_{nt-1}) - \frac{1}{2}\int_0^{h_{nt}}\left\{\frac{\partial E^o_t[Y_{i1nt}|\omega_{nt}]}{\partial h} + E^o_t[Y_{i2nt}|\omega_{nt}]\right\} \mathrm{d}h \tag{76}$$

Let

$$C_{1it} = E^o_t\left\{\sum_{s=1}^{\rho(\omega_{nt})}\beta^s\left[u_{1i}\left(1, z^{(s)}_{1nt}, H^{(s)}_{1nt-1}\right) - u_{1i}\left(1, z^{(s)}_{0nt}, H^{(s)}_{0nt-1}\right)\right]\right\},$$

then substituting (76) into (70) and rearranging gives

$$u_{0it}(z_{nt}) = C_{1it} - E^o_t[Y_{i1nt}|\omega_{nt}] + \frac{1}{2}\int_0^{h_{nt}}\left\{E^o_t[Y_{i2nt}|\omega_{nt}] + \frac{\partial E^o_t[Y_{i1nt}|\omega_{nt}]}{\partial h}\right\} \mathrm{d}h \tag{77}$$

and we obtain the result in the Lemma. ■

D Estimation

D.1 Estimation of Consumption and Earnings Equations

In the first step, we use the Euler equation for consumption to form the moment condition:

$$E\left[\frac{\partial u_{2i}(c_{nt}, z_{nt}, \varepsilon_{2nt}, \theta_c)}{\partial c_{nt}} - \eta_n\lambda_t \;\middle|\; z_{nt}\right] = 0. \tag{78}$$

Here, we are assuming that the functional form of $u_2()$ is known up to a finite-dimensional parameter vector, θ_c. Recall that we assume that

$$u_{2i}(c_{nt}, z_{nt}, \varepsilon_{2nt}, \theta_c) = \exp(z'_{nt}B_4 + \varepsilon_{2nt})c^{\alpha}_{nt}/\alpha.$$

Let $\triangle$ denote the first-difference operator. Taking the logarithm of each side of this expression, differencing, and rearranging implies

(79) $$(1-\alpha)^{-1} \triangle \varepsilon_{2nt} = \triangle \ln(c_{nt}) - (1-\alpha)^{-1} \triangle z'_{nt} B_4 + \triangle (1-\alpha)^{-1} \ln(\lambda_t).$$

Let Θ_c denote the $(K+T-1)$-dimensional vector of parameters to be estimated, defined as

$$\Theta_c = \begin{pmatrix} (1-\alpha)^{-1} B_4 \\ \triangle (1-\alpha)^{-1} \ln(\lambda_2) \\ \vdots \\ \triangle (1-\alpha)^{-1} \ln(\lambda_T) \end{pmatrix}.$$

We also define $Y_n = (\triangle \ln(c_{n2}), \ldots, \triangle \ln(c_{nT}))'$ as a vector of endogenous variables and Z_n^c as the exogenous variables:

$$Z_n^c = \begin{bmatrix} \triangle z'_{n2} & D_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \triangle z'_{nT} & 0 & \ldots & D_T \end{bmatrix},$$

where D_t denotes a time dummy for $t \in \{2, \ldots, T\}$. The assumptions in Section 2 imply that the unobserved variable ε_{5nt} is independent of individual-specific characteristics. Therefore $E((1-\alpha)^{-1} \triangle \varepsilon_{2nt} \mid z_{nt}) = 0$. Using equation (79), one can obtain a set of orthogonality conditions,

$$E\left[\left(Y_n - Z_n^c \Theta_c\right) Z_n^c\right] = 0,$$

that can be exploited to estimate Θ_c using an optimal instrumental-variable estimation technique.

We use a traditional fixed-effect estimator to estimate $(1-\alpha)^{-1} \ln(\eta_n)$. Let T_1 be the number of time periods for which the marginal utility of consumption equation is estimated. Let

(80) $$(1-\alpha)^{-1} \ln(\eta_n) \equiv \sum_{t \in T_1} \left[\ln(c_{nt}) - (1-\alpha)^{-1} z'_{nt} B_4 + (1-\alpha)^{-1} \ln(\lambda_t)\right] / T_1.$$

The fixed-effects estimates of $(1-\alpha)^{-1} \ln(\eta_n)$ are obtained as the simple time averages of the estimated residuals of the consumption equation, which correspond to the sample counterparts of $(1-\alpha)^{-1} \ln(\eta_n)$ defined above. In order to form the sample counterpart of (80), we need an estimate of $\{(1-\alpha)^{-1}\ln(\lambda_t)\}_{t=1}^{T_1}$. From the estimate of Θ_c, however, we can only obtain estimates of $\{\triangle(1-\alpha)^{-1}\ln(\lambda_2)\}_{t=2}^{T_1}$. This requires us to make the additional assumption that $E_n[\eta_n \mid Z_{nt}] = 0$, where $E_n[.]$ is the expectation operator over individuals. This assumption enables us to obtain an estimate of $(1-\alpha)^{-1}\ln(\lambda_1)$ as the sample analogue of

$$(1-\alpha)^{-1}\ln(\lambda_1) = -E_n\left[\ln(c_{n1}) - (1-\alpha)^{-1}z'_{n1}B_4\right].$$

We now have estimates of $\{(1-\alpha)^{-1}\ln(\lambda_t)\}_{t=1}^{T_1}$ and $(1-\alpha)^{-1}\ln(\eta_n)$, enabling us to recover α in the third step of our estimation.

Next, we turn our attention to the estimation of the earnings equations. Let $d_{n\tau t} = I_{n\tau t} \times d_{nt}$. Since all the information set in equation (44) is public at period t, we have

(81)

$$E_t\{d_{n\tau t}d_{n\tau t-1}[\triangle S_{nt} - \triangle b_{0\tau t} - b_\tau \triangle HC_{nt} - \triangle z_{nt}^{p'}B_{\tau 5} - \beta\gamma_\tau \triangle d_{n\tau t+1}] \mid z_{nt}^p, H_{nt}, h_{nt}^*\} = 0,$$

where $\triangle HC_{nt} = (\triangle h_{nt}, \triangle h_{nt}^2, \triangle h_{nt-1}, \ldots, \triangle h_{nt-\rho})'$ and $b_\tau = (b_{\tau 1}, b_{\tau 2}, b_{\tau 31}, \ldots, b_{\tau 3\rho})$.

Let $\Theta_{e\tau}$ denote the $(2+K+\rho+T)$-dimensional vector of parameters to be estimated,

$$\Theta_{e\tau} = \begin{pmatrix} b_\tau \\ B_{\tau 5} \\ \beta\gamma_\tau \\ \triangle b_{0\tau 2} \\ \vdots \\ \triangle b_{0\tau T} \end{pmatrix}.$$

We also define $Y_{n\tau} = (d_{n\tau 2}d_{n\tau 1}\triangle S_{n2}, \ldots, d_{n\tau T}d_{n\tau T-1}\triangle S_{nT})'$ as a vector of endogenous variables and $X_{\tau n}$ as the exogenous variables,

$$X_{n\tau} = \begin{bmatrix} \triangle x'_{\tau 2} & D_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \triangle x'_{\tau T} & 0 & \ldots & D_T \end{bmatrix},$$

where $\triangle x'_{\tau nt} = d_{n\tau t}d_{n\tau t-1}(\triangle h_{nt}, \triangle h_{nt}^2, \triangle h_{nt-1}, \ldots, \triangle h_{nt-\rho}, \triangle z_{nt}^{p'}, \triangle d_{n\tau t+1})$. Letting

Z_n be the matrix of conditioning variables,

$$Z_n = \begin{bmatrix} z_{n2}^{p\prime} & H_{n2} & h_{n2} \\ \vdots & \vdots & \vdots \\ z_{nT}^{p\prime} & H_{nT} & h_{n2T} \end{bmatrix},$$

and using equation (81), one can obtain a set of orthogonality conditions:

$$E\left[\left(Y_{n\tau} - X_{n\tau}\Theta_{e\tau}\right) Z_n\right] = 0,$$

which can be exploited to estimate $\Theta_{e\tau}$ using an optimal instrumental-variable technique. The aggregate effect and fixed effect in the earnings equation are estimated in a similar way to those in the consumption equation.

D.2 Estimation of the Final Stage

Note that from the second step, we have estimates of $b_{\tau 1}$, $b_{\tau 2}$, β,γ_τ, and all the other parameters of the production function. In addition, from the first step, we have an estimate of ϕ_{nt},

$$\phi_{nt} = (1-\alpha)^{-1}\ln(\eta_n\lambda_t).$$

The third step yields estimates of p_{nt}, $p_{1nt}^{(s)}$, $\widetilde{p}_{n\tau t+1}$, $\frac{\partial p_{1nt}^{(s)}}{\partial h_{nt}}$, and $\frac{\partial \widetilde{p}_{n\tau t+1}}{\partial h_{nt}}$. We can form the moment conditions:

$$\begin{aligned} m_{1nt}\left(\Theta u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right) \quad = \quad & \sigma \ln\left[p_{nt}^{(N)} / \left(1 - p_{nt}^{(N)}\right)\right] - B_{0t} - z_{nt}' B_1 + z_{nt}' h_{nt} B_2 \\ & + \theta_0\left(1 - l_{nt}^2\right) + \sum_{s=1}^{\rho} \theta_s h_{nt}(l_{nt-s} + \beta^s) \\ & - \sigma \sum_{s=1}^{\rho} \beta^s \ln\left(\frac{1 - p_{1nt}^{(s)(N)}}{1 - p_{0nt}^{(s)(N)}}\right) \\ & - \exp\left((1-\alpha)\phi_{nt}^{(N)}\right) \sum_{\tau=1}^{M} I_{n\tau t}\left[y_{\tau t}\left(h_{nt}, H_{nt-1}, z_{nt}^p, \theta_e^{(N)}\right)\right. \\ (82) \qquad & \left. - \gamma_\tau^{(N)} + \beta \gamma_\tau^{(N)} \widetilde{p}_{n\tau t+1}^{(N)}\right] \end{aligned}$$

and

$$m_{2nt}\left(\Theta u,\Theta_c^{(N)},\Theta_e^{(N)},\psi^{(N)}\right) = d_{nt}\Bigg\{\sigma\sum_{s=1}^{\rho}\beta^s\left(1-p_{1nt}^{(s)(N)}\right)^{-1}\frac{\partial p_{1nt}^{(s)(N)}}{\partial h_{nt}}$$
$$-z'_{nt}B_2-2\theta_0 l_{nt}-\sum_{s=1}^{\rho}\theta_s(l_{nt-s}+\beta^s)$$
$$+\exp\left((1-\alpha)\phi_{nt}^{(N)}\right)\sum_{\tau=1}^{M}d_{n\tau t}\Bigg[b_{\tau 1}^{(N)}+2b_{\tau 2}^{(N)}h_{nt}$$
$$+\ \beta\gamma_\tau^{(N)}\frac{\partial\widetilde{p}_{n\tau t+1}^{(N)}}{\partial h_{nt}}\Bigg]\Bigg\}, \quad (83)$$

where $\psi^{(N)} = \left(p_{nt}^{(N)},p_{0nt}^{(s)(N)},p_{1nt}^{(s)(N)},\widetilde{p}_{n\tau t+1}^{(N)}\right)$ are the nonparametric second-step estimates and $\Theta_u = (\sigma,\alpha,\beta,B_{01},\ldots,B_{0T},B_1,B_2,\theta_0,\ldots,\theta_\rho)$ are the structural parameters left to be estimated.

There are now two sources of errors in evaluating the sample counterparts of (82) and (83). The first is the forecast errors from replacing the expectations of future variables with their realizations. The second is the approximation error that arises from replacing the true values of the conditional choice probabilities, conditional expectation, and time-invariant individual-specific effects with their estimates. Let us define the 2×1 vector $m_{3nt}\left(\Theta_u,\Theta_c^{(N)},\Theta_e^{(N)},\psi^{(N)}\right)\equiv\left[m_{1nt}\left(\Theta_u,\Theta_c^{(N)},\Theta_e^{(N)},\psi^{(N)}\right),m_{2nt}\left(\Theta_u,\Theta_c^{(N)},\Theta_e^{(N)},\psi^{(N)}\right)\right]'$ and let T_3 denote the set of periods for which the hours and participation equations are valid. Define the vector

$$m_{3n}^{(N)}\left(\Theta_u,\Theta_c^{(N)},\Theta_e^{(N)},\psi^{(N)}\right)\equiv\left(m_{3n1}\left(\Theta_u,\Theta_c^{(N)},\Theta_e^{(N)},\psi^{(N)}\right)',\ldots,m_{3nT_3}\left(\Theta_u,\Theta_c^{(N)},\Theta_e^{(N)},\psi^{(N)}\right)'\right)'$$

as the vector of the idiosyncratic errors for a given individual over time. Define $\Omega_{nt}^{(N)}\equiv E_t\left[m_{3nt}\left(\Theta_u,\Theta_c^{(N)},\Theta_e^{(N)},\psi^{(N)}\right)m_{3nt}\left(\Theta_u,\Theta_c^{(N)},\Theta_e^{(N)},\psi^{(N)}\right)'\right]$. The off-diagonal elements of $\Omega_{nt}^{(N)}$ are zero because $E_t\left[m_{3nt}\left(\Theta_u,\Theta_c^{(N)},\Theta_e^{(N)},\psi^{(N)}\right)m_{3nr}\left(\Theta_u,\Theta_c^{(N)},\Theta_e^{(N)},\psi^{(N)}\right)'\right]=0$ for $r\neq t$, $r<t$. The 2×2 conditional heteroscedasticity matrix $\Omega_{nt}^{(N)}$ associated with the individual-specific errors, $m_{3nt}\left(\Theta_u,\Theta_c^{(N)},\Theta_e^{(N)},\psi^{(N)}\right)$, is evaluated using a nonparametric esti-

mator based on the estimated moments, $m_{3nt}\left(\Theta_{1u}^{(N)}, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)$, derived from an initial consistent estimate of $\Theta_{1u}^{(N)}$. The optimal instrumental-variables estimator for $\Theta_u^{(N)}$ is

(84)

$$\Theta_u^{(N)} \equiv \arg\min_{\Theta u} \frac{\sum_{n=1}^{N} m_{3n}^{(N)}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)\left(\Omega_n^{(N)}\right)^{-1} m_{3n}^{(N)}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)}{N}.$$

D.3 Asymptotic Properties

It is well known in the econometric literature that under certain regularity conditions, pre-estimation does not have any impact on the consistency of the parameters in the subsequent steps of a multistage estimation (Newey, 1984; Newey and McFadden, 1994; Newey, 1994). The asymptotic variance, however, is affected by the pre-estimation. In order to conduct inference in this type of estimation, one has to correct the asymptotic variance for the pre-estimation. The method used for correcting the variance in the final step of estimation depends on whether the pre-estimation parameters are of finite or infinite dimension. Unfortunately, our estimation strategy combines both finite- and infinite-dimensional parameters. Combining results from two sources (Newey, 1984; Newey and McFadden, 1994), however, allows us to derive the corrected asymptotic variance for our estimator.

Following Newey (1984), we can write the sequential-moments conditions for the first- and third-step estimation as a set of joint moment conditions:

$$m_n(\Theta_u, \Theta_c, \Theta_e, \psi) = \begin{bmatrix} (Y_n - Z_n\Theta_c)\, Z_n^c \\ (Y_{n1} - X_{1n}\Theta_{e1})\, Z_n \\ (Y_{n2} - X_{Mn}\Theta_{eM}) Z_n \\ m_{3n}(\Theta_u, \Theta_c, \Theta_e, \psi) \end{bmatrix},$$

where $(Y_n - Z_n\Theta_c)\, Z_n^c$ is the orthogonality condition from the estimation of the consumption equation, $(Y_{n\tau} - X_{n\tau}\Theta_{e\tau})Z_n$ is the orthogonality condition from the estimation of the earnings equation, and $m_{3n}(\Theta_u, \Theta_c, \Theta_e, \psi)$ is the moment conditions from the third-step estimation. Let $\Theta = (\Theta_u, \Theta_c, \Theta_e)'$, with the true value denoted by Θ_0. Note that each element of ψ is a conditional expectation. Redefine each element as $\psi^j(z^j) = f_{z^j}(z^j)E\left[\widetilde{d}_n^j \mid z^j\right]$, where $\widetilde{d}_{nt}^j = [1, d_{nt}]'$ for the estimation of p_{nt}, $\widetilde{d}_{nt}^j = [d_{knt}^{(r)}, d_{knt}^{(r)} d_{nt}]'$ for the estimation of $p_{knt}^{(r)}$, and $\widetilde{d}_{nt}^j = [d_{n\tau t}, d_{n\tau t} d_{n\tau t+1}]'$ for the estimation of $\widetilde{p}_{n\tau t+1}$. Therefore, $\psi^{j(N)}(z^j) = \frac{1}{N}\sum_{n=1}^{N} \widetilde{d}_n^j J_{\delta_N}(z^j - z_n^j)$. The conditions below ensure that $\psi^{(N)}$ is close enough to ψ_0 for N large enough, in particular that $\sqrt{N}\left\|\psi^{(N)} - \psi_0\right\|^2$ converges to zero.

A3: *There is a version of $\psi_0(z)$ that is continuously differentiable of order κ, greater than the dimension of z and $\psi_{10}(z) = f_z(z)$ is bounded away from 0.*

A4: $\int J(u)\,\mathrm{d}u = 1$ *and for all* $j < \kappa$, $\int J(u)\left(\bigotimes_{s=1}^{j} u\right)\mathrm{d}u = 0$.

A5: *The bandwidth, δ_N, satisfies* $N\delta_N^{2\dim(z)}/(\ln(N))^2 \to \infty$ *and* $N\delta_N^{2\kappa} \to 0$.

A6: *There exists a* $\Psi(\omega)$, $\epsilon > 0$, *such that*

$$\|\nabla_\Theta m_n(\omega.\Theta, \psi) - \nabla_\Theta m_n(\omega.\Theta_0, \psi_0)\| \leq \Psi(\omega)\left[\|\Theta - \Theta_0\|^\epsilon + \|\psi - \psi_0\|^\epsilon\right]$$

and $E[\Psi(\omega)] < \infty$.

A7: $\Theta^{(N)} \to \Theta_0$ *with* Θ_0 *in the interior of its parameter space.*

A8: (Boundedness)

(i) Each element of $m_n(\Theta, \psi)$ is bounded almost surely: $E[\|m_n(\Theta, \psi)\|^2] < \infty$;

(ii) $E[Z'_{n'}Z_n] < \infty$, $E[X'_{\tau n}Z_n] < \infty$, $E[\exp((1-\alpha)\phi_{nt})] < \infty, E[z_{nt}] < \infty, E[y_{\tau t}(h_{nt}, H_{nt-1}, z_{nt}^p, \theta_e)] < \infty, \gamma_\tau < \infty$, $E[\nabla_{h_{nt}}\widetilde{p}_{n\tau t+1}] < \infty$, $E[X_{n\tau}] < \infty$ for $\tau = 1, 2$;

(iii) $p_{nt}, p_{knt}^{(r)}, \widetilde{p}_{n\tau t+1}, \in (0,1)$, for $k \in \{0,1\}$, $r = 1, \ldots, \rho$, and $\tau = 1, 2$;

(iv) $E[\nabla_h f_{z^j}(z^j)] < \infty$ and $E[\nabla_h E[\widetilde{d}_n^j \mid z^j]] < \infty$;

Theorem 1 *Under A1–A8 and $\Phi(\omega)$, defined below,*

$$\sqrt{N}\left(\Theta^{(N)} - \Theta_0\right) \Rightarrow N(0, \Sigma(\Theta_0)),$$

where

$$
\begin{aligned}
\Sigma(\Theta_0) &= E\left[\nabla_\Theta m_n(\omega)\Omega_n^{-1}\nabla_\Theta m_n(\omega)'\right]^{-1} \\
&\quad \times E\left[\nabla_\Theta m_n(\omega)\Omega_n^{-1}\left\{m_n(\omega)+\Phi(\omega)\right\}\left\{m_n(\omega)+\Phi(\omega)\right\}'\Omega_n^{-1}\nabla_\Theta m_n(\omega)'\right] \\
&\quad \times E\left[\nabla_\Theta m_n(\omega)\Omega_n^{-1}\nabla_\Theta m_n(\omega)'\right]^{-1}.
\end{aligned}
$$

Assumptions A3–A8 are standard in the semiparametric literature, see Newey and McFadden (1994) for details. One can now use Theorem 1 to calculate the standard for all the parameters in our estimation.

The proof of Theorem 1 will follow from checking the conditions for Theorem 8.12 in Newey and McFadden (1994). We Assume A1–A7 and add the following additional assumption.

Proof of Theorem 1. We first check the various boundedness requirements of Theorem 8.12 in Newey and McFadden (1994). By assumption A8(i), we have that $E[\|m_n(\Theta,\psi)\|^2]<\infty$. It obvious by inspection that $m_n(\Theta,\psi)$ is continuously differentiable in Θ and by A8(ii–iv) that $E[\nabla_\Theta m_n(\Theta,\psi)]<\infty$. Additionally, $\nabla_{\psi\psi}m_n(\Theta_0,\psi_0)$ is also bounded: $E[\|\nabla_{\psi\psi}m_n(\Theta_0,\psi_0)\|]<\infty$.

Second, consider a pointwise Taylor expansion for the j^{th} element of m_n,

$$
\begin{aligned}
m^j(\omega,\psi) &= m^j(\omega,\psi_0)+\nabla_\psi m^j(\omega,\psi_0)(\psi(z)-\psi_0(z)) \\
&\quad + (\psi(z)-\psi_0(z))'\nabla_{\psi\psi}m^j(\omega,\psi_0)(\psi(z)-\psi_0(z))+o(\|\psi(z)-\psi_0(z)\|^2),
\end{aligned}
$$

where the norm over ψ is the sup-norm. Next, note that

$$
\begin{aligned}
&\left|m^j(\omega,\psi)-m^j(\omega,\psi_0)\nabla_\psi m^j(\omega,\psi_0)(\psi(z)-\psi_0(z))\right| \\
&\qquad \le \left\|(\psi(z)-\psi_0(z))'\nabla_{\psi\psi}m^j(\omega,\psi_0)(\psi(z)-\psi_0(z))\right\| \\
&\qquad\quad + o(\|\psi(z)-\psi_0(z)\|^2) \\
&\qquad \le \|\psi-\psi_0\|^2\left\|\nabla_{\psi\psi}m^j(\omega,\psi_0)\right\|+o(\|\psi-\psi_0\|^2),
\end{aligned}
$$

using the triangle inequality and the Cauchy-Schwartz inequality. Therefore, for $\|\psi-\psi_0\|$ small enough,

$$
\left|m^j(\omega,\psi)-m^j(\omega,\psi_0)-\nabla_\psi m^j(\omega,\psi_0)(\psi(z)-\psi_0(z))\right| \le \|\psi-\psi_0\|^2\left\|\nabla_{\psi\psi}m^j(\omega,\psi_0)\right\|.
$$

So that

$$
\begin{aligned}
\|m(\omega,\psi)-m(\omega,\psi_0)-\nabla_\psi m(\omega,\psi_0)(\psi(z)-\psi_0(z))\| &\leq \|\psi-\psi_0\|^2 \|\nabla_{\psi\psi} m(\omega,\psi_0)\|
\end{aligned}
$$

$$
\|m(\omega,\psi)-m(\omega,\psi_0)-\nabla_\psi m(\omega,\psi_0)(\psi(z)-\psi_0(z))\| \leq \|\psi-\psi_0\|^2 \|\nabla_{\psi\psi} m(\omega,\psi_0)\|
$$

Hence $\Gamma(\omega,\psi-\psi_0)=\nabla_\psi m(\omega,\psi_0)(\psi(z)-\psi_0(z))$ and $\Psi(\omega)=\|\nabla_{\psi\psi} m(\omega,\psi_0)\|$. It follows that both $\Gamma(\omega,\psi-\psi_0)$ and $\Psi(\omega)$ are bounded from the boundedness conditions established above.

Next we establish the form of the influence function. Note that we have

$$
\begin{aligned}
\int \Gamma(\omega,\psi)F_0(\mathrm{d}\omega) &= \int f_z(z)E[\nabla_\psi m(\omega,\psi_0)\mid z]\psi(z)\,\mathrm{d}z \\
&= \int \upsilon(z)\psi(z),
\end{aligned}
$$

where $\upsilon(z)=f_z(z)E[\nabla_\psi m(\omega,\psi_0)\mid z]$. So, by the arguments on page 2208 of Newey and McFadden (1994), we have the influence function for $m(\omega,\psi^{(N)})$:

$$
\begin{aligned}
\Phi(\omega) &= \upsilon(z)-E\left[\upsilon(z)\widetilde{d}\right] \\
&= f_z(z)E\left[\nabla_\psi m(\omega,\psi_0)\mid z\right]-E\left[f_z(z)E[\nabla_\psi m(\omega,\psi_0)\mid z]\widetilde{d}\right].
\end{aligned}
$$

Again by the boundedness of $\nabla_\psi m(\omega,\psi_0)$, it follows that $\int\|\upsilon(z)\|\,\mathrm{d}z<\infty$. Finally Assumption A7 guarantees that the Jacobian term converges. ■

References

[1] Aguirregabiria, V. (2005): "Nonparametric Identification of Behavioral Responses to Counterfactual Policy Interventions in Dynamic Discrete Decision Processes," *Economics Letters*, 87, 393–398.

[2] Aguirregabiria, V. (2005b): *Notes on Industry Dynamics*, Department of Economics, Boston University.

[3] Aguirregabiria, V., and P. Mira (2002): "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70, 1519–1543.

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

[5] Albanesi, S., and C. Olivetti (2005): *Home Production, Market Production and the Gender Wage Gap: Incentives and Expectations*, mimeo, Boston University.

[6] Altonji, J. (2005): " Employer Learning, Statistical Discrimination and Occupational Attainment", American Economic Review, 95(2) pp. 112-117.

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

[8] Altonji and Paxton (1992) "Labor Supply, Hours Constraints and Job Mobility," Journal of Human Resources 27(2):256-78.

[9] Altonji, J., and C. Pierret (2001): "Employer Learning and Statistical Discrimination," *The Quarterly Journal of Economics*, 116, 313–350.

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

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

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

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

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

[15] Bajari, P., L. Benkard, and J. Levin (forthcoming): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*.

[16] Bajari, P., and H. Hong (2006): *Semiparametric Estimation Game of a Dynamic Incomplete Information*, NBER Technical Working Paper No. T0320.

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

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

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

[20] Bisin, A., and P. Gottardi (2006): "Efficient Competitive Equilibria With Adverse Selection," *Journal of Political Economy*, 114, 485–515.

[21] Blau, F., and L. Kahn (1997): "Swimming Upstream: Trends in the Gender Wage Differential in the 1980s," *Journal of Labor Economics*, 15, 1–42.

[22] Blundell, R., and T. MaCurdy (2007): "Labor Supply: A Review of Alternative Approaches," in: O. Ashenfelter and D. Card ed., *Handbook of Labor Economics*, vol. 3, Elsevier, Amsterdam, pp. 1559–1695.

[23] Blundell, R., L. Pistaferri and I. Preston (2004): *Imputing Consumption in the PSID Using Food Demand Estimates from the CEX*, IFS Working Paper, W04/27.

[24] Bowlus, A., and Z. Eckstein (2002): "Discrimination and Skill Differences in an Equilibrium Search Model," *International Economic Review*, 43, 1309–1345.

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

[26] Chesher, A. (2003): "Identification in Nonseparable. Models," *Econometrica*, 71, 1405–1441.

[27] Chesher, A. (2005): "Nonparametric Identification under Discrete Variation," *Econometrica*, 73, 1525–1550.

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

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

[30] Dey, J. G., and C. Hill (2007): *Behind the Pay Gap*, American Association of University Women Educational Foundation, Washington, DC.

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

[32] Eckstein, Z., and E. Nagypal (2004): "The Evolution of U.S. Earnings Inequality: 1961–2002," *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(2), 10–29.

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

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

[35] Fernández, R. (2007): *Culture as Learning: The Evolution of Female Labor Force Participation over a Century,* Department of Economics, New York University.

[36] Fogli, A. and L. Veldkamp (2007): *Nature or Nurture? Learning and Female Labor Force Participation*," Federal Reserve Bank of Minneapolis Staff Report 386.

[37] Fudenberg, D., and J. Tirole (1996): *Game Theory*. Cambridge: MIT Press.

[38] Gale D. (1992): "A Walrasian Theory of Markets With Adverse Selection" *The Review of Economic Studies*, 59, 229–255.

[39] Gayle, G.-L., and R. A. Miller (2004): *Life-Cycle Fertility Behavior and Human Capital Accumulation*, mimeo, Tepper School of Business, Carnegie Mellon University.

[40] Haavelmo, T. M. (1944): "The Probability Approach in Econometrics," *Econometrica*, 12, Supplement, iii–vi 1–115.

[41] Hagedorn, M. and I. Manovskii (2005): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," Department of Economics, University of Pennsylvania.

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

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

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

[45] Hotz, V. J., R. A. Miller, S. Sanders, and J. Smith (1994): "A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies*, 61, 256–289.

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

[47] Jofre-Bonet, M., and M. Pesendorfer (2003): "Estimation of a Dynamic Auction," *Econometrica*, 71, 1443-1489.

[48] Jovanovic, B. (1989): "Observable Implications of Models With Multiple Equilibria," *Econometrica*, 57, 1431–1437.

[49] Lewis, G. (1996): "Gender Integration of Occupations in Federal Civil Service: Extent and Effects on Male–Female Earnings," *Industrial and Labor Relations Review*, 49, 472–483.

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

[51] Light, A., and M. Ureta (1992): "Panel Estimates of Male and Female Job Turnover Behavior: Can Female Nonquitters Be Identified?" *Journal of Labor Economics*, 10, 156–181.

[52] Light and Ureta (1992),: "Early Career Work Experience and Gender Wage Differential" *Journal of Labor Economics 13(1), pp. 121-154*

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

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

[55] Magnac, T., and D. Thesmar (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 70, 801–816.

[56] Maskin, E., and J. Tirole (2001): "Markov Perfect Equilibrium I. Observable Actions ", Journal of Economic Theory, 100(2), pp. 191-219.

[57] Miller, R. A., and H. Sieg (1997): "A Macroeconomic Comparison of Household Behavior Between Countries," *Journal of Business and Economic Statistics*, 15, 237–254.

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

[59] 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.

[60] Moro, A., and P. Norman (2004): "A General Equilibrium Model of Statistical Discrimination," *Journal of Economic Theory*, 114, 1–30.

[61] Newey, W. (1984): "A Method of Moments Interpretation of Sequential Estimators," *Economic Letters*, 14, 201–206.

[62] Newey, W. (1994): "The Asymptotic Variance of Semiparametric Estimators," *Econometrica*, 16, 1–32.

[63] Newey, W., and D. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," in: R. F. Engle and D. L. McFadden eds. *Handbook of Econometrics.* North-Holland, Amsterdam, pp. 2111–2245.

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

[65] Pakes, A., M. Ostrovsky, and S. Berry (forthcoming: "Simple Estimators for the Parameters of Discrete Dynamic Games", The Rand Journal of Economics.

[66] Pesendorfer, M., and P. Schmidt-Dengler (2003): *Identification and Estimation of Dynamic Games*, mimeo, London School of Economics.

[67] Pesendorfer, M., and P. Schmidt-Dengler (2006): *Asymptotic Least Squares Estimators for Dynamic Games*, NBER Working Paper w9726.

[68] Phelps, E. (1972): "The statistical theory of racism and sexism," *American Economic Review*, 62, 659–661.

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

[70] Prescott, E., and R. Townsend (1984): "Pareto Optima and Competitive Equilibria With Adverse Selection and Moral Hazard" *Econometrica*, 52, 21–46.

[71] Riley J. (1979): "Informational Equilibrium" *Econometrica*, 47, 331–359.

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

[73] Tamer, E. (2003): "Incomplete Simultaneous Discrete Response Model With Multiple Equilibria," *Review of Economic Studies.*, 70, 147–167.

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

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

[76] Weinberger, C. and P. Kuhn (2006): *The Narrowing of the U.S. Gender Earnings Gap, 1959-1999: A Cohort-Based Analysis*, Department of Economics, University of California at Santa Barbara.

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

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

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

Table 2: Summary of Demographic and Wealth Variables

	Household Income	Food Consumption	Family Size	Age	Number of Kids ≤ 5 years old	> 5 and < 17 years old	Marital Status
1968	46.7 (27.7)		4.0 (1.9)	37.8 (10.7)	0.56 (0.82)	0.94 (1.3)	0.85 (0.35)
1969	49.9 (31.1)	7.7 (3.7)	4.0 (1.9)	38.5 (10.9)	0.53 (0.83)	0.93 (1.3)	0.86 (0.35)
1970	50.3 (30.0)	7.7 (3.6)	3.8 (1.8)	38.6 (11.3)	0.49 (0.79)	0.87 (1.3)	0.85 (0.36)
1971	51.0 (31.4)	7.5 (3.5)	3.7 (1.8)	39.0 (11.6)	0.44 (0.76)	0.85 (1.2)	0.83 (0.37)
1972	51.4 (31.5)	7.4 (3.5)	3.7 (1.8)	39.3 (11.8)	0.42 (0.72)	0.80 (1.2)	0.82 (0.38)
1973	53.3 (35.2)	7.4 (3.4)	3.6 (1.8)	39.5 (12.2)	0.39 (0.69)	0.77 (1.1)	0.82 (0.38)
1974	54.3 (35.2)	7.3 (3.4)	3.4 (1.8)	39.8 (12.5)	0.37 (0.68)	0.71 (1.1)	0.82 (0.39)
1975	52.9 (34.6)	6.9 (3.2)	3.3 (1.7)	39.8 (12.6)	0.35 (0.67)	0.66 (1.0)	0.81 (0.39)
1976	53.4 (35.6)	6.8 (3.2)	3.3 (1.7)	39.7 (12.6)	0.35 (0.68)	0.62 (1.0)	0.80 (0.39)
1977	52.1 (35.6)	6.7 (3.3)	3.2 (1.6)	39.7 (12.6)	0.34 (0.68)	0.60 (0.96)	0.79 (0.40)
1978	52.1 (35.6)	6.5 (3.5)	3.2 (1.6)	38.9 (12.7)	0.41 (0.72)	0.51 (0.87)	0.77 (0.42)
1979	55.5 (55.2)	6.7 (3.3)	3.1 (1.5)	39.8 (12.5)	0.34 (0.66)	0.53 (0.88)	0.77 (0.42)
1980	55.1 (39.1)	6.6 (3.3)	3.1 (1.5)	39.9 (12.5)	0.35 (0.69)	0.50 (0.84)	0.78 (0.42)
1981	56.2 (68.0)	6.4 (3.1)	3.1 (1.4)	38.8 (11.9)	0.39 (0.69)	0.50 (0.82)	0.80 (0.39)
1982	54.1 (40.3)	6.3 (3.1)	3.1 (1.4)	38.9 (11.8)	0.38 (0.69)	0.50 (0.82)	0.80 (0.40)
1983	53.1 (39.4)	6.3 (3.1)	3.1 (1.4)	39.0 (11.8)	0.38 (0.67)	0.51 (0.84)	0.80 (0.40)
1984	54.8 (43.2)	6.3 (3.1)	3.1 (1.4)	39.1 (11.7)	0.38 (0.70)	0.51 (0.84)	0.80 (0.40)
1985	57.9 (51.3)	6.5 (3.8)	3.1 (1.4)	39.6 (11.5)	0.37 (0.68)	0.53 (0.85)	0.80 (0.40)
1986	59.2 (48.8)	6.4 (3.2)	3.1 (1.4)	40.2 (11.2)	0.37 (0.69)	0.56 (0.87)	0.81 (0.40)
1987	61.9 (51.4)	6.5 (3.1)	3.1 (1.3)	40.6 (10.8)	0.36 (0.67)	0.57 (0.87)	0.81 (0.40)
1988	64.1 (63.6)	6.6 (3.0)	3.1 (1.3)	41.4 (10.6)	0.34 (0.66)	0.58 (0.87)	0.81 (0.40)
1989	65.9 (69.7)	6.5 (2.9)	3.1 (1.3)	42.1 (10.3)	0.33 (0.63)	0.59 (0.87)	0.81 (0.39)
1990	66.3 (63.7)	6.6 (3.2)	3.1 (1.3)	42.7 (10.1)	0.30 (0.62)	0.61 (0.89)	0.81 (0.39)
1991	61.4 (52.3)	6.4 (3.1)	3.0 (1.4)	43.7 (10.2)	0.29 (0.60)	0.58 (0.87)	0.77 (0.43)
1992	65.6 (62.1)	6.7 (3.8)	3.1 (1.3)	44.0 (9.6)	0.29 (0.61)	0.62 (0.89)	0.82 (0.39)

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

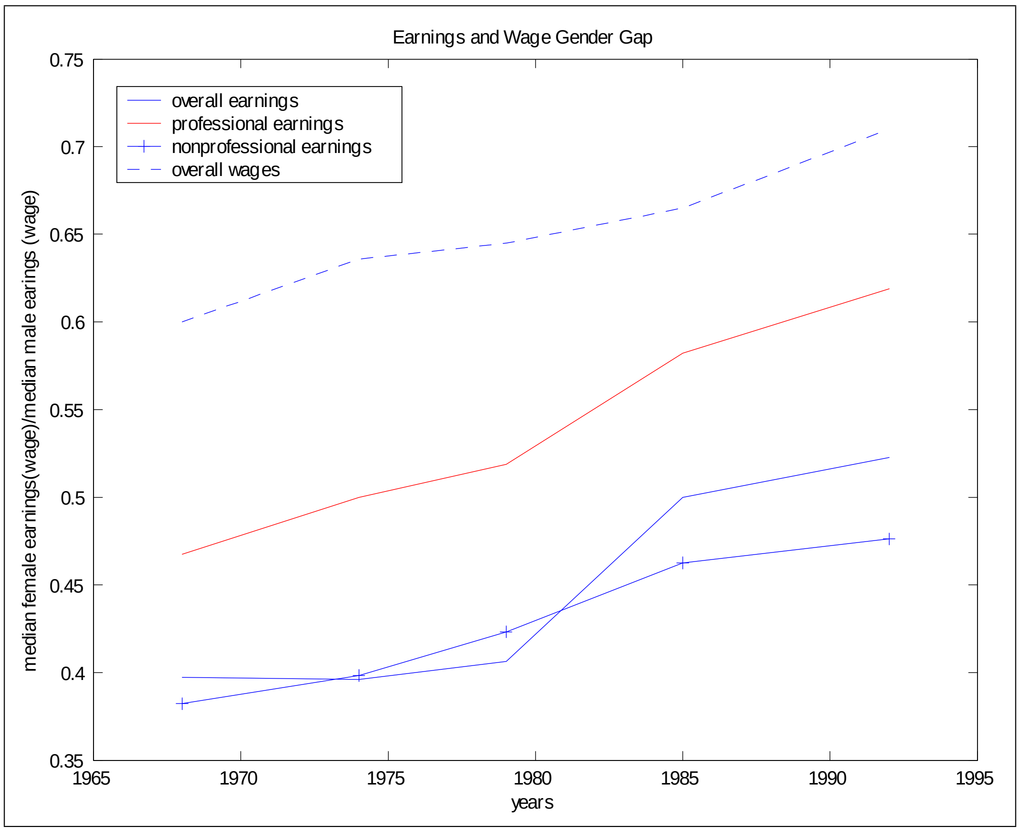

Figure 1: Earnings and Wage Gaps

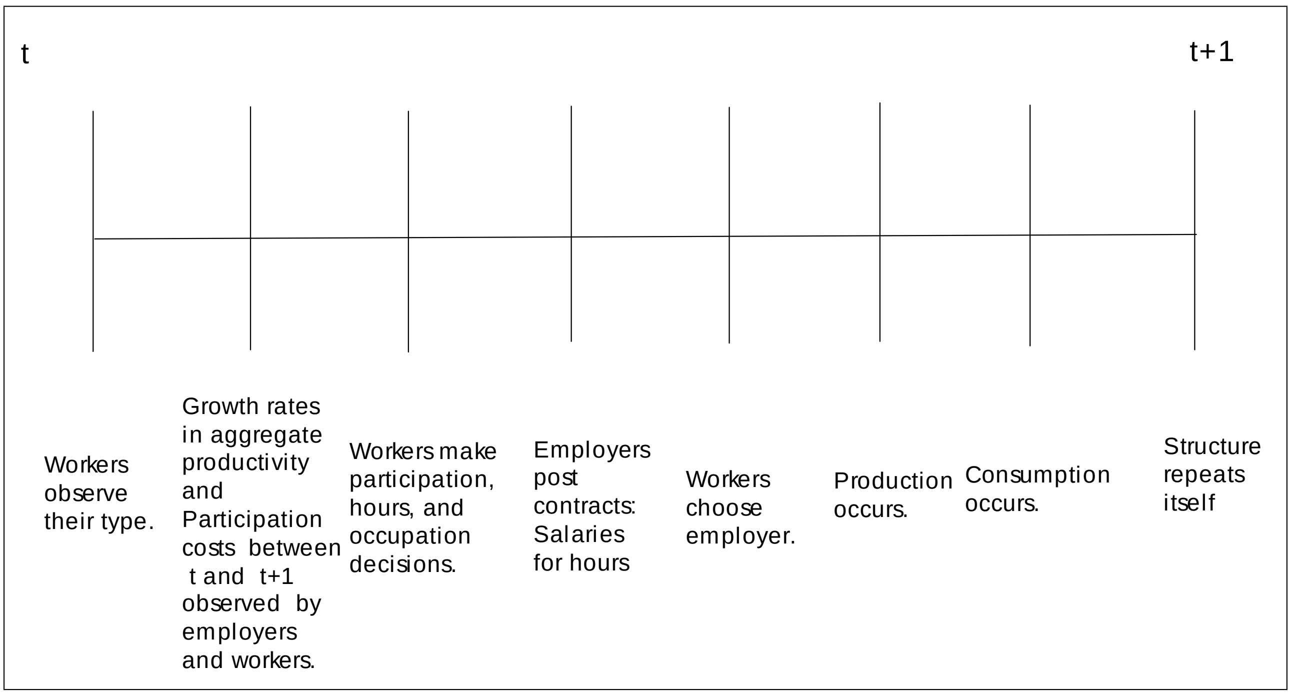

Figure 2: Timeline

Table 3: Consumption Equation

$\ln(c_{nt}) = 1/(1-\alpha)[z'_{nt}B_4 - \ln(\eta_n \lambda_t) + \epsilon_{2nt}]$

Variable	Parameter	Estimate
risk aversion	α	0.636 (2.0E–04)
Socioeconomic variables		
FAM_{nt}	$(1-\alpha)^{-1}B_{41}$	0.0253 (3.4E–04)
$YKID_{nt}$	$(1-\alpha)^{-1}B_{42}$	0.0014 (0.0015)
$OKID_{nt}$	$(1-\alpha)^{-1}B_{43}$	-0.0013 (0.0014)
AGE^2_{nt}	$(1-\alpha)^{-1}B_{24}$	-1.20E–04 (4.03E–05)
Region Dummies		
NE_{nt}	$(1-\alpha)^{-1}B_{45}$	-0.0076 (0.0032)
SO_{nt}	$(1-\alpha)^{-1}B_{46}$	-0.0041 (0.0022)
W_{nt}	$(1-\alpha)^{-1}B_{26}$	-0.0023 (0.0025)

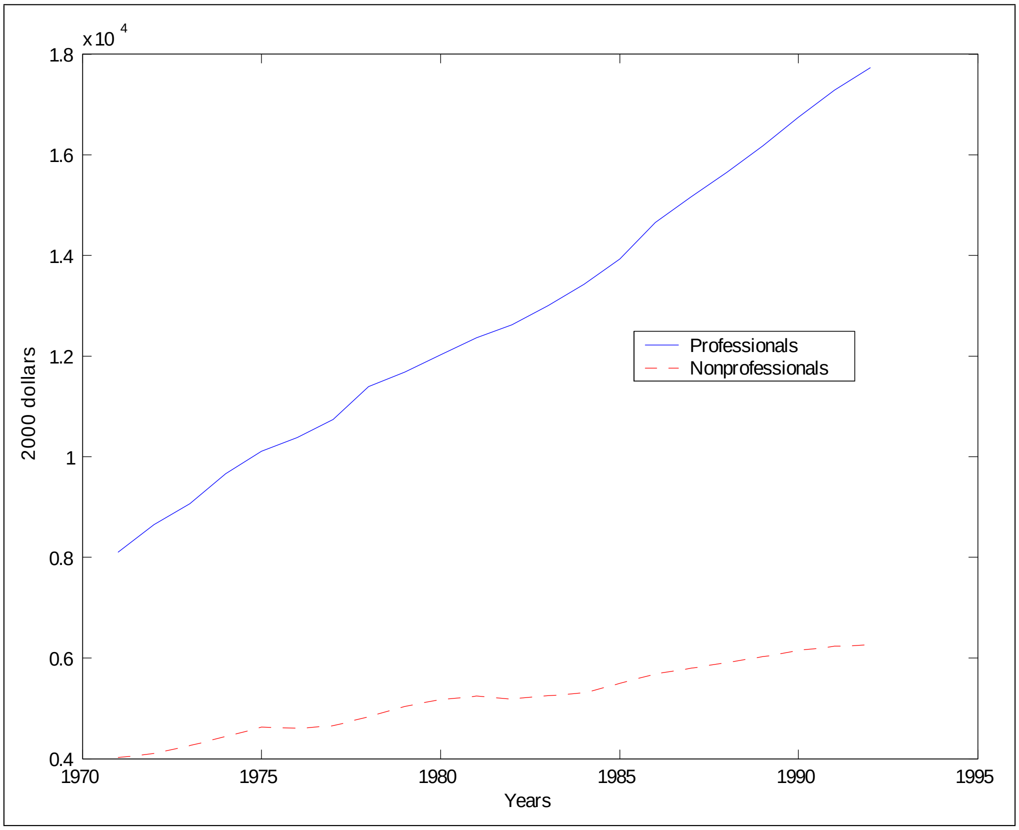

Figure 3: Estimated Aggregate Productivity

Table 4: Earnings Equation

Variable	Professional	Nonprofessional
Hours and Lagged hours		
h_{nt}	183,392 (2,560)	100,688 (967)
h_{nt}^2	−251,162 (4,908)	−88,891 (2,152)
h_{nt-1}	14,252 (808)	12,394 (340)
h_{nt-2}	6,086 (730)	3,969 (330)
Age and Education		
AGE_{nt}^2	−36 (1.5)	−13 (0.7)
$AGE_{nt} \times EDU_{nt}$	−23 (14)	25 (6.6)
Hiring cost	3,032 (171)	875 (70)

Table 5: Fixed Cost to Labor Participation

Variable	Estimate
Socioeconomic variables	
FAM_{nt}	-0.0625
	(0.001)
$YKID_{nt}$	-0.713
	(0.0001)
$YKID_{nt} \times male\ dummy_{nt}$	0.863
	(0.0001)
$OKID_{nt}$	-0.413
	(0.0001)
$OKID_{nt} \times male\ dummy_{nt}$	0.477
	(0.0001)
AGE_{nt}	0.163
	(0.01)
AGE_{nt}^2	-0.003
	(0.008)
$EDUC_{nt}$	0.08
	(0.0004)
$EDUC_{nt} \times male\ dummy_{nt}$	-0.03
	(0.04)
MS_{nt}	0.205
	(0.006)
$SP.EDUC_{nt} \times MS_{nt}$	-0.088
	(0.005)
$SP.EDUC_{nt} \times MS_{nt} \times male\ dummy_{nt}$	0.145
	(0.003)

Table 6: Utility of Leisure/Home Production

Variable	Estimate
l_{nt}	-4.4558 (0.004)
$FAM_{nt} \times l_{nt}$	0.082 (0.01)
$YKID_{nt} \times l_{nt}$	-0.1033 (0.001)
$YKID_{nt} \times l_{nt} \times male\ dummy_{nt}$	0.933 (0.001)
$OKID_{nt} \times l_{nt}$	-0.141 (0.001)
$OKID_{nt} \times l_{nt} \times male\ dummy_{nt}$	0.098 (0.001)
$AGE_{nt} \times l_{nt}$	-0.045 (0.19)
$AGE_{nt}^2 \times l_{nt}$	0.0005 (9.4)
$EDUC_{nt} \times l_{nt}$	0.0504 (0.04)
$EDUC_{nt} \times l_{nt} \times male\ dummy_{nt}$	-0.225 (0.004)
$MS_{nt} \times l_{nt}$	0.198 (0.06)
$MS_{nt} \times SP.EDUC_{nt} \times l_{nt}$	-0.0398 (0.05)
$MS_{nt} \times SP.EDUC_{nt} \times l_{nt} \times male\ dummy_{nt}$	0.0956 (0.04)

Table 7: Nonseparability in Utility of Leisure/Home Production

Variable	Estimate
Lagged Leisure	
l_{nt}^2	-0.214
	(0.002)
$l_{nt} \times l_{nt-1}$	2.423
	(0.004)
$l_{nt} \times l_{nt-1} \times male\ dummy_{nt}$	3.479
	(0.004)
$l_{nt} \times l_{nt-2}$	2.357
	(0.004)
$l_{nt} \times l_{nt-2} \times male\ dummy_{nt}$	-2.575
	(0.004)
Standard deviation	42,553
	(12,376)

Table 8: Decomposition of the Gender Wage Gap
(Median Women Wage over Median Men Wage (%))

Source	Professional	Nonprofessional
Raw	87	76
Predicted	92	81
Explained	95	93
Unexplained	5	7
Decomposition of Explained Gap		
Source	Professional	Nonprofessional
Human Capital[21]	69	74
Beliefs	14	12
Fixed Effect	6	8
Age-education	11	6

[21] Human capital includes the effect of current and past hours on the production function.

Table 9: Decomposition of Change in the Gender Earnings Gap
(Median Women Earnings over Median Men Earnings(%))
1974–1978:1984–1988

Source	Professional	Nonprofessional
Raw	30	24
Predicted	29	22
Human Capital	67	65
Beliefs	8	6
Other	25	29

Table 10: Decomposition of Change in Human Capital as a Source of Gender Earnings Gap
(Median Women Earnings over Median Men Earnings(%))
1974–1978:1984–1988

Source	Professional	Nonprofessional
Hiring Cost	38	32
Private Information	12	13
Demographic	28	34
Home Production Shock	2	3
Production Shock	18	11

Table 11: The Gender Earnings Gap
(Median Women Earnings over Median Men Earnings (%))
1974–1978:1984–1988

Source	Professional		Nonprofessional	
	1974:1978	1984:1988	1974:1978	1984:1988
Raw	48	62	38	47
Predicted	52	67	41	50
Hiring Cost	81	96	68	79
Symmetric Information	71	82	57	65

Table 12: Participation and Occupation Composition (1974–1978:1984–1988)

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

Table 13: Average Hours Worked 1974–1978:1984–1988

Source	Professional				Nonprofessional			
	1974:1978		1984:1988		1974:1978		1984:1988	
	Women	Men	Women	Men	Women	Men	Women	Men
Raw	1640	2201	1904	2226	1424	1998	1635	2117
Hiring Cost	1980	2010	2050	2100	1580	2000	1790	2060
Private Information	1820	2080	1990	2090	1510	1970	1640	1930